Chapter 6

Trust-Region Method for Set Optimization Problems with Set-Valued Mapping Being Finitely Many Vector-Valued Functions

6.1 Introduction

Trust-Region based method is a powerful optimization technique which can be very effective for a variety of optimization tasks, including linear optimization, nonlinear optimization, convex optimization, non-convex optimization etc. The main idea behind this method is to select a region around the current point where we can sufficiently trust the model function's approximation in depicting the objective function. This region is called the trust-region from where the method gets its name. Upon region selection, a step is determined based on minimizing the approximation of the objective function within the trust-region. Then, based on a criterion of how good the minimization was, the computed step is either accepted or rejected. On rejection, the radius of the trust-region is shrinked and a new step is computed until an acceptable step is found.

Trust-Region methods can be very effective for solving nonconvex, ill-conditioned, and noisy problems because of its reliability and robustness. Especially for non-convex problems, these methods has been proven to be very efficient for single-objective as well as multi-objective optimization framework. For example, for single-objective optimization, trust-region based schemes have been studied by Sun et al. [170], Powell [151], Hoseini and Nobakhtiana [88], Friedlander et al. [60], Wang et al. [181], Shi et al. [162] etc. For multi-objective optimization, these methods have been explored by Qu et al. [153], Thomann et al. [175], Villacorta et al. [180], Carrizo et al. [20], Mohammadi et al. [141], etc.

6.2 Motivation

Due to the difficulty of the task, literature on numerical methods for set optimization is limited, containing only Newton method, SetOpt, derivative-free methods, branch-and-bound methods, sorting based methods, and steepest-descent method. These methods, however, suffer from their own drawback. SetOpt [160] is applicable only for polyhedral and convex set-valued maps. Newton method [40] is restricted because of the regularity assumption of the metric. Derivative-free methods [93] ignore the important first and second order derivative information. Sorting-based method [116] is applicable only when the feasible set is finite, and for branch and bound method [47], the problem is needed to be box constrained. These constraints render the two methods inapplicable for the type of unconstrained SOP considered in this thesis. Lastly, steepest descent method [18] may have slow rate of convergence and might get stuck at local minima.

Trust-Region method can overcome these above shortcomings. It does not require any kind of convexity, polyhedral, metric regularity assumptions on set-valued map. It also does not need the feasible set to be finite or the constraint to be bounded by a box. Therefore, based on the success of trust-region methods in single and multi-objective optimization, the trust-region method can be considered to be promising for set optimization as well. However, in the literature, trust-region method for set optimization has not been developed yet. Further, trust-region methods are concerned with finding critical points instead of minimal solutions. Therefore, such a concept for set optimization problems needs to be first defined before designing any trust-region algorithm. However, we find that the notion of critical point that has been defined for multi-objective optimization [21, 30, 55, 133] cannot be directly extended to SOP. Therefore, first introducing the concept of critical point for SOP and then using it to design trust-region method can offer new powerful techniques for set optimization.

6.3 Contributions

In this chapter, we develop a trust-region based method to find critical points of a discrete, non-convex, unconstrained set optimization problem (SOP). For defining the notion of criticality, we employ a vectorization technique to convert the SOP into a family of vector optimization problems using the partition set of weakly minimal elements. We then solve the vector optimization problem by scalarizing it with the help of oriented distance technique. Next, based on the idea of critical points, we present a necessary condition of optimality for either a critical point or weakly minimal solution. For moving from one iterate to next, a VOP is chosen from the family of VOPs and its scalarized form is solved using the trust-region algorithm. Finally, we defined a new reduction rule that is used as the step acceptance criterion.

The main contributions of this chapter can be summarized as follows:

(i) We propose a trust-region based algorithm to minimize a particular type of set optimization problem and show the well-definedness of all the steps of the algorithm.

(ii) We also prove the global convergence property of the algorithm where we show that all the limiting points of the sequence of noncritical iterates generated by the algorithm are critical points.

(iii) Finally, to verify the efficacy of our proposed method, we perform numerical experiments on six type of test problems. For each problem, we aggregate the results by running the experiments for 100 randomly sampled initial points and tabulate different statistics for the number of iterations and execution time.

6.4 Auxiliary concepts to solve set optimization problem

Before entering into main content of this chapter, we first introduce some auxiliary definitions and Lemmas. Also note that, throughout the chapter, for a given $\bar{x} \in \mathbb{R}^n$, we denoted $\bar{\omega} = \omega(\bar{x})$, and a typical element of the partition set $P_{\bar{x}}$ is represented by $\bar{a} = (\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_{\bar{\omega}})$.

Definition 6.1 (Critical point for SOP) *A point $\bar{x} \in \mathbb{R}^n$ is called a critical point of ($\mathcal{SOP}^l_K$) if there does not exist any $\bar{a} \in P_{\bar{x}}$ and $\bar{s} \in \mathbb{R}^n$ satisfying*

$$\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s} \prec_K 0_{m\times 1} \text{ for all } j \in [\omega(\bar{x})],$$

where $0_{m\times 1}$ is the null vector in $\mathbb{R}^m$, and for $j \in [\omega(\bar{x})]$, the notation $\nabla f^{a_j}(\bar{x})^\top s$ represents

$$\nabla f^{a_j}(\bar{x})^\top s = \left(\nabla f^{a_j,1}(\bar{x})^\top s, \nabla f^{a_j,2}(\bar{x})^\top s, \ldots, \nabla f^{a_j,m}(\bar{x})^\top s\right)^\top, s \in \mathbb{R}^n.$$

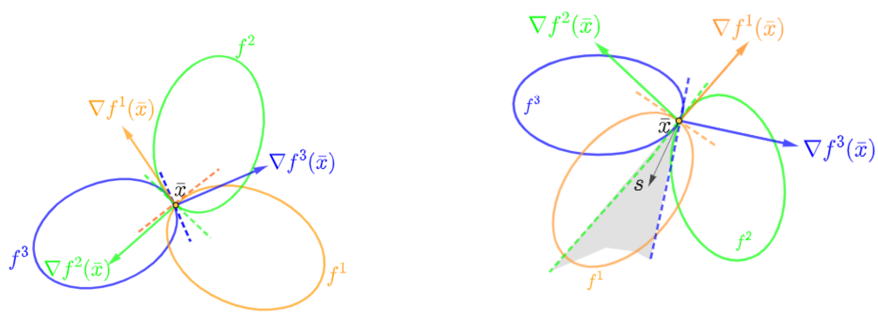
Remark 6.4.1 *A geometrical meaning of critical points for ($\mathcal{SOP}^l_K$) can be understood as follows. Let $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}$ be defined as $F(x) = \{f^1(x), f^2(x), f^3(x)\}$ and $\mathbb{R}$ be equipped with pre-ordering relation $\preceq^l_{\mathbb{R}_+}$. According to Definition 6.1, a point $\bar{x}$ is a critical point of the set optimization problem with F as the objective function if and only if for any $\bar{a} \in P_{\bar{x}}$,*

$$\left\{s \in \mathbb{R}^2 : \nabla f^{\bar{a}_j}(\bar{x})^\top s < 0,\ j \in [\omega(\bar{x})]\right\} = \emptyset.$$

In Figure 6.1(a), the orange, green and blue curves passing through $\bar{x}$ represent the level-curves of f^1, f^2 and f^3, respectively. Let $\omega(\bar{x}) = 3$. From Figure 6.1(a), we see that for any $\bar{a} \in P_{\bar{x}}$,

$$\left\{s \in \mathbb{R}^2 : \nabla f^{\bar{a}_j}(\bar{x})^\top s < 0, j \in [\omega(\bar{x})]\right\} = \cap_{i=1}^{3} \left\{s \in \mathbb{R}^2 : \nabla f^i(\bar{x})^\top s < 0\right\} = \emptyset$$

is an empty set. This indicates that the point $\bar{x}$ shown in Figure 6.1(a) is a critical point of F. However, the point $\bar{x}$ in Figure 6.1(b) is not a critical point because the set $\left\{s \in \mathbb{R}^2 : \nabla f^i(\bar{x})^\top s < 0, i = 1, 2, 3\right\}$, represented by the grey-shaded region, is nonempty.



(a) Critical point $\bar{x}$ of F

(b) Noncritical point $\bar{x}$ of F

Figure 6.1: Geometrical view of critical and noncritical points for a set-valued map $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}$

Definition 6.2 (Descent direction for set-valued maps) *A vector $s \in \mathbb{R}^n$ is called as a descent direction of the set-valued map $F = \{f^i\}_{i \in [p]}$ at $\bar{x}$ if there exists $t_0 > 0$ such that*

$$\{f^i(\bar{x} + ts)\}_{i \in [p]} \prec^l_K \{f^i(\bar{x})\}_{i \in [p]} \text{ for all } t \in (0, t_0].$$

Lemma 6.1 *Let $\bar{x} \in \mathbb{R}^m$, $\bar{\omega} = \omega(\bar{x})$, and $\widetilde{K} \in \mathscr{P}(\mathbb{R}^{m\bar{\omega}})$ be the cone $\Pi_{j=1}^{\bar{\omega}} K$. Suppose $\preceq_{\widetilde{K}}$ and $\prec_{\widetilde{K}}$ be the partial order and the strict order, respectively, in $\mathbb{R}^{m\bar{\omega}}$ induced by $\widetilde{K}$. Furthermore, consider the partition set $P_{\bar{x}}$ associated to $\bar{x}$ and define, for any $a = (a_1, a_2, \ldots, a_{\bar{\omega}}) \in P_{\bar{x}}$, the function $\widetilde{f}^a : \mathbb{R}^n \to \Pi_{j=1}^{\bar{\omega}} \mathbb{R}^m$ by*

$$\widetilde{f}^a(x) = \left(f^{a_1}(x), f^{a_2}(x), \ldots, f^{a_{\bar{\omega}}}(x)\right)^\top. \tag{6.1}$$

Then, $\bar{x}$ is a critical point of ($\mathcal{SOP}^l_K$) if and only if for every $a \in P_{\bar{x}}$, $\bar{x}$ is a critical point of the following vector optimization problem:

$$(\preceq_{\widetilde{K}}) \min_{x \in \mathbb{R}^n} \widetilde{f}^a(x), \tag{$\mathcal{VOP}_a$}$$

Proof: Let $\bar{x}$ be a critical point of ($\mathcal{SOP}^l_K$). On contrary, suppose $\bar{x}$ is not a critical point of ($\mathcal{VOP}_a$) for some $a \in P_{\bar{x}}$. Then, there exists $\bar{s} \in \mathbb{R}^n$ such that for all $j \in [\bar{\omega}]$,

$$\nabla f^{a_j}(\bar{x})^\top \bar{s} \in -\text{int}(K)$$
$$\text{i.e., } \nabla f^{a_j}(\bar{x})^\top \bar{s} \prec_K 0_{m\times 1}, \tag{6.2}$$

which contradicts the fact that $\bar{x}$ is a critical point of ($\mathcal{SOP}^l_K$). So, $\bar{x}$ is a critical point of ($\mathcal{VOP}_a$) for every $a \in P_{\bar{x}}$.

Conversely, assume that $\bar{x}$ is a critical point of ($\mathcal{VOP}_a$) for every $a \in P_{\bar{x}}$. If possible, let $\bar{x}$ is not a critical point of ($\mathcal{SOP}^l_K$). Then, there exists $\bar{a} \in P_{\bar{x}}$ and $\bar{s} \in \mathbb{R}^n$ such that

$$\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s} \prec_K 0_{m\times 1} \text{ for all } j \in [\omega(\bar{x})],$$

which implies that $\bar{x}$ is a not a critical point of ($\mathcal{VOP}_a$) with $a = \bar{a}$. This is contradictory to the assumption, and hence the result follows. □

As per Lemma 6.1, to perceive the critical points of ($\mathcal{SOP}^l_K$), one can necessarily focus to identify critical points of ($\mathcal{VOP}_a$) for each $a \in P_{\bar{x}}$. However, it is important to note that ($\mathcal{VOP}_a$) slightly differs from the conventional vector optimization problem because each component of $\widetilde{f}^a$ is vector-valued, making the gradients of each component f^{a_j} of $\widetilde{f}^a$ Jacobian matrix, which causes complexity that challenges the straightforward identification of critical points for ($\mathcal{VOP}_a$) as conventionally defined. To address this issue, we employ a scalarization approach using oriented distance function, that can facilitate to determine if a point is critical for ($\mathcal{VOP}_a$) by transforming a vector into a scalar value.

We recall the properties of the oriented distance function Δ_{-K} in characterizing critical points and descent directions of the vector-valued map $\widetilde{f}^a$ in ($\mathcal{VOP}_a$).

Definition 6.3 [6] *Let $\mathcal{A} \in \mathscr{P}(\mathbb{R}^m)$. A function $\Delta_{\mathcal{A}} : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\} =: \overline{\mathbb{R}}$ defined by*

$$\Delta_{\mathcal{A}}(y) := d_{\mathcal{A}}(y) - d_{\mathcal{A}^c}(y) \text{ for all } y \in \mathbb{R}^m,$$

is called the oriented distance function, where $d_{\mathcal{A}}(y) := \inf_{a\in\mathcal{A}} \|y - a\|$. By convention, for any $y \in \mathbb{R}^m$, $d_{\emptyset}(y) = +\infty$, and hence, $\Delta_{\emptyset}(y) = +\infty$, whereas $\Delta_{\mathbb{R}^m}(y) = -\infty$.

Lemma 6.4.1 ([6, 192])

(i) Δ_{-K} is Lipschitz continuous with Lipschitz constant 1;

(ii) $\Delta_{-K}(y) < 0$ *if and only if* $y \in int(-K)$;

(iii) $\Delta_{-K}(y) = 0$ *if and only if* $y \in bd(-K)$;

(iv) $\Delta_{-K}(y) > 0$ *if and only if* $y \in int[(-K)^c]$;

(v) $\Delta_{-K}(y + z) = \Delta_{-K-z}(y)$ *for all* $y, z \in \mathbb{R}^m$;

(vi) $\Delta_{-K}(\lambda y) = \lambda \Delta_{\lambda^{-1}(-K)}(y)$ *for all* $y \in \mathbb{R}^m$ *and* $\lambda > 0$;

(vii) $\Delta_{-K}(y_1 + y_2) \leq \Delta_{-K}(y_1) + \Delta_{-K}(y_2)$ *and* $\Delta_{-K}(y_1) - \Delta_{-K}(y_2) \leq \Delta_{-K}(y_1 - y_2)$ *for all* $y_1, y_2 \in \mathbb{R}^m$.

(viii) *Given* $y_1, y_2 \in \mathbb{R}^m$, *if* $y_1 \prec y_2$ *(resp.,* $y_1 \preceq y_2$*), then* $\Delta_{-K}(y_1) < \Delta_{-K}(y_2)$ *(resp.,* $\Delta_{-K}(y_1) \leq \Delta_{-K}(y_2)$*).*

With the support of Δ_{-K} and its properties, the concept of critical point for ($\mathcal{VOP}_a$) is given as follows.

Definition 6.4 (Critical point for ($\mathcal{VOP}_a$)) *A point* $\bar{x} \in \mathbb{R}^n$ *is said to be a critical point for (*$\mathcal{VOP}_a$*) if for any* $s \in \mathbb{R}^n$ *there exists* $j_0 \in \{1, 2, 3, \ldots, \bar{\omega}\}$ *such that* $\Delta_{-K}(\nabla f^{a_{j_0}}(\bar{x})^\top s) \geq 0$, *i.e.,* $\nabla f^{a_{j_0}}(\bar{x})^\top s \notin -int(K)$.

Definition 6.5 (Descent direction [57]) *A vector* $s \in \mathbb{R}^n$ *is said to be a descent direction of the objective function* $\widetilde{f}^a : \mathbb{R}^n \to \mathbb{R}^{m\bar{\omega}}$ *of (*$\mathcal{VOP}_a$*) at* $\bar{x}$ *if there exists* $t_0 > 0$ *such that*

$$\widetilde{f}^a(\bar{x} + ts) \prec_{\widetilde{K}} \widetilde{f}^a(\bar{x}) \text{ for all } t \in (0, t_0],$$

i.e., if there exists $t_0 > 0$ *such that for all* $j \in [\bar{\omega}]$,

$$f^{a_j}(\bar{x} + ts) \prec_K f^{a_j}(\bar{x}) \text{ for all } t \in (0, t_0],$$

i.e., if $\nabla f^{a_j}(\bar{x})^\top s \in -int(K)$ *for all* $j \in [\bar{\omega}]$.

The following result indicates that a local weakly-minimal point of an SOP is a critical point, and Furthermore, assuming convexity of the objective function, critical points and weakly-minimal points coincide.

Theorem 6.1 *(i) If* $\bar{x}$ *is a local weak minimal point of (*$\mathcal{SOP}^l_K$*), then* $\bar{x}$ *is a critical point for (*$\mathcal{SOP}^l_K$*).*

(ii) If each f^i, $i \in [p]$, is K-convex and $\bar{x}$ is a critical point for ($\mathcal{SOP}^l_K$), then $\bar{x}$ is a weakly-minimal point of ($\mathcal{SOP}^l_K$).

Proof: (i) On contrary, let $\bar{x}$ is not a critical point of ($\mathcal{SOP}^l_K$). By Lemma 6.1, there exists $\bar{a} = (\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_{\bar{\omega}})^\top \in P_{\bar{x}}$ such that $\bar{x}$ is not a critical point of ($\mathcal{VOP}_{\bar{a}}$). Therefore, there exists $\bar{s} \in \mathbb{R}^n$ such that

$$\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s} \in -\text{int}(K) \text{ for all } j \in [\bar{\omega}]. \tag{6.3}$$

Since $f^{\bar{a}_j} \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ for all $j \in [\bar{\omega}]$, we have

$$f^{\bar{a}_j}(x) = f^{\bar{a}_j}(\bar{x}) + \nabla f^{\bar{a}_j}(\bar{x})^\top (x - \bar{x}) + o(\|x - \bar{x}\|), \tag{6.4}$$

where $\lim\limits_{x \to \bar{x}} \frac{o(\|x-\bar{x}\|)}{\|x-\bar{x}\|} = 0$. Since $\bar{x}$ is a locally weak minimal element of ($\mathcal{SOP}^l_K$), there exists a neighbourhood $\mathcal{N}(\bar{x}, \delta)$ of $\bar{x}$ such that

$$\nexists\ x \in \mathcal{N}(\bar{x}, \delta) \text{ with } \{f^i(x)\}_{i \in [p]} \prec^l_K \{f^i(\bar{x})\}_{i \in [p]}.$$

Thus, by Lemma 3.1 in [18], $\bar{x}$ is a local weakly-minimal point of ($\mathcal{VOP}_a$) with $a = \bar{a}$. This implies, $\nexists\ x \in \mathcal{N}(\bar{x}, \delta)$ with $f^{\bar{a}_j}(x) \prec_K f^{\bar{a}_j}(\bar{x})$ for all $j \in [\bar{\omega}]$, i.e., $f^{\bar{a}_j}(x) - f^{\bar{a}_j}(\bar{x}) \notin -\text{int}(K)$ for all $j \in [\bar{\omega}]$.

From (6.4), $\exists\ \bar{\delta} < \delta$ such that for all $j \in [\bar{\omega}]$,

$$\nabla f^{\bar{a}_j}(\bar{x})^\top (x - \bar{x}) \notin -\text{int}(K) \ \ \forall\ x \in \mathcal{N}(\bar{x}, \bar{\delta}). \tag{6.5}$$

Take any $x' \in \mathcal{N}(\bar{x}, \bar{\delta})$ such that $x' - \bar{x} = \bar{s}$. Then, from (6.5), it implies that $\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s} \notin -\text{int}(K)$ for all $j \in [\bar{\omega}]$. This is contradictory to (6.3). Therefore, $\bar{x}$ is a critical point for ($\mathcal{SOP}^l_K$).

(ii) By the given hypothesis, it follows that for any $\bar{a} \in P_{\bar{x}}$, for all $j \in [\bar{\omega}], x \in \mathbb{R}^n, \mu \in (0, 1]$,

$$\begin{aligned} &\mu f^{\bar{a}_j}(x) + (1-\mu) f^{\bar{a}_j}(\bar{x}) - f^{\bar{a}_j}(\mu x + (1-\mu)\bar{x}) \in K \\ \implies &(f^{\bar{a}_j}(x) - f^{\bar{a}_j}(\bar{x})) - \tfrac{1}{\mu}\left(f^{\bar{a}_j}(\bar{x} + \mu(x - \bar{x})) - f^{\bar{a}_j}(\bar{x})\right) \in K. \end{aligned}$$

As $\mu \to 0^+$, we obtain for all $j \in [\bar{\omega}]$ and $x \in \mathbb{R}^n$ that

$$(f^{\bar{a}_j}(x) - f^{\bar{a}_j}(\bar{x})) - \nabla f^{\bar{a}_j}(\bar{x})(x - \bar{x}) \in K$$

$$\implies \nabla f^{\bar{a}_j}(\bar{x})(x-\bar{x}) \preceq_K (f^{\bar{a}_j}(x) - f^{\bar{a}_j}(\bar{x})) \tag{6.6}$$

$$\implies \Delta_{-K}(\nabla f^{\bar{a}_j}(\bar{x})(x-\bar{x})) \leq \Delta_{-K}(f^{\bar{a}_j}(x) - f^{\bar{a}_j}(\bar{x})). \tag{6.7}$$

Since $\bar{x}$ is a critical point for $(\mathcal{SOP}_K^l)$, it follows from Lemma 6.1 that $\bar{x}$ is critical for $(\mathcal{VOP}_a)$. Hence, for any $s \in \mathbb{R}^n$, there exists $j_0 \in [\bar{\omega}]$ such that $\Delta_{-K}(\nabla f^{\bar{a}_{j_0}}(\bar{x})^\top s) \geq 0$. In particular, taking $s = x - \bar{x}$, we have for any $x \in \mathbb{R}^n$ that

$$\begin{aligned} & 0 \leq \Delta_{-K}(\nabla f^{\bar{a}_{j_0}}(\bar{x})(x-\bar{x})) \leq \Delta_{-K}(f^{\bar{a}_{j_0}}(x) - f^{\bar{a}_{j_0}}(\bar{x})) \\ \implies & f^{\bar{a}_{j_0}}(x) - f^{\bar{a}_{j_0}}(\bar{x}) \notin -\text{int}(K) \text{ by Lemma 6.4.1} \\ \implies & f^{\bar{a}_{j_0}}(x) \not\prec_K f^{\bar{a}_{j_0}}(\bar{x}). \end{aligned}$$

Therefore, there exists $j_0 \in [\bar{\omega}]$ for which there does not exist any $x \in \mathbb{R}^n$ such that $f^{a_{j_0}}(x) \prec_K f^{a_{j_0}}(\bar{x})$. This implies $\bar{x}$ is a weakly-minimal element of $(\mathcal{VOP}_{\bar{a}})$ for all $\bar{a} \in P_{\bar{x}}$. Hence, using Lemma 3.1 in [18], $\bar{x}$ is weak minimal element for $(\mathcal{SOP}_K^l)$.

□

6.5 Trust-Region method for set optimization

Based on the above definitions and lemmas, in this section, we derive the trust-region scheme to obtain the critical points of $(\mathcal{SOP}_K^l)$.

As per Lemma 6.1, any critical point of $(\mathcal{SOP}_K^l)$ is a critical point of $(\mathcal{VOP}_a)$ and vice-versa. Therefore, one may think that the set optimization problem $(\mathcal{SOP}_K^l)$ with the objective function represented by finitely many vector-valued functions is just a vector optimization problem, and thus, all the classical methods for vector optimization problems can be straightly implemented to solve $(\mathcal{SOP}_K^l)$. However, this is not true. The argument for this is as follows. Note that the formulation of $(\mathcal{VOP}_a)$ based on the partition set $P_{\bar{x}}$ at the point $\bar{x}$. If we aim to approach a critical point $x^\star$ of $(\mathcal{SOP}_K^l)$ through an iterative sequence $\{x_k\}$ that converges to $x^\star$, then note that at each iterative point x_k, the formulation of $(\mathcal{VOP}_a)$ varies according to changing the partition set P_{x_k} across the iterates. Neatly, the vector optimization problem analogous to $(\mathcal{VOP}_a)$ at x_k is the following problem:

$$(\preceq_{\widetilde{K}}) \quad \min_{x \in \mathbb{R}^n} \widetilde{f}^{a^k}(x), \; a^k \in P_k, \tag{$\mathcal{VOP}_{a^k}(x_k)$}$$

and commonly, the problem $(\mathcal{VOP}_{a^k}(x_k))$ at x_k differs from that at x_{k+1}. Therefore, solving ($\mathcal{SOP}^l_K$) does not simply means solving just one vector optimization problem ($\mathcal{VOP}_a$) for all $a \in P_{x^\star}$.

However, one may perceive from the observation that 'if $\bar{x}$ is not a critical point of ($\mathcal{VOP}_a$) for at least one $a \in P_{\bar{x}}$, then $\bar{x}$ is not a critical point of ($\mathcal{SOP}^l_K$),' that the computational effort in computing $P_{\bar{x}}$ and then solving ($\mathcal{VOP}_a$) at a noncritical point is ineffective for identifying critical points of ($\mathcal{SOP}^l_K$). This assumption is incorrect. In the following, we explore how the formulation of ($\mathcal{VOP}_a$) can be leveraged to address *a sequence of vector optimization problems* ($\mathcal{VOP}_{a^k}(x_k)$)'s at noncritical points x_k's, solving which we can arrive at a critical point of ($\mathcal{SOP}^l_K$).

We initiate the process from any arbitrarily chosen initial point x_0. Through the trust-region iterative scheme outline below, we aim to generate a sequence $\{x_k\}$ that converges to a critical point of ($\mathcal{SOP}^l_K$). Below, we describe how we proceed from the k-th iterate to the next iterate x_{k+1}. For this, we tactfully find a direction of movement s_k and then take the next iterate x_{k+1} as $x_k + s_k$. The process of searching a direction s_k at the k-th iterate is the following. At first, at the current iterate x_k, we generate the partition set P_{x_k}. Then, we chose a specific a^k from P_{x_k} and formulate the corresponding ($\mathcal{VOP}_{a^k}(x_k)$); a descent direction s_k of ($\mathcal{VOP}_{a^k}(x_k)$) happens to be a descent direction of ($\mathcal{SOP}^l_K$) at x_k (Theorem 6.2). The process of choosing such an a^k from P_{x_k} is described below, which is based on identifying a necessary optimality condition of a weakly-minimal element of ($\mathcal{SOP}^l_K$). Throughout the rest of the paper, we denote $P_k := P_{x_k}$ and $\omega_k := \omega(x_k) = |P_k|$.

6.5.1 Choice of 'a^k' from P_k

Before choosing an a from P_k, we present a necessary condition for weak minimal solution of ($\mathcal{SOP}^l_K$). From Theorem 6.1(i), we observe that a weak minimal solution of ($\mathcal{SOP}^l_K$) is a critical point of ($\mathcal{SOP}^l_K$). From Definition 6.1 and 6.4, we see that

$$\begin{aligned} &\text{a point } x_k \text{ is a critical point of } (\mathcal{SOP}^l_K) \\ \iff &\text{ for any } a \in P_k \text{ and } s \in \mathbb{R}^n,\ \exists j_0 \in [\omega_k] \text{ such that } \Delta_{-K}(\nabla f^{a_{j_0}}(x_k)^\top s) \geq 0. \end{aligned}$$

So, at a critical point x_k, for any $a \in P_k$ and $s \in \mathbb{R}^n$, there exists $j_0 \in [\omega_k]$ such that

$$\max_{j\in[\omega_k]} \left\{\Delta_{-K}(\nabla f^{a_j}(x_k)^\top s + \tfrac{1}{2}s^\top \nabla^2 f^{a_j}(x_k)s), \Delta_{-K}(\nabla f^{a_j}(x_k)^\top s)\right\} \geq \Delta_{-K}(\nabla f^{a_{j_0}}(x_k)^\top s) \geq 0. \tag{6.8}$$

Let $\Omega_{\max}$ be the maximum allowed trust-region step-size across the iterates x_k' s. Denote $\mathcal{B} := \{s \in \mathbb{R}^n : \|s\| \leq \Omega_{\max}\}$. Then, for any given $x \in \mathbb{R}^n$, if we introduce a function

$\Theta_x : P_x \times \mathcal{B} \to \mathbb{R}$ by

$$\Theta_x(a, s) := \max_{j\in[\omega(x)]} \left\{\Delta_{-K}\left(\nabla f^{a_j}(x)^\top s + \tfrac{1}{2} s^\top \nabla^2 f^{a_j}(x) s\right), \Delta_{-K}(\nabla f^{a_j}(x)^\top s)\right\}, a \in P_x, s \in \mathcal{B}, \tag{6.9}$$

then by (6.8), at a critical point x_k, we have

$$\begin{aligned} &\Theta_{x_k}(a, s) \geq 0 \;\forall\; a \in P_k \text{ and } s \in \mathcal{B} \\ \Longrightarrow\; &\inf_{s\in\mathcal{B}} \Theta_{x_k}(a, s) \geq 0 \;\forall\; a \in P_k \\ \Longrightarrow\; &\forall a \in P_k : 0 \leq \inf_{s\in\mathcal{B}} \Theta_{x_k}(a, s) \leq \Theta_{x_k}(a, 0) = 0 \\ \Longrightarrow\; &\forall a \in P_k : \min_{s\in\mathcal{B}} \Theta_{x_k}(a, s) = 0 = \Theta_{x_k}(a, 0). \end{aligned} \tag{6.10}$$

As $\mathcal{B}$ is compact, and for any $x \in \mathbb{R}^n$, the partition set P_x is finite, Θ_x attains its minimum over the set $P_x \times \mathcal{B}$. Let us define a function $\theta : \mathbb{R}^n \to \mathbb{R}$ by

$$\theta(x) = \min_{(a,s)\in P_x\times\mathcal{B}} \Theta_x(a, s). \tag{6.11}$$

Then, in view of (6.10), at a critical point x_k of ($\mathcal{SOP}^l_K$), if for $(a^k, s_k) \in P_k \times \mathcal{B}$ we have $\theta(x_k) = \Theta_{x_k}(a^k, s_k)$, then

$$\theta(x_k) = 0, \tag{6.12}$$

where $\mathcal{B}_k := \{x \in \mathbb{R}^n : \|x - x_k\| \leq \Omega_k\}$ is a region centered at x_k with $\Omega_k > 0$ as the radius. By assembling everything, we obtain the following result:

Proposition 6.5.1 (Necessary condition for weakly-minimal points). *If x_k is either a weak minimal point or a critical point of ($\mathcal{SOP}^l_K$), and $a^k \in P_k$ and $s_k \in \mathcal{B}$ be such that $\theta(x_k) = \Theta_{x_k}(a^k, s_k)$, where Θ_x and θ are as defined in (6.9) and (6.11), respectively, then $\theta(x_k) = 0$.*

As a consequence of Proposition 6.5.1 and discussion in this Subsection 6.5.1, if we can select an $a^k = \left(a^k_1, a^k_2, \ldots, a^k_{\omega_k}\right)$ from the partition P_k of the current iterate x_k such that $\theta(x_k) = \Theta_{x_k}(a^k, s_k)$, i.e.,

$$(a^k, s_k) \in \operatorname*{argmin}_{(a,s)\in P_k\times\mathcal{B}_k} \max_{j\in[\omega_k]} \left\{\Delta_{-K}\left(\nabla f^{a_j}(x_k)^\top s + \tfrac{1}{2} s^\top \nabla^2 f^{a_j}(x_k) s\right), \Delta_{-K}\left(\nabla f^{a_j}(x_k)^\top s\right)\right\}, \tag{6.13}$$

then $\theta(x_k) \neq 0$ implies x_k is a noncritical point. Next, we employ a trust-region scheme to solve ($\mathcal{VOP}_{a^k}(x_k)$) for the particular $a^k \in P_k$ which satisfies (6.13). Thereby, we get a trust-region step s_k to progress to the next step x_{k+1} by $x_k + s_k$.

6.5.2 Choice of the trust-region step 's_k'

Once an a^k is chosen that satisfies (6.13), we attempt to find a descent direction, if it exists, at x_k of the objective function of ($\mathcal{VOP}_{a^k}(x_k)$) in a region what we call a trust-region at the current iterate x_k; later, in the next subsection, we prove that (see Theorem 6.2) a descent direction of the objective function of $\mathcal{VOP}_{a^k}(x_k)$ is a descent direction of the objective function of ($\mathcal{SOP}^l_K$). A trust-region at x_k for the objective function of ($\mathcal{VOP}_{a^k}(x_k)$) is a region where a representative model function of the objective function of ($\mathcal{VOP}_{a^k}(x_k)$) is a good approximation of it. Let the trust-region at x_k be

$$\mathcal{B}_k = \{x \in \mathbb{R}^n : \|x - x_k\| \leq \Omega_k\},$$

where $\Omega_k > 0$; the value of Ω_k is called the trust-region radius at x_k, and x_k is called the center of the trust-region. By the transformation $s = x - x_k$, the trust-region $\mathcal{B}_k$ can be presented by $\mathcal{B}_k := \{s \in \mathbb{R}^n : \|s\| \leq \Omega_k\}$.

We aim to choose an s_k from $\mathcal{B}_k$ such that (a^k, s_k) satisfies (6.13). If with this (a^k, s_k) we get $\Theta_{x_k}(a^k, s_k) = 0$, then we have reached a point x_k at which a necessary condition for critical points of ($\mathcal{SOP}^l_K$) is satisfied, and we stop the process. If, however, $\Theta_{x_k}(a^k, s_k) \neq 0$, then x_k is a noncritical point of ($\mathcal{SOP}^l_K$), and in this case, we revise the trust-region radius and the current iterate as described in the next subsections.

For a noncritical point x_k of ($\mathcal{SOP}^l_K$), to progress the iterate to the next iterate, we describe a process to find an s_k from $\mathcal{B}_k$ for which (a^k, s_k) satisfies (6.13). Such an s_k is identified by solving a vector optimization problem whose objective is a model function of the objective function of ($\mathcal{VOP}_{a^k}(x_k)$). At x_k, the model function of the objective function $\widetilde{f}^{a^k} = (f^{a^k_1}, f^{a^k_2}, \ldots, f^{a^k_{\omega_k}})^\top$ of ($\mathcal{VOP}_{a^k}(x_k)$) inside $\mathcal{B}_k$ is taken as the following quadratic approximation $\widetilde{m}^{a^k} = (m^{a^k_1}, m^{a^k_2}, \ldots, m^{a^k_{\omega_k}})^\top$:

$$m_k^{a^k_j}(s) = \nabla f^{a^k_j}(x_k)^\top s + \tfrac{1}{2} s^\top \nabla^2 f^{a^k_j}(x_k) s, \;\; s \in \mathcal{B}_k, \quad \text{for each } j \in [\omega_k], \tag{6.14}$$

where

$$\nabla f^{a^k_j}(x_k)^\top s = \left(\nabla f^{a^k_j,1}(x_k)^\top s, \nabla f^{a^k_j,2}(x_k)^\top s, \, \ldots, \nabla f^{a^k_j,m}(x_k)^\top s\right)^\top$$

and

$$s^\top \nabla^2 f^{a^k_j}(x_k) s = \left(s^\top \nabla^2 f^{a^k_j,1}(x_k) s, s^\top \nabla^2 f^{a^k_j,1}(x_k) s, \, \ldots, s^\top \nabla^2 f^{a^k_j,m}(x_k) s\right)^\top.$$

To generate a search direction s_k in $\mathcal{B}_k$, similar to the conventional trust-region scheme, we solve the following vector problem:

$$\min_{s \in \mathcal{B}_k} \max_{j \in [\omega_k]} m_k^{a^k_j}(s). \tag{6.15}$$

To find a weakly-minimal solution of the problem 6.15, we find a solution of the following scalar representation of (6.15) with the help of the oriented distance nonlinear scalarizing function:

$$\min_{s\in\mathcal{B}_k}\max_{j\in[\omega_k]}\left[\Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s+\tfrac{1}{2}s^\top\nabla^2 f^{a_j^k}(x_k)s\right)\right]. \tag{6.16}$$

Lemma 6.5.1 *Any solution $s_k\in\mathcal{B}_k$ of the problem (6.16) is a weakly-minimal point of the vector-valued function $\widetilde{m}^{a^k}(s)=\left(m_k^{a_1^k}(s),m_k^{a_2^k}(s),\ldots,m_k^{a_{\omega_k}^k}(s)\right)^\top$ in $\mathcal{B}_k$.*

Proof: Let s_k be an optimal solution of the problem (6.16). Then for all $s\in\mathcal{B}_k$,

$$\begin{aligned}
&\max_{j\in[\omega_k]}\Delta_{-K}(m^{a_j^k}(s_k))\le\max_{j\in[\omega_k]}\Delta_{-K}(m^{a_j^k}(s))\\
\Longrightarrow\ &0\le\max_{j\in[\omega_k]}\Delta_{-K}(m^{a_j^k}(s))-\max_{j\in[\omega_k]}\Delta_{-K}(m^{a_{kj}}(s_k))\le\max_{j\in[\omega_k]}\left(\Delta_{-K}(m^{a_j^k}(s))-\Delta_{-K}(m^{a_j^k}(s_k))\right)\\
\Longrightarrow\ &0\le\Delta_{-K}(m^{a_j^k}(s))-\Delta_{-K}(m^{a_j^k}(s_k))\text{ for some } j\in[\omega_k]\\
\Longrightarrow\ &0\le\Delta_{-K}\left(m^{a_j^k}(s)-m^{a_j^k}(s_k)\right)\text{ for some } j\in[\omega_k],\ \text{by Lemma 6.4.1 (vii)}\\
\Longrightarrow\ &m^{a_j^k}(s)-m^{a_j^k}(s_k)\notin-\text{int}(K)\text{ for some } j\in[\omega_k],\ \text{by Lemma 6.4.1 (ii)}\\
\Longrightarrow\ &m^{a_j^k}(s)\not\prec_K m^{a_j^k}(s_k)\text{ for some } j\in[\omega_k].
\end{aligned}$$

Then, there exists no $s\in\mathcal{B}_k$ such that $m^{a_j^k}(s)\prec_K m^{a_j^k}(s_k)$ for all $j\in[\omega_k]$. Therefore, s_k is a weakly-minimal point of $\widetilde{m}^{a^k}$ in $\mathcal{B}_k$. □

Note 6.1 *It is to be mentioned that if $K=\mathbb{R}^m_+$, then $\Delta_{-K}(y)=\max_{1\le i\le m}y_i$, where $y=(y_1,y_2,\ldots,y_m)\in\mathbb{R}^m$. Thus, for $K=\mathbb{R}^m_+$, the subproblem (6.15) reduces to the problem (3.5) in [55], which finds a point along the Newton direction s_k^N inside $\mathcal{B}_k$. As Newton method admits a fast rate of convergence, if we choose s_k by solving (6.15), we expect to have a fast convergence rate of the proposed method.*

As the objective function of the problem (6.16) is nonsmooth, we recast it by the following problem:

$$\left.\begin{aligned}
&\min && t\\
&\text{subject to} && \Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s+\tfrac{1}{2}s^\top\nabla^2 f^{a_j^k}(x_k)s\right)-t\le 0,\ j=1,2,\ldots,\omega_k,\\
& && \|s\|\le\Omega_k.
\end{aligned}\right\} \tag{6.17}$$

However, as explained in [20, Example 1], to overcome the difficulty in finding a descent direction for a nonconvex problem at the current iterate x_k, we add a sublinear inequality in the constraint set of (6.17) as below:

$$\left.\begin{array}{ll} \min & t \\ \text{subject to} & \Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s + \frac{1}{2} s^\top \nabla^2 f^{a_j^k}(x_k) s\right) - t \le 0,\ j = 1, 2, \ldots, \omega_k, \\ & \Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s\right) - t \le 0,\ j = 1, 2, \ldots, \omega_k, \\ & \|s\| \le \Omega_k. \end{array}\right\} \tag{6.18}$$

Observe that the problem (6.18) can be compactly rewritten as

$$\min_{s \in \mathcal{B}_k} \max_{j \in [\omega_k]} \left[\Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s + \frac{1}{2} s^\top \nabla^2 f^{a_j^k}(x_k) s\right),\ \Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s)\right]. \tag{6.19}$$

Although to find a suitable s_k, we solve the problem (6.18), not the equivalent non-smooth problem (6.19), we have rewritten the problem (6.18) into (6.19) because the expression of the objective function of (6.19) will be used later (in Subsection 6.5.5) to find a stopping condition. Also, it is noteworthy that the objective function of the problem (6.19) is the objective function in (6.13). Hence, any s_k that is found by solving the problem (6.18) essentially satisfies (6.13). However, it is not always the case that a solution of (6.18) is a descent direction of the objective function F of ($\mathcal{SOP}^l_K$) at x_k. If or not a solution s_k of (6.18) is a descent direction of F at x_k is essentially dependent on how good is the model function $\widetilde{m}^{a^k}$ to represent the objective function $\widetilde{f}^{a^k}$ inside the region $\mathcal{B}_k$. The goodness of the model function $\widetilde{m}^{a^k}$ is identified by a reduction ratio, as explained in the next Subsection 6.5.3. It is shown below (see Propositions 6.5.2 and 6.2) that a positive value of the reduction ratio indicates that a solution s_k of is a descent direction of $F : \mathcal{B}_k \rightrightarrows \mathbb{R}^m$ at x_k, and a significant threshold positive value (η_1) of the reduction ratio indicates $m_k^{a_j^k}$ a good model function of $f_k^{a_j^k}$ in $\mathcal{B}_k$ (see the equation (6.27)).

Throughout the rest of the paper, we call the vector s_k in $\mathcal{B}_k$, which is a solution to (6.18), as the trust-region step at x_k. If for a trust-region step s_k the value of the function $F : \mathcal{B}_k \rightrightarrows \mathbb{R}^m$ gets significant decrement by the movement from x_k to $x_k + s_k$, then we call the current iteration a successful iteration and thus update the current iterate x_k by $x_k + s_k$. Otherwise, we keep hold of the iterate at x_k and revise the trust-region radius as explained in Subsection 6.5.4 and reevaluate a trust-region step in the revised trust-region.

6.5.3 Calculation of reduction ratio

Once a step s_k is identified, by solving (6.18), from the trust-region $\mathcal{B}_k$, we aim to check if the direction s_k is a descent direction for $F : \mathcal{B}_k \rightrightarrows \mathbb{R}^m$ at x_k; if descent, then we may update the iterate by $x_{k+1} = x_k + s_k$. To check if the identified s_k is a descent direction, we evaluate a measurement of the movement of values of component functions in $\widetilde{f}^{a^k}$ for the movement of the argument point from x_k to $x_k + s_k$, which is executed by $-\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k))$. Note that if $-\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k)) > 0$, then

$$0 < -\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k)) \leq \Delta_{-K}(f^{a_j^k}(x_k)) - \Delta_{-K}(f^{a_j^k}(x_k + s_k)),$$

i.e., $f^{a_j^k}(x_k + s_k) \prec_K f^{a_j^k}(x_k)$, and hence $s_k \in \mathcal{B}_k$ is a step along a descent direction of $f^{a_j^k}$ at x_k. Thus, we call the value of $-\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k))$ as a measurement of the actual reduction of the vector-valued function $f^{a_j^k}$ due to the step s_k.

A predicted value of the reduction of $f^{a_j^k}$ due to the step s_k at x_k is measured by a measurement of the movement of the function value of its model function $m^{a_j^k}$, which we measure by the positive value $\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))$. The reason behind the value $\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))$ being positive is given in the proof of the following Corollary 6.6.1 (see (6.47)).

With the values of the measurements of the actual reduction and predicted reduction, we define a reduction ratio $\rho_k^{a_j^k}$ as follows:

$$\sigma_k^{a_j^k} = \frac{\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k))}{\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))} \text{ and } \rho_k^{a_j^k} = \frac{\text{actual reduction}}{\text{predicted reduction}} = -\sigma_k^{a_j^k} \text{ for all }$$

$j \in [\omega_k]$.

Proposition 6.5.2 *Let x_k be a noncritical point of ($\mathcal{SOP}^l_K$). Then, for any $j \in [\omega_k]$, a trust-region step s_k satisfies $f_k^{a_j^k}(x_k + s_k) \prec_K f_k^{a_j^k}(x_k)$ if and only if $\rho_k^{a_j^k} > 0$.*

Proof: Let s_k satisfy $f_k^{a_j^k}(x_k + s_k) \prec_K f_k^{a_j^k}(x_k)$. Then, $\Delta_{-K}(f_k^{a_j^k}(x_k + s_k) - f_k^{a_j^k}(x_k)) < 0$. As s_k is identified by solving (6.18), we get

$$\begin{aligned} \max_{j \in [\omega_k]} \Delta_{-K}(m_k^{a_j^k}(s_k)) &\leq \max_{j \in [\omega_k]} \left\{ \Delta_{-K}(m_k^{a_j^k}(s_k)),\ \Delta_{-K}(\nabla f_k^{a_j^k}(x_k)^\top s_k) \right\} \\ &\leq \max_{j \in [\omega_k]} \left\{ \Delta_{-K}(m_k^{a_j^k}(0)),\ \Delta_{-K}(\nabla f_k^{a_j^k}(x_k)^\top 0) \right\} = 0, \end{aligned}$$

i.e., $\Delta_{-K}(m_k^{a_j^k}(s_k)) \leq 0$. Thus,

$$\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k)) \geq \Delta_{-K}(m_k^{a_j^k}(0)) - \Delta_{-K}(m_k^{a_j^k}(s_k)) \geq 0. \tag{6.20}$$

In fact, for a noncritical point x_k, we prove in Corollary 6.6.1 (see (6.47)) that $\Delta_{-K}(m_k^{a_j^k}(0)-$

$m_k^{a_j^k}(s_k)) > 0$. Thus,

$$\rho_k^{a_j^k} = \frac{-\Delta_{-K}(f_k^{a_j^k}(x_k+s_k) - f_k^{a_j^k}(x_k))}{\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k))} > 0.$$

Conversely, suppose $\rho_k^{a_j^k} > 0$. Then, $\Delta_{-K}(f_k^{a_j^k}(x_k+s_k) - f_k^{a_j^k}(x_k)) < 0$, and hence $f_k^{a_j^k}(x_k+s_k) \prec_K f_k^{a_j^k}(x_k)$. □

Note 6.2 *If we take* $\rho_k^{a_j^k}$ *as*

$$\rho_k^{a_j^k} = \frac{\Delta_{-K}(f_k^{a_j^k}(x_k) - f_k^{a_j^k}(x_k+s_k))}{\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))} \text{ for all } j \in [\omega_k], \tag{6.21}$$

instead of $\rho_k^{a_j^k} = -\sigma_k^{a_j^k}$*, then* $\rho_k^{a_j^k}$ *cannot be a correct choice of the reduction ratio. This is due to the fact that if we consider the formula (6.21), then for a given* $\eta_1 \in (0,1)$*,* $\rho_k^{a_j^k} \geq \eta_1$ *may not imply that* s_k *is a decent direction for each of* $f_k^{a_j^k}$*'s,* $j \in [\omega_k]$*. To realize this, let us take the following simple instance:*

$$m = 2, n = 1, p = 1, K = -\mathbb{R}_+^2, F_k(x) = \{f_k^1(x)\},$$

where $f_k^1(x) = (f_{k1}^1(x), f_{k2}^1(x))^\top$*. Then, according to the formula (6.21), for a given* $\eta_1 \in (0,1)$*,*

$$\rho_k^{a_j^k} \geq \eta_1 \Longrightarrow \max\{f_{k1}^1(x_k) - f_{k1}^1(x_k+s_k), f_{k2}^1(x_k) - f_{k2}^1(x_k+s_k)\} > 0, \tag{6.22}$$

which does not necessarily imply that $f_{k1}^1(x_k+s_k) < f_{k1}^1(x_k)$ *and* $f_{k2}^1(x_k+s_k) < f_{k2}^1(x_k)$*.*

For example, consider

$$\eta_1 = \tfrac{1}{8}, \Omega_k = \tfrac{1}{2}, x_k = 0, f_{k1}^1(x) = 2\sin x - 8\cos x - 10^4 x \sin x^2, \text{ and } f_{k2}^1(x) = \sin x - \tfrac{32}{5}\cos x.$$

Then, at x_k*,* $\omega_k = 1$*, and the model functions corresponding to* f_{k1}^1 *and* f_{k2}^1 *are given by*

$$m_{k1}^1(s_k) = 2s_k + 4s_k^2 \text{ and } m_{k2}^1(s_k) = s_k + \tfrac{16}{5}s_k^2, \text{ respectively.}$$

Hence, the solution to the problem (6.18) is obtained at $s_k = -\frac{5}{32}$*. With this value of* s_k*, we see that* $\rho_k^{a_j^k} = 0.3612 \geq \eta_1$*, but* $f_{k1}^1(x_k+s_k) = 29.9294 > -8 = f_{k1}^1(x_k)$*. Hence, the formula (6.21) cannot be a correct choice of* $\rho_k^{a_j^k}$*.*

It is noteworthy that once $\rho_k^{a_j^k}$ *is taken as* $-\sigma_k^{a_j^k}$*, then (6.22) gives that* $f_{k1}^1(x_k+s_k) < f_{k1}^1(x_k)$ *and* $f_{k2}^1(x_k+s_k) < f_{k2}^1(x_k)$*, i.e,* s_k *is a descent direction at* x_k *for both the* *functions* f_{k1} *and* f_{k2}.

The following theorem ensures that for the chosen a^k that satisfies (6.13), a descent direction of the objective function of ($\mathcal{VOP}_{a^k}(x_k)$) is a descent of the objective function $F_k = \{f_k^i\}_{i\in[p]}$ of ($\mathcal{SOP}_K^l$) at x_k.

Theorem 6.2 *If a trust-region step* s_k *at* x_k *is a descent direction of the objective function* $\widetilde{f}_k^{a^k} = (f_k^{a_1^k}, f_k^{a_2^k}, \ldots, f_k^{a_{\omega_k}^k})^\top$ *of* ($\mathcal{VOP}_{a^k}(x_k)$), *then* s_k *is a descent direction of the objective function* F_k *of* ($\mathcal{SOP}_K^l$) *at* x_k.

Proof: Let s_k be a descent direction of $\widetilde{f}_k^{a^k}$ at x_k. Then, there exists $t_0 > 0$ such that for all $t \in (0, t_0]$,

$$\begin{aligned} & f_k^{a_j^k}(x_k + ts_k) \prec_K f_k^{a_j^k}(x_k) \text{ for all } j \in [\omega_k] \\ \Longrightarrow & \{f_k^{a_j^k}(x_k)\}_{j\in[\omega_k]} \subseteq \{f_k^{a_j^k}(x_k + ts_k)\}_{j\in[\omega_k]} + \text{int}(K) \\ \Longrightarrow & \{f_k^{a_j^k}(x_k)\}_{j\in[\omega_k]} + K \subseteq \{f_k^{a_j^k}(x_k + ts_k)\}_{j\in[\omega_k]} + \text{int}(K) + K. \end{aligned} \tag{6.23}$$

Hence, it gives that

$$\begin{aligned} F_k(x_k) & \subseteq \{f_k^{a_j^k}(x_k)\}_{j\in[\omega_k]} + K \text{ by Proposition 2.1 in [18]} \\ & \overset{(6.23)}{\subseteq} \{f_k^{a_j^k}(x_k + ts_k)\}_{j\in[\omega_k]} + \text{int}(K) + K \\ & \subseteq F_k(x_k + ts_k) + \text{int}(K) \\ \Longrightarrow F_k(x_k + ts_k) & \prec_K^l F_k(x_k). \end{aligned}$$

Then, s_k be a descent direction of $F_k = \{f_k^i\}_{i\in[p]}$ of ($\mathcal{SOP}_K^l$) at x_k. □

6.5.4 Trust-Region radius update

Based on the choice of reduction ratio $\rho_k^{a_j^k}$, we decide the acceptance of the trial step s_k and update the trust-region radius Ω_k. Our update criterion is very similar to the criterion used normally for the multiobjective trust-region methods [20,61]. We compare the reduction ratio $\rho_k^{a_j^k}$ with two pre-specified threshold parameters, $\eta_1, \eta_2 \in (0, 1)$, where η_1 is near to 0 and η_2 is close to 1, and classify the current iteration into following three cases.

Case 1 **(Successful iteration)**: when $\rho_k^{a_j^k} \geq \eta_1$ for all $j \in [\omega_k]$ and $\exists\ l \in [\omega_k]$ such that $\rho_k^{a_l^k} < \eta_2$. If satisfies $\rho_k^{a_j^k} \geq \eta_1$, then for all $j \in [\omega_k]$, we obtain that

$$-\sigma_k^{a_j^k} = \frac{-\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k))}{\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))} > \eta_1$$

$$\text{i.e.,} \ -\Delta_{-K}(f^{a_j^k}(x_k+s_k)-f^{a_j^k}(x_k)) > \eta_1 \Delta_{-K}(-m^{a_j^k}(s_k)). \tag{6.24}$$

Since $\eta_1 > 0$, from (6.24), we have

$$f^{a_j^k}(x_k+s_k) \preceq_K f^{a_j^k}(x_k) \ \forall \ j \in [\omega_k]. \tag{6.25}$$

This means that the trial step s_k can be accepted because the function $f^{a_j^k}$ at the new iterate $x_k + s_k$ is at least equal or less than the function at the current iterate x_k, with respect to K, for each j.

We then have

$$\begin{aligned}\forall \ k \in \mathbb{N} \cup \{0\} : F(x_k) \ &\subseteq \{f^{a_1^k}(x_k), f^{a_2^k}(x_k), \ldots, f^{a_{w_k}^k}(x_k)\} + K \text{ by Proposition 2.1 in [18]} \\ &\overset{(6.25)}{\subseteq} \{f^{a_1^k}(x_k+s_k), f^{a_2^k}(x_k+s_k), \ldots, f^{a_{w_k}^k}(x_k+s_k)\} + \text{int}(K) + K \\ &\subseteq F(x_k+s_k) + \text{int}(K). \end{aligned} \tag{6.26}$$

Therefore, $\{F(x_k)\}$ is a monotonically non-increasing sequence. Moreover, by Lemma 6.4.1 (vii), we have

$$(\Delta_{-K}(f^{a_j^k}(x_k)) - \Delta_{-K}(f^{a_j^k}(x_k+s_k))) > \eta_1(\Delta_{-K}(m^{a_j^k}(0)) - \Delta_{-K}(m^{a_j^k}(s_k))), \tag{6.27}$$

i.e. decrement of each $f^{a_j^k}$, with respect to K, is at least η_1 times the decrement of its associated model $m^{a_j^k}$.

In addition, if there are some $l \in [\omega_k]$ satisfying $\rho_k^{a_l^k} < \eta_2$, then

$$-\sigma_k^{a_l^k} = \frac{-\Delta_{-K}(f^{a_l^k}(x_k+s_k) - f^{a_l^k}(x_k))}{\Delta_{-K}(m^{a_l^k}(0) - m^{a_l^k}(s_k))} < \eta_2 \ \text{,i.e.,} \ \frac{\Delta_{-K}(f^{a_l^k}(x_k+s_k) - f^{a_l^k}(x_k))}{-\Delta_{-K}(m^{\widetilde{a}_l}(0) - m^{\widetilde{a}_l}(s_k))} < \eta_2. \tag{6.28}$$

From Lemma 6.4.1 (vii) and from (6.28), we have

$$(\Delta_{-K}(f^{a_l^k}(x_k)) - \Delta_{-K}(f^{a_l^k}(x_k+s_k))) < \eta_2(\Delta_{-K}(m^{a_l^k}(0)) - \Delta_{-K}(m^{a_l^k}(s_k))),$$

that is, decrements of some of the objectives $f^{a_l^k}$ are smaller than η_2 times the decrement of their model $m^{a_l^k}$. Hence, although decrements of all $f^{a_j^k}$ are bigger than η_1 times the decrements of $m^{a_j^k}$, not all decrements are bigger than η_2 times $m^{a_j^k}$, and since η_1 is a small number closer to 0, we get a small reduction in each objective function $f^{a_j^k}$ compared to the predicted reduction in their models $m^{a_j^k}$. Since this step is not entirely satisfactory, we call this case "successful" (instead of "very successful") and consider

s_k to be acceptable to update x_k to $x_k + s_k$. For the trust-region radius, following the conventional update rule [20,61], Ω_k is updated to $\Omega_{k+1} \in (\gamma_2\Omega_k, \Omega_k]$, where γ_2 is close to 1. Next, we consider the case where all the objectives are reduced fully satisfactorily.

Case 2 (**Very successful iteration**): when iterate x_k satisfies $\rho_k^{a_j^k} \geq \eta_2$, for all $j \in [\omega_k]$. In this case, we have

$$- \sigma_k^{a_j^k} = \frac{-\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k))}{\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))} > \eta_2$$

$$\text{i.e., } \Delta_{-K}(f^{a_l^k}(x_k)) - \Delta_{-K}(f^{a_l^k}(x_k + s_k)) > \eta_2(\Delta_{-K}(m^{a_l^k}(0)) - \Delta_{-K}(m^{a_l^k}(s_k))),$$

which means that decrements of all the objectives $f^{a_j^k}$ with respect to K are η_2 times the decrements of their respective models $m^{a_j^k}$. Since η_2 is closer to 1, every model function $m^{a_j^k}$ acts as a good local approximation of its associated objective functions $f^{a_j^k}$. Therefore, this trust-region step s_k is considered "very successful" for updating x_k to $x_k + s_k$. Following the conventional update rule [20,61], trust-region radius Ω_k is enlarged to $\Omega_{k+1} \in (\Omega_k, \infty)$.

Case 3 (**Unsuccessful iteration**): when $\exists\ l \in [\omega_k]$ such that $\rho_k^{a_l^k} \leq \eta_1$ i.e. when there exists even one objective function whose decrement is less than η_1. In this case, the trust-region update step s_k is declined and we keep hold of the current iterate x_k. The trust- region radius Ω_k, similar to [20, 61], is contracted to the new region $\Omega_{k+1} \in [\gamma_1\Omega_k, \gamma_2\Omega_k]$, where $0 < \gamma_1 \leq \gamma_2 < 1$.

6.5.5 Stopping condition

Let's denote θ, s at x_k by $\theta(x_k)$ and $s(x_k)$, where $\theta(x_k)$ and $s(x_k)$ are the optimal value of the subproblem (6.18) in the image space and solution of (6.18) in variable space, respectively, and given by

$$\theta(x_k) = \min_{s \in \mathcal{B}_k} \max_{j \in [\omega_k]} \left\{ \Delta_{-K} \left(\nabla f^{a_j^k}(x_k)^\top s + \tfrac{1}{2} s^\top \nabla^2 f^{a_j^k}(x_k) s \right), \Delta_{-K} \left(\nabla f^{a_j^k}(x_k)^\top s \right) \right\} \tag{6.29}$$

and

$$s(x_k) = \operatorname*{argmin}_{s \in \mathcal{B}_k} \max_{j \in [\omega_k]} \left\{ \Delta_{-K} \left(\nabla f^{a_j^k}(x_k)^\top s + \tfrac{1}{2} s^\top \nabla^2 f^{a_j^k}(x_k) s \right), \Delta_{-K} \left(\nabla f^{a_j^k}(x_k)^\top s \right) \right\}. \tag{6.30}$$

In the following theorem, we study some properties of the function θ and analyse its relation with the descent direction $s(x)$. Alongside, we present a characterization of criticality at x_k for $(\mathcal{SOP}^l_K)$ in terms of θ that is important for defining the stopping condition and for the convergence analysis of our algorithm that follows.

Theorem 6.3 *For the functions θ and s in (6.11) and (6.30), respectively, the following results hold:*

(a) The mapping θ is continuous $\forall\ x \in \mathcal{R}$, where $\mathcal{R}$ is the set of all regular point for ($\mathcal{SOP}^l_K$), and well defined $\forall\ x \in \mathbb{R}^n$.

(b) The following three statements satisfy (i) $\Longleftrightarrow$ (ii) $\Longrightarrow$ (iii):

(i) x_k is not a critical point for ($\mathcal{SOP}^l_K$);

(ii) $\theta(x_k) < 0$;

(iii) $s(x_k) \neq 0$.

Proof: (a) Let $\bar{x} \in \mathcal{R}$ and $\varepsilon > 0$. Since $\bar{x}$ is a regular point of F, there exists a neighbourhood U of $\bar{x}$ such that

$$\omega(z) = \bar{\omega} \text{ and } P_z \subseteq P_{\bar{x}}\ \forall\ z \in U.$$

For any given $x \in \mathbb{R}^n$, $\bar{a} = (\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_{\bar{\omega}}) \in P_z$ and $j \in [\bar{\omega}]$, we define two functions $\phi_{x,\bar{a}_j} : \mathbb{R}^n \to \mathbb{R}$ and $\psi_{x,\bar{a}_j} : \mathbb{R}^n \to \mathbb{R}$ by

$$\begin{aligned} \phi_{x,\bar{a}_j}(z) &= \Delta_{-K}\left(\nabla f^{\bar{a}_j}(z)^\top s(x) + \tfrac{1}{2}s(x)^\top \nabla^2 f^{\bar{a}_j}(z)s(x)\right), \\ \text{and } \psi_{x,\bar{a}_j}(z) &= \Delta_{-K}\left(\nabla f^{\bar{a}_j}(z)^\top s(x)\right). \end{aligned}$$

Then, we find that $\theta(z) = \min\limits_{(\bar{a},s)\in P_z\times\mathcal{B}} \max\limits_{j\in[\bar{\omega}]} \left\{\phi_{x,\bar{a}_j}(z), \psi_{x,\bar{a}_j}(z)\right\}$. Let $\mathcal{W} \subseteq \mathbb{R}^n$ be a compact set containing $\bar{x}$. We note for any $z \in U \cap \mathcal{W}$ that

$$\begin{aligned} &\max_{j\in[\bar{\omega}]} |\phi_{x,\bar{a}_j}(z) - \phi_{x,\bar{a}_j}(\bar{x})| \\ \leq\ &\max_{j\in[\bar{\omega}]} \left\| \left(\nabla f^{\bar{a}_j}(z)^\top s(x) + \tfrac{1}{2}s(x)^\top \nabla^2 f^{\bar{a}_j}(z)s(x)\right) - \left(\nabla f^{\bar{a}_j}(\bar{x})^\top s(x) + \tfrac{1}{2}s(x)^\top \nabla^2 \right.\right. \\ &\left.\left. f^{\bar{a}_j}(\bar{x})s(x)\right)\right\| \quad \text{(by Lemma 6.4.1 ((i)))} \\ \leq\ &\max_{j\in[\bar{\omega}]} \left\{\left\|\left(\nabla f^{\bar{a}_j}(z) - \nabla f^{\bar{a}_j}(\bar{x})\right)^\top s(x)\right\| + \tfrac{1}{2}\left\|s(x)^\top\left(\nabla^2 f^{\bar{a}_j}(z) - \nabla^2 f^{\bar{a}_j}(\bar{x})\right)s(x)\right\|\right\} \\ \leq\ &\|s(x)\| \max_{j\in[\bar{\omega}]}\|\nabla f^{\bar{a}_j}(z) - \nabla f^{\bar{a}_j}(\bar{x}))\| + \tfrac{1}{2}\|s(x)\|^2 \max_{j\in[\bar{\omega}]}\|\nabla^2 f^{\bar{a}_j}(z) - \nabla^2 f^{\bar{a}_j}(\bar{x}))\|. \end{aligned} \tag{6.31}$$

Since $f^{\bar{a}_j}$ is twice continuously differentiable for any $j \in [\bar{\omega}]$, the functions $\nabla f^{\bar{a}_j}$ and $\nabla^2 f^{\bar{a}_j}$ are uniformly continuous in $\mathcal{W}$. Hence, for the given ε, there is $\delta_1 > 0$

such that $\|z-\bar{x}\| < \delta_1$ and $z \in U \cap \mathcal{W}$ imply $\|\nabla f^{\bar{a}_j}(z) - \nabla f^{\bar{a}_j}(\bar{x}))\| \leq \frac{\varepsilon}{2\Omega_{\max}}$. Similarly, there is $\delta_2 > 0$ such that $\|z-\bar{x}\| < \delta_2$ and $z \in U \cap \mathcal{W}$ imply $\|\nabla^2 f^{\bar{a}_j}(z) - \nabla^2 f^{\bar{a}_j}(\bar{x}))\| \leq \frac{\varepsilon}{\Omega_{\max}^2}$. Let $\delta = \min\{\delta_1, \delta_2\} > 0$. As $\|s(x)\| \leq \Omega_{\max}$, from (6.31), we have

$$\|z-\bar{x}\| < \delta \text{ and } z \in U \cap \mathcal{W} \implies \max_{j \in [\bar{\omega}]} |\phi_{x,\bar{a}_j}(z) - \phi_{x,\bar{a}_j}(\bar{x})| \leq \varepsilon.$$

This means that $\{\phi_{x,\bar{a}_j}\}_{x \in \mathbb{R}^n, j \in [\bar{\omega}]}$ is equicontinuous at $\bar{x}$. Similarly, $\{\psi_{x,\bar{a}_j}\}_{x \in \mathbb{R}^n, j \in [\bar{\omega}]}$ is equicontinuous at $\bar{x}$. Therefore, the family $\{\Phi_x\}_{x \in \mathbb{R}^n}$, where

$$\Phi_x(z) = \max_{j \in [\bar{\omega}]} \{\phi_{x,\bar{a}_j}(z), \psi_{x,\bar{a}_j}(z)\}, \; z \in U \cap \mathcal{W},$$

is equicontinuous. Then, for the given $\varepsilon > 0$, there exists $\bar{\delta} > 0$ such that for all $z \in U \cap \mathcal{W}$ satisfying $\|z - \bar{x}\| < \delta$, we have

$$|\Phi_x(z) - \Phi_x(\bar{x})| < \varepsilon \; \forall \; x \in \mathbb{R}^n.$$

Hence, for all $z \in U \cap \mathcal{W}$ satisfying $\|z - \bar{x}\| < \bar{\delta}$, it follows that

$$\begin{aligned}
\theta(z) &\leq \max_{j \in [\bar{\omega}]} \left\{ \Delta_{-K} \left(\nabla f^{\bar{a}_j}(z)^\top s(\bar{x}) + \tfrac{1}{2} s(\bar{x})^\top \nabla^2 f^{\bar{a}_j}(z) s(\bar{x}) \right), \Delta_{-K} \left(\nabla f^{\bar{a}_j}(z)^\top s(\bar{x}) \right) \right\} \\
&= \Phi_{\bar{x}}(z) \\
&\leq \Phi_{\bar{x}}(\bar{x}) + |\Phi_{\bar{x}}(z) - \Phi_{\bar{x}}(\bar{x})| \\
&< \theta(\bar{x}) + \varepsilon,
\end{aligned}$$

i.e., $\theta(z) - \theta(\bar{x}) < \varepsilon$. By altering the role of z and $\bar{x}$, we find that $|\theta(\bar{x}) - \theta(z)| < \epsilon$. Thus, the continuity of θ at $\bar{x}$ follows. Hence, θ is continuous in $\mathcal{R}$.

As for any $x \in \mathcal{R}^n$, the value of $\theta(x)$ is obtained by minimization of a maximum function of continuous functions over the compact set $P_x \times B$, function θ is well-defined.

(b) $\boxed{(i) \text{ implies } (ii)}$: Since x_k is not critical for ($\mathcal{SOP}^l_K$), there exists $a^k \in P_k$ and $s_k \in \mathcal{B}_k$ such that

$$\begin{aligned}
&\nabla f^{a_j}(x_k)^\top s_k \prec_K 0_{m \times 1} \; \forall \; j \in [\omega_k] \\
\implies &\nabla f^{a^k_j}(x_k)^\top s_k \in -\text{int}(K) \; \forall \; j \in [\omega_k] \\
\implies &\Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top s_k) < 0 \; \forall \; j \in [\omega_k].
\end{aligned} \tag{6.32}$$

Case 1. Let for all $j \in [\omega_k]$, $\Delta_{-K}({s_k}^\top \nabla^2 f^{a_j^k}(x_k) s_k) \leq 0$. In this case, it is obvious that

$$\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k) + \tfrac{1}{2}\Delta_{-K}({s_k}^\top \nabla^2 f^{a_j^k}(x) s_k) < 0,$$

which, by Lemma 6.4.1 (vii), implies that

$$\Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s_k + \tfrac{1}{2}{s_k}^\top \nabla^2 f^{a_j^k}(x_k) s_k\right) < 0. \tag{6.33}$$

Combining (6.32) and (6.33), we get $\theta(x_k) < 0$.

Case 2. Let $\Delta_{-K}({s_k}^\top \nabla^2 f^{a_j^k}(x_k) s_k) > 0$ for some $j \in [\omega_k]$. Then, by considering $s'_k = \alpha s_k$, with $\alpha > 0$ small enough, we have

$$\Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s'_k\right) + \tfrac{1}{2}\Delta_{-K}\left({s'_k}^\top \nabla^2 f^{a_j^k}(x_k) s'_k\right) < 0. \tag{6.34}$$

For instance, any α that satisfies $0 < \alpha < \frac{-2\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\Delta_{-K}(s_k^\top \nabla^2 f^{a_j^k}(x_k) s_k)}$ holds the inequality (6.34). As s_k lies in $\mathcal{B}_k$, a small enough choice of α leads to $s'_k \in \mathcal{B}_k$. With such an $s'_k \in \mathcal{B}_k$, we note from (6.32) that

$$\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s'_k) = \alpha \Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s'_k) < 0 \;\forall\; j \in [\omega_k]. \tag{6.35}$$

From the construction of (6.34), we see by Lemma 6.4.1 ((vii)) that

$$\Delta_{-K}\left(\nabla f^{a_j^k}(x_k)^\top s'_k + \tfrac{1}{2}{s'_k}^\top \nabla^2 f^{a_j^k}(x_k) s'_k\right) < 0 \;\forall\; j \in [\omega_k]. \tag{6.36}$$

In view of (6.35) and (6.36), we get $\theta(x_k) < 0$.

$\boxed{(ii) \text{ implies } (i)}$: If $\theta(x_k) < 0$, then there exists a feasible point (t_k, s_k), with $t_k < 0$, of (6.18) such that

$$\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k) \leq t_k < 0 \text{ for all } j \in [\omega_k].$$

So, x_k is not critical.

$\boxed{(ii) \text{ implies } (iii)}$: On contrary, assume that $s(x_k) = 0$. Then, from the definition of $s(x_k)$ and (6.29), we get $\theta(x_k) = 0$. This contradicts $\theta(x_k) < 0$. So, $s(x_k) \neq 0$.

□

Algorithm 3 Trust-Region algorithm for solving ($\mathcal{SOP}^l_K$)

1: *Input and initialization*
Provide the functions $f^i : \mathbb{R}^n \to \mathbb{R}^m$, $i = 1, 2, \ldots, p$, in the problem ($\mathcal{SOP}^l_K$).
Choose an initial point x_0, an initial trust-region radius Ω_0, a maximum allowed trust-region radius $\Omega_{\max}$, a tolerance value ϵ for stopping criterion, threshold parameters $\eta_1, \eta_2 \in (0, 1)$, and fractions $\gamma_1, \gamma_2 \in (0, 1)$ to shrink the trust-region radius.
Set the iteration counter $k = 0$.

2: *Computation of minimal elements*
Compute $M_k = \text{Min}(F(x_k), K) = \{r_1, r_2, \ldots, r_{\omega_k}\}$.
Compute the partition set $P_k = I_{r_1} \times I_{r_2} \times \cdots \times I_{r_{\omega_k}}$.
Find $p_k = |P_k|$ and $\omega_k = |\text{Min}(F(x_k), K)|$.

3: *Choice of an 'a^k' from P_k*
Select an element $a^k = (a^k_1, a^k_2, \ldots, a^k_{\omega_k}) \in P_k$ such that

$$(a^k, s_k) \in \underset{(a,s)\in P_k\times\mathcal{B}_k}{\text{argmin}} \max_{j\in[\omega_k]} \left\{ \Delta_{-K}\left(\nabla f^{a_j}(x_k)^\top s + \tfrac{1}{2}s^\top \nabla^2 f^{a_j}(x_k)s\right), \Delta_{-K}\left(\nabla f^{a_j}(x_k)^\top s\right) \right\}.$$

4: *Model definition*
Compute the model functions $m_k^{a_j^k}$ for all $j \in [\omega_k]$ by the formula

$$m_k^{a_j^k}(s) = \nabla f^{a_j^k}(x_k)^\top s + \tfrac{1}{2}s^\top \nabla^2 f^{a_j^k}(x_k)s, \ \|s\| \leq \Omega_k.$$

5: *Step size calculation*
Find (t_k, s_k) by solving the subproblem (6.18).

6: *Termination criterion*
If $|t_k| < \epsilon$, then stop and provide x_k as an approximate critical point of ($\mathcal{SOP}^l_K$).
Else, proceed to the next step.

7: *Calculation of reduction ratio*
Execute

$$\sigma_k^{a_j^k} = \frac{\Delta_{-K}(f^{a_j^k}(x_k + s_k) - f^{a_j^k}(x_k))}{\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k))} \text{ and } \rho_k^{a_j^k} = -\sigma_k^{a_j^k} \text{ for all } j \in [\omega_k].$$

8: *Acceptance of the computed step size*
If $\rho_k^{a_j^k} \geq \eta_1$ for all $j \in [\omega_k]$, then update $x_{k+1} = x_k + s_k$.

9: *Rejection of the computed step size*
If $\exists j \in [\omega_k], \rho_k^{a_j^k} < \eta_1$, then $x_{k+1} = x_k$.

10: *Trust-region radius update*
Choose

$$\Omega_{k+1} \in \begin{cases} (\gamma_2\Omega_k, \Omega_k], & \text{if } \eta_1 \leq \rho_k^{a_j^k} \ \forall\, j \in [\omega_k] \text{ and } \exists\, l \in [\omega_k] \text{ such that } \rho_k^{a_l^k} < \eta_2 \ (\textbf{Successful}) \\ (\Omega_k, \infty), & \text{if } \eta_2 \leq \rho_k^{a_j^k} \ \forall\, j \in [\omega_k], \ (\textbf{Very successful}) \\ [\gamma_1\Omega_k, \gamma_2\Omega_k], & \text{if } \exists\, l \in [\omega_k] \text{ such that } \rho_k^{a_l^k} \leq \eta_1 \ (\textbf{Unsuccessful}). \end{cases}$$

11: *Move to next iteration*
$k \leftarrow k + 1$ and go to Step 2:.

6.5.6 Well-definedness of Algorithm 3

The well-definedness of Algorithm 3 depends upon Step 3:, Step 5:, Step 6:, Step 7: and Step 8:.

In Step 3:, we choose a specific a^k from P_k as defined in (6.13). This a^k is then used (in Step 5:) to solve the problem (6.18) and obtain a possible descent direction s_k. In case a descent direction is not obtained, the point of criticality x_k for ($\mathcal{SOP}^l_K$) is reached, which readily follows from the direct implication of items $(i)-(ii)$ and $(ii)-(iii)$ of Theorem 6.3 (b). Hence, the specific a^k is chosen so that the necessary condition of weak minimal point or critical point of ($\mathcal{SOP}^l_K$) is met. The well-definedness of Step 3: is discussed in more details in Subsection 6.5.1.

In Step 5:, we solve the subproblem (6.18) to find a solution (t_k, s_k). For this step to be well-defined, it is required that a solution exists for the subproblem (6.18). From Lemma 6.4.1 (i), we recall that Δ_{-K} is continuous function. Since the subproblem (6.18) is a minimization of a continuous function over the compact region $\{s : \|s\| \leq \Omega_k\}$, then a solution (t_k, s_k) to the subproblem (6.18) is assured in Step 5:.

In Step 6:, we check for the termination criterion $|t_k| < \epsilon$. In general, the parameter ϵ used in Algorithm 3 can be arbitrarily small. Therefore, the condition in Step 6: implies $t_k = 0$, which in turn implies that $\theta(x_k) = 0$. Thus x_k is an approximate critical point of ($\mathcal{SOP}^l_K$), and this conclusion follows from the contrapositive statement of item $(i) - (ii)$ of Theorem 6.3 (b).

In Step 7:, we calculate the reduction ratio. It is important to note that the predicted reduction due to the identified solution s_k of the subproblem (6.18), $\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k))$ is positive at the noncritical point x_k as shown in Corollary 6.6.1 (see (6.47)). Then, the reduction ratio $\rho_k^{a_j^k}$ is well-defined.

In Step 8:, we check if the computed step size can be accepted or not. Due to Proposition 6.5.2, there exists a stepsize s_k along descent direction, which satisfies $\rho_k^{a_j^k} \geq \eta_1 > 0$ for all $j \in [\omega_k]$. This concludes the well-definedness of Step 8:.

Therefore, we can conclude that s_k is a descent direction and Algorithm 3 generates an infinite sequence of noncritical points $\{x_k\}$. As a consequence, we obtain $f^{a_j^k}(x_k + s_k) \preceq_K f^{a_j^k}(x_k),\ \forall\ j \in [\omega_k]$. This concludes that $F(x_k + s_k) \preceq^l_K F(x_k)$ for every $k \in \mathbb{N}$ as shown previously through (6.26) in Subsection 6.5.4. Therefore, Algorithm 3 is an $\preceq^l_K$-decreasing method.

6.6 Global convergence analysis

In this section, we analyze the proof that Algorithm 3 converges to a critical point for ($\mathcal{SOP}^l_K$). For this, we first make a few assumptions on the given set-valued function

$F = \{f^i\}_{i\in[p]}$. Note that these assumptions are very natural for single and multi-objective trust-region schemes.

Assumption 6.6.1 *Function f^i is continuously differentiable for all $i \in [p]$.*

Assumption 6.6.2 *The Hessian of function f^i, for all $i \in [p]$, is uniformly bounded, which means that there exists a $\mathcal{K}_1$ such that*

$$\|\nabla^2 f^i(x)\| \leq \mathcal{K}_1 \text{ for all } i \in [p].$$

Assumption 6.6.3 *The level set for F, defined as $\mathcal{L}_c = \{x \in \mathbb{R}^n : F(x) \preceq^l_K F(x_0)\}$, is bounded.*

Assumption 6.6.4 *Function f^i, for all $i \in [p]$, is bounded below.*

Assumption 6.6.5 *There is a $\Omega_{\max} > 0$ such that*

$$\Omega_k \leq \Omega_{\max} \ \forall \ k \geq 0.$$

Assumption 6.6.6 *There is a $\mathcal{K}_3 > 0$ such that, for all $i \in [p]$,*

$$\|\nabla f^i(x)\| \leq \mathcal{K}_3.$$

For our objective function f^{a^k} of ($\mathcal{VOP}_{a^k}(x_k)$), with these assumptions, we prove the convergence of Algorithm 3 with the support of a few other results as outlined next. Theorem 4.1 and Corollary 4.1 demonstrate results regarding sufficient decrease of the model function. Theorem 4.2 addresses the accuracy of model function $m^{a^k_j}$ in approximating its objective $f^{a^k_j}$. Theorem 4.3 discusses criteria for accepting trial step and updating trust-region radius. Theorem 4.4, employing Lemma 4.1, shows the convergence of algorithm for finitely many successful iterations. Finally, Theorem 4.5, using Lemma 4.2, establishes the algorithm's convergence over infinitely many successful iteration.

In the first step, akin to classical trust-region schemes, we derive a theorem concerning the adequate decrease of the model function m^{a^k}, within the trust-region Ω_k, for our considered ($\mathcal{VOP}_{a^k}(x_k)$). Specifically, we show that the predicted reduction of the objective function f^{a^k} along any descent direction v, denoted by $\Delta_{-K}(m^{a^k_j}(0) - m^{a^k_j}(\bar{t}v))$, is lower bounded by some sufficient amount.

Theorem 6.4 *Let x_k be a non-critical point of ($\mathcal{VOP}_{a^k}(x_k)$) at iteration $k \in \mathbb{N}$. Further, let $v \in \mathbb{R}^n$ be a descent direction of f^{a^k} at x_k. Then, for each $j \in [\omega_k]$, there exists*

$\bar{t} > 0$ *satisfying* $\|\bar{t}v\| \leq \Omega_k$, *such that*

$$\Delta_{-K}(m^{a^k_j}(\bar{t}v)) \leq \Delta_{-K}(m^{a^k_j}(tv)) \ \ \forall\, t > 0 \text{ satisfying } \|tv\| \leq \Omega_k.$$

In addition, for all $j \in [\omega_k]$,

$$\Delta_{-K}(m^{a^k_j}(0) - m^{a^k_j}(\bar{t}v)) \geq -\tfrac{1}{2}\frac{\Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top v)}{\|v\|} \min\left\{-\frac{\Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top v)}{\|v\|\mathcal{K}_1},\ \Omega_k\right\},$$

and $\Delta_{-K}(m^{a^k_j}(0)) \geq \Delta_{-K}(m^{a^k_j}(\bar{t}v))$.

Proof: Suppose that x_k is a non-critical point of ($\mathcal{VOP}_{a^k(x_k)}$). For all $j \in [\omega_k]$, let us define a scalar function $\phi : \left[0, \frac{\Omega_k}{\|v\|}\right] \to \mathbb{R}$ as

$$\phi(t) = \Delta_{-K}(m^{a^k_j}(tv)) \quad \forall\, t \in \left[0, \frac{\Omega_k}{\|v\|}\right], \forall\, j \in [\omega_k].$$

Since Δ_{-K} and $m^{a^k_j}$ are continuous on $\mathbb{R}^n$, ϕ is also continuous on the compact set $\left[0, \frac{\Omega_k}{\|v\|}\right]$. Assuming that $\phi'(t)$ exists, using (6.14), $\phi'(t)$ can be written as

$$\begin{aligned}\phi'(t) = {} & \Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top v + \frac{1}{2} t v^\top \nabla^2 f^{a^k_j}(x_k) v) \\ & - t \inf_{c \in (-K)^c} \frac{\nabla f^{a^k_j}(x_k)^\top v + \frac{1}{2} t v^\top \nabla^2 f^{a^k_j}(x_k) v - c}{\|\nabla f^{a^k_j}(x_k)^\top v + \frac{1}{2} t v^\top \nabla^2 f^{a^k_j}(x_k) v - c\|}.\end{aligned}$$

Now, since v is a descent direction of f^{a^k}, it holds that $\phi'(0) = \Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top v) < 0$. This implies that ϕ is strictly decreasing on $\left[0, \frac{\Omega_k}{\|v\|}\right]$. Also, since ϕ is continuous on $\left[0, \frac{\Omega_k}{\|v\|}\right]$, ϕ is bounded on $\left[0, \frac{\Omega_k}{\|v\|}\right]$. Further, since ϕ is strictly decreasing on $\left[0, \frac{\Omega_k}{\|v\|}\right]$, $\sup\phi = \phi(0)$ and $\inf\phi = \phi\left(\frac{\Omega_k}{\|v\|}\right)$. Then, for any t, t' satisfying $0 \leq t < t' \leq \frac{\Omega_k}{\|v\|}$, we have

$$\phi\left(\frac{\Omega_k}{\|v\|}\right) < \phi(t') < \phi(t) \leq \phi(0). \tag{6.37}$$

Therefore, ϕ is injective on $\left[0, \frac{\Omega_k}{\|v\|}\right]$ and ϕ follows intermediate value property. Now, if we take a $\bar{t} \in \left[0, \frac{\Omega_k}{\|v\|}\right]$ such that $\phi(\bar{t}) = \phi(t')$, then from injectivity of ϕ, it must be that $\bar{t} = t' > 0$, and, from (6.37), we have

$$\phi(\bar{t}) \leq \phi(t)$$

for all $t > 0$ satisfying $\|tv\| \leq \Omega_k$. This concludes the first part of the proof that, for each $j \in [\omega_k]$, there exists $\bar{t} > 0$ satisfying $\|\bar{t}v\| \leq \Omega_k$, such that

$$\Delta_{-K}(m^{a_j^k}(\bar{t}v)) \leq \Delta_{-K}(m^{a_j^k}(tv)) \quad \forall\, t > 0 \text{ satisfying } \|tv\| \leq \Omega_k.$$

For the second part of the proof, since x_k is a non-critical point, it follows from Theorem 6.3 that $\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top tv + \frac{1}{2}t^2 v^\top \nabla^2 f^{a_j^k}(x_k)v) < 0$. This implies that $\nabla f^{a_j^k}(x_k)^\top tv + \frac{1}{2}t^2 v^\top \nabla^2 f^{a_j^k}(x_k)v \in -\text{int}(K)$ due to property (ii) of Lemma 6.4.1. Thus, we can rewrite

$$\phi(t) = -\inf_{\widetilde{a} \in (-K)^c} \|(\nabla f^{a_j^k}(x_k)^\top tv + \tfrac{1}{2}t^2 v^\top \nabla^2 f^{a_j^k}(x_k)v) - \widetilde{a}\|. \tag{6.38}$$

Next, we estimate the parameter t at which ϕ is optimal, denoted by t^*. Let us write $\widetilde{a} = ta'$, where $t > 0$. Note that, since a' is arbitrary, $\widetilde{a}$ is also arbitrary. Assuming again that $\phi'(t)$ exists, differentiating (6.38) with respect to t and equating it to zero, we get

$$\phi'(t^*) = -\inf_{a' \in (-K)^c} \tfrac{(\nabla f^{a_j^k}(x_k)^\top t^* v + \frac{1}{2}(t^*)^2 v^\top \nabla^2 f^{a_j^k}(x_k)v - t^* a')^\top}{\|\nabla f^{a_j^k}(x_k)^\top t^* v + \frac{1}{2}(t^*)^2 v^\top \nabla^2 f^{a_j^k}(x_k)v - t^* a'\|} \{\nabla f^{a_j^k}(x_k)^\top v + t^* v^\top \nabla^2 f^{a_j^k}(x_k)v - a'\} = 0 \tag{6.39}$$

From (6.39), it implies that

$$\begin{aligned} &-\inf_{a' \in (-K)^c} \frac{(\nabla f^{a_j^k}(x_k)^\top \frac{t^*}{2}v + \left(\frac{t^*}{2}\right)^2 v^\top \nabla^2 f^{a_j^k}(x_k)v - \frac{t^*}{2}a')^\top}{\|(\nabla f^{a_j^k}(x_k)^\top \frac{t^*}{2}v + \left(\frac{t^*}{2}\right)^2 v^\top \nabla^2 f^{a_j^k}(x_k)v - \frac{t^*}{2}a')\|} \{\nabla f^{a_j^k}(x_k)^\top v + t^* v^\top \nabla^2 f^{a_j^k}(x_k)v - a'\} = 0 \\ \implies &\inf_{a' \in (-K)^c} \frac{\left| \left(\nabla f^{a_j^k}(x_k)^\top \frac{t^*}{2}v + \left(\frac{t^*}{2}\right)^2 v^\top \nabla^2 f^{a_j^k}(x_k)v - \frac{t^*}{2}a'\right)^\top \left(\nabla f^{a_j^k}(x_k)^\top v + t^* v^\top \nabla^2 f^{a_j^k}(x_k)v - a'\right) \right|}{\|(\nabla f^{a_j^k}(x_k)^\top \frac{t^*}{2}v + \left(\frac{t^*}{2}\right)^2 v^\top \nabla^2 f^{a_j^k}(x_k)v - \frac{t^*}{2}a')\|} = 0. \end{aligned}$$

Applying Cauchy-Schwarz inequality, we can recast (6.39) as

$$\begin{aligned} & 0 \leq \inf_{a' \in (-K)^c} \|\nabla f^{a_j^k}(x_k)^\top v + t^* v^\top \nabla^2 f^{a_j^k}(x_k)v - a'\| \\ \implies & -\inf_{a' \in (-K)^c} \|\nabla f^{a_j^k}(x_k)^\top v + t^* v^\top \nabla^2 f^{a_j^k}(x_k)v - a'\| \leq 0 \\ \implies & \Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v + t^* v^\top \nabla^2 f^{a_j^k}(x_k)v) \leq 0 \\ \overset{\text{Lemma 6.4.1 (ii)+(iii)}}{\implies} & \nabla f^{a_j^k}(x_k)^\top v + t^* v^\top \nabla^2 f^{a_j^k}(x_k)v \in -K \\ \implies & -\nabla f^{a_j^k}(x_k)^\top v = t^* v^\top \nabla^2 f^{a_j^k}(x_k)v - p \text{ for some } p \in -K \\ \overset{\text{Lemma 6.4.1 (v)}}{\implies} & \Delta_{-K}(-\nabla f^{a_j^k}(x_k)^\top v) = \Delta_{-K+p}(t^* v^\top \nabla^2 f^{a_j^k}(x_k)v) \\ & \qquad = t^* \Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v) \\ \implies & t^* = \tfrac{\Delta_{-K}(-\nabla f^{a_j^k}(x_k)^\top v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)}. \end{aligned}$$

Now, we split into two possible cases.
Case 1 : when $\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v) > 0$. In this case,

$$t^* = \frac{\Delta_{-K}(-\nabla f^{a_j^k}(x_k)^\top v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)} > -\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)} > 0.$$

Here, we have two further subcases. (1) Subcase 1: when t^* lies within the trust-region, i.e. $\|t^* v\| \leq \Omega_k$. We choose $\bar{t} = -\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)} < \frac{\Delta_{-K}(-\nabla f^{a_j^k}(x_k)^\top v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)} = t^*$, then

$$\begin{aligned}
&\Delta_{-K}(m^{a_j^k}(\bar{t}v)) \\
&= \Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top \bar{t}v + \tfrac{1}{2}(\bar{t}v)^\top \nabla^2 f^{a_j^k}(x_k)(\bar{t}v)) \\
&\leq \bar{t}\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v) + \tfrac{1}{2}\bar{t}^2 \Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v) \\
&= -\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)}\Delta_{-K}(\nabla f^{a_j^k}(x_k)v) + \tfrac{1}{2}\left(\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)}\right)^2 \Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v) \\
&= -\tfrac{1}{2}\frac{(\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)^2}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)}.
\end{aligned}$$

Owing to the property ((i)) of Lemma 6.4.1, we get

$$|\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v) - \Delta_{-K}(0)| \leq \frac{\|v\|^2 \|\nabla^2 f^{a_j^k}(x_k)\|}{\|v\|^2} \leq \mathcal{K}_1.$$

Now, using the fact $\Delta_{-K}(m_k^{a_j^k}(0)) = 0$, we obtain

$$\Delta_{-K}(m_k^{a_j^k}(0)) - \Delta_{-K}(m_k^{a_j^k}(\bar{t}v)) \geq \tfrac{1}{2}\frac{(\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v))^2}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)} \geq \tfrac{1}{2}\frac{(\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v))^2}{\|v\|^2 \mathcal{K}_1}. \tag{6.40}$$

(2) Subcase 2: when t^* lies outside the trust-region, i.e. $\|t^* v\| > \Omega_k$ when $-\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)\|v\|}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)}$ $> \Omega_k$. In this subcase, we take $\bar{t} = \frac{\Omega_k}{\|v\|}$, then

$$\Delta_{-K}(m_k^{a_j^k}(\bar{t}v)) \leq \frac{\Omega_k}{\|v\|}\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v) + \tfrac{1}{2}\left(\frac{\Omega_k}{\|v\|}\right)^2 \Delta_{-K}(v^\top \nabla^2 f^{a_{kj}}(x_k)v).$$

From $-\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)}\|v\| > \Omega_k = \bar{t}\|v\|$, we have $\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\bar{t}} < -\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v)$. Therefore, using the last inequality,

$$\Delta_{-K}(m_k^{a_j^k}(0)) - \Delta_{-K}(m_k^{a_j^k}(\bar{t}v)) \geq -\frac{\Omega_k}{\|v\|}\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v) + \tfrac{1}{2}\left(\frac{\Omega_k}{\|v\|}\right)^2 \frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\bar{t}}$$

$$= -\tfrac{1}{2}\frac{\Omega_k}{\|v\|}\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v). \tag{6.41}$$

Case 2: when $\Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)v) \le 0$. In this case, since the minimum value of the model function is on the boundary of the trust-region, we take $\bar{t} = \frac{\Omega_k}{\|v\|}$ and

$$\begin{aligned}\Delta_{-K}(m^{a_j^k}(\bar{t}v)) &\le \frac{\Omega_k}{\|v\|}\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v) + \tfrac{1}{2}\left(\frac{\Omega_k}{\|v\|}\right)^2 \Delta_{-K}(v^\top \nabla^2 f^{a_j^k}(x_k)^\top v)\\ &\le \frac{\Omega_k}{\|v\|}\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v).\end{aligned}$$

Then,

$$\Delta_{-K}(m_k^{a_j^k}(0)) - \Delta_{-K}(m_k^{a_j^k}(\bar{t}v)) \ge -\tfrac{1}{2}\frac{\Omega_k}{\|v\|}\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v). \tag{6.42}$$

Finally, from (6.40), (6.41), and (6.42), we have, at the non-critical point x_k, that

$$\begin{aligned}&\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(\bar{t}v))\\ \overset{\text{Lemma 6.4.1 (vii)}}{\ge}\ &\Delta_{-K}(m_k^{a_j^k}(0)) - \Delta_{-K}(m_k^{a_j^k}(\bar{t}v))\\ \ge\ &-\tfrac{1}{2}\left(\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\|v\|}\right)\min\left\{-\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top v)}{\|v\|\mathcal{K}_1}, \Omega_k\right\} > 0,\end{aligned}$$

which completes the proof.

□

Next, we show that the result in Theorem 6.4, applicable to a descent direction, also holds true for step s_k obtained by solving the subproblem (6.18). This is mainly because x_k being non-critical implies that step s_k is also a descent direction.

Corollary 6.6.1 *If s_k is a solution of the subproblem (6.18), and x_k is a non-critical point of ($\mathcal{VOP}_{a^k}(x_k)$), then there exists a positive constant β such that, for all $j \in [\omega_k]$,*

$$\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k)) \ge -\tfrac{\beta}{2}\tfrac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\|s_k\|}\min\left\{-\tfrac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\|s_k\|\mathcal{K}_1}, \Omega_k\right\}$$

and

$$\Delta_{-K}(m_k^{a_j^k}(0)) \ge \Delta_{-K}(m_k^{a_j^k}(s_k)).$$

Proof: Let (s_k, t_k) be a solution of the subproblem (6.18). Then, from Theorem 6.3, $t_k \le 0$ and $\Delta_{-K}(m_k^{a_j^k}(s_k)) \le t_k$ for all $j \in [\omega_k]$. If Algorithm 3 doesn't terminate at the step 6:, then $-t_k = |t_k| \geq \text{tol}$. Hence using Lemma 6.4.1 (vii), we can write

$$\Delta_{-K}(m_k^{a_j^k}(0) - m^{a_j^k}{}_k(s_k)) \geq \Delta_{-K}(m_k^{a_j^k}(0)) - \Delta_{-K}(m_k^{a_j^k}(s_k)) \geq -t_k \geq \text{tol}, \tag{6.43}$$

for all $j = 1, 2, \ldots, \omega_k$. Moreover, if we choose a $\bar{t}$ as in Theorem 6.4, then by Assumptions 6.6.2, 6.6.5 and 6.6.6, we have

$$\begin{aligned}
\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(\bar{t}s_k)) &= \Delta_{-K}(-m_k^{a_j^k}(\bar{t}s_k)) \\
&= \Delta_{-K}(-\bar{t}\nabla f^{a_j^k}(x_k)^\top s_k - \tfrac{1}{2}\bar{t}^2 s_k^\top \nabla^2 f^{a_j^k}(x_k) s_k) \\
&\leq \|\bar{t}\nabla f^{a_j^k}(x_k)^\top s_k + \tfrac{1}{2}\bar{t}^2 s_k^\top \nabla^2 f^{a_j^k}(x_k) s_k\| \text{ by Lemma 6.4.1 ((i))} \\
&\leq \|\bar{t}s_k\|\|\nabla f^{a_j^k}(x_k)\| + \tfrac{1}{2}\|ts_k\|^2\|\nabla^2 f^{a_j^k}(x_k)\| \\
&\leq \mathcal{K}_3\Omega_{\max} + \tfrac{1}{2}\Omega_{\max}^2\mathcal{K}_1 \text{ for all } j \in [\omega_k].
\end{aligned} \tag{6.44}$$

Then, using (6.43) and (6.44) we have, for all $j \in [\omega_k]$, that

$$\frac{\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k))}{\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(\bar{t}s_k))} \geq \frac{\text{tol}}{\mathcal{K}_3\Omega_{\max} + \frac{1}{2}\Omega_{\max}^2\mathcal{K}_1} = \beta, \text{say} \tag{6.45}$$

$$\implies \Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k)) \geq \beta\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(\bar{t}s_k)). \tag{6.46}$$

Finally, from (6.40)–(6.42) and (6.46), we can obtain that, for all $j \in [\omega_k]$,

$$\begin{aligned}
&\Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k)) \\
\overset{\text{Lemma 6.4.1 (vii)}}{\geq}\ &\Delta_{-K}(m_k^{a_j^k}(0)) - \Delta_{-K}(m_k^{a_j^k}(s_k)) \\
\geq\ & -\frac{\beta}{2}\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\|s_k\|}\min\left\{-\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\|s_k\|\mathcal{K}_1}, \Omega_k\right\} > 0,
\end{aligned} \tag{6.47}$$

which concludes the proof. □

Next, analogous to single and multi-objective trust-region schemes, we present a result about the accuracy of model functions $m^{a_j^k}$ in approximating their respective objectives $f^{a_j^k}$ in terms of the oriented distance function.

Theorem 6.5 *If Assumptions 6.6.1-6.6.6 are true, then for every iteration $k \in \mathbb{N}$ and for all $j \in [\omega_k]$, it holds that*

$$|\Delta_{-K}(f^{a_j^k}(x_k + s_k) - m^{a_j^k}(x_k + s_k))| \leq \frac{1}{2}(\mathcal{K}_1 + \mathcal{K}_2)\Omega_k^2,$$

where $x_k + s_k \in \Omega_k$ and γ is a constant.

Proof: Because of Assumption 6.6.1, we can employ mean value theorem on the objective function $f^{a_j^k}$, for all $j \in [\omega_k]$, to get

$$f^{a_j^k}(x_k + s_k) = f^{a_j^k}(x_k) + \nabla f^{a_j^k}(x_k)^\top s_k + \tfrac{1}{2} s_k^\top \nabla^2 f^{a_j^k}(\xi_k) s_k, \tag{6.48}$$

for some ξ_k in the segment $[x_k, x_k + s_k]$, and likewise on the model function $m^{a_j^k}$, for all $j \in [\omega_k]$, to get

$$m^{a_j^k}(x_k + s_k) = f^{a_j^k}(x_k) + \nabla f^{a_j^k}(x_k)^\top s_k + \tfrac{1}{2} s_k^\top \nabla^2 f^{a_j^k}(\zeta_k) s_k, \tag{6.49}$$

for some ζ_k in the segment $[x_k, x_k + s_k]$. Applying Δ_{-K} on the subtraction between (6.48) from (6.49) and taking the absolute value yields

$$\begin{aligned}
|\Delta_{-K}(f^{a_j^k}(x_k + s_k) - m^{a_j^k}(x_k + s_k))| &= \tfrac{1}{2}|\Delta_{-K}(s_k^\top \nabla^2 f^{a_j^k}(\xi) s_k - s_k^\top \nabla^2 m^{a_j^k}(\zeta) s_k)| \\
&\leq \tfrac{1}{2}\|s_k^\top \nabla^2 f^{a_j^k}(\xi) s_k - s_k^\top \nabla^2 m^{a_j^k}(\zeta) s_k\| \\
&\leq \tfrac{1}{2}\|s_k^\top \nabla^2 f^{a_j^k}(\xi) s_k\| + \|s_k^\top \nabla^2 m^{a_j^k}(\zeta) s_k\| \\
&\leq \tfrac{1}{2}(\|\nabla^2 f^{a_j^k}(\xi)\|\|s_k\|^2 + \|\nabla^2 m^{a_j^k}(\zeta)\|\|s_k\|^2) \\
&\leq \tfrac{1}{2}(\mathcal{K}_1 + \mathcal{K}_2)\Omega_k^2, \qquad (6.50)
\end{aligned}$$

by consecutively using Lemma 6.4.1 (i), Assumption 6.6.2, Triangle and Cauchy-Schwarz inequalities, and the fact that $x_k + s_k \in \mathcal{B}_k$ implies that $\|s_k\| \leq \Omega_k$. This concludes the proof. □

From above, we see that the error between objective $f^{a_j^k}$ and model $m^{a_j^k}$ is proportional to the square of the trust-region radius. More precisely, $m^{a_j^k}$ represents the objective $f^{a_j^k}$ very well within a sufficiently small trust-region such that a decrease in $m^{a_j^k}$ will also minimize $f^{a_j^k}$.

In the next result, we show that for a non-critical iterate x^k if trust-region radius falls below the fraction $\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)$, then iteration k is considered very successful (Case 2 in Subsection 6.5.4), step s_k is accepted and trust-region radius Ω_k either remains the same or increases.

Theorem 6.6 *Under Assumption 6.6.2, for each iteration $k \in \mathbb{N}$, if x_k is non-critical, s_k is the solution of the subproblem (6.18) with $\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k) \neq 0$, and*

$$\Omega_k \leq \frac{\beta|\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|(1 - \eta_2)}{2\mathcal{K}_1\|s_k\|} \tag{6.51}$$

for all $j \in [\omega_k]$, then iteration k is very successful and $\Omega_{k+1} \geq \Omega_k$.

Proof: Since $0 < \eta_2 < 1$ and $0 < \beta < 1$, we have $\frac{\beta}{2}(1-\eta_2) < 1$. Using Assumption 6.6.2 and (6.51), we have

$$\Omega_k < \frac{|\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|}{\mathcal{K}_1 \|s_k\|} \ \forall\ j \in [\omega_k]. \tag{6.52}$$

Further, since x_k is non-critical, from (6.52) and Corollary 6.6.1, one can obtain

$$\begin{aligned} \Delta_{-K}(m_k^{a_j^k}(0) - m_k^{a_j^k}(s_k)) &\geq -\beta\tfrac{1}{2}\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\|s_k\|}\Omega_k \ \forall\ j \in [\omega_k] \\ &= \beta\frac{1}{2}\frac{|\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|}{\|s_k\|}\Omega_k \ \forall\ j \in [\omega_k] \\ &= \mathcal{K}_1 > 0. \end{aligned} \tag{6.53}$$

Now, from the definition of reduction ratio $\rho^{a_j^k}$, we can write

$$\begin{aligned} 1 - \rho_k^{a_j^k} &= 1 + \frac{(\Delta_{-K}(-(f^{a_j^k}(x_k) - f^{a_j^k}(x_k + s_k)))}{(\Delta_{-K}(-m_k^{a_j^k}(s_k))} \\ &= \frac{\Delta_{-K}(-(f^{a_j^k}(x_k) - f^{a_j^k}(x_k + s_k))) + \Delta_{-K}(-m_k^{a_j^k}(s_k))}{(\Delta_{-K}(-m_k^{a_j^k}(s_k))}. \end{aligned} \tag{6.54}$$

Here, applying mean value theorem on $f^{a_j^k}$ around x_k, for all $j \in [\omega_k]$, with $\xi = x_k + \lambda s_k$ and $\lambda \in (0,1)$, we can write

$$f^{a_j^k}(x_k + s_k) = f^{a_j^k}(x_k) + \nabla f^{a_j^k}(x_k)^\top s_k + \tfrac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(\xi)s_k.$$

Then, using above and (vii) of Lemma 6.4.1 in (6.54), we obtain

$$\begin{aligned} 1 - \rho_k^{a_j^k} &= \frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k + \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(\xi)s_k) + \Delta_{-K}(-m_k^{a_j^k}(s_k))}{(\Delta_{-K}(-m_k^{a_j^k}(s_k))} \\ &\leq \frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k + \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(x_k)s_k) + \Delta_{-K}(-\nabla f^{a_j^k}(x_k)^\top s_k - \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(\xi_k)s_k)}{-(\Delta_{-K}(m_k^{a_j^k}(s_k))} \\ &= \frac{-\Delta_{-K}(-\nabla f^{a_j^k}(x_k)^\top s_k - \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(\xi)s_k) - \Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k + \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(x_k)s_k)}{(\Delta_{-K}(m_k^{a_j^k}(s_k))}. \\ &\leq \frac{-\Delta_{-K}(-\nabla f^{a_j^k}(x_k)^\top s_k - \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(\xi)s_k + \nabla f^{a_j^k}(x_k)^\top s_k + \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(x_k)s_k)}{(\Delta_{-K}(m^{a_j^k}(s_k))} \\ &\leq \frac{|\Delta_{-K}(\frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(x_k)s_k - \frac{1}{2}s_k^\top \nabla^2 f^{a_j^k}(\xi)s_k)|}{\Delta_{-K}(m^{a_j^k}(0)) - \Delta_{-K}(m^{a_j^k}(s_k))}. \end{aligned}$$

Next, by using property (i) of Lemma 6.4.1, Cauchy-Schwartz inequality, and (6.53), we get

$$1 - \rho_k^{a_j^k} \leq \frac{1}{2} \frac{\|s_k\|^2 \|\nabla^2 f^{a_j^k}(x_k) - \nabla^2 f^{a_j^k}(\xi)\|}{(-\Delta_{-K}(m^{a_j^k}(s_k))} \leq \frac{2\Omega_k \|s_k\| \mathcal{K}_1}{\beta |\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|} < 1 - \eta_2,$$

which gives $\rho_k^{a_j^k} > \eta_2$ for all $j \in [\omega_k]$. Hence, as per Case 2 of Step 10: of Algorithm 3, k is a very successful iteration with step s_k accepted and trust-region radius updated with $\Omega_{k+1} \geq \Omega_k$. □

In the next lemma, we use Theorem 6.6 above to show that the trust-region radius cannot shrink to zero for non-critical points and Ω_k is bounded below by a positive constant. More precisely, we show that it cannot simultaneously hold that x_k is non-critical and $\Omega_k \to 0$ as $k \to \infty$.

Lemma 6.2 *Under Assumption 6.6.2, if for every iteration $k \in \mathbb{N}$ and for every $j \in [\omega_k]$, there exists a σ such that $0 < \sigma < \frac{|\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|}{\|s_k\|}$, then there exists a constant $\Omega_{\min}$ such that $\Omega_k > \Omega_{\min}$.*

Proof: We prove this by contradiction. Let us assume a subsequence of trust-region radius $\{\Omega_{k_i}\}$ such that $\Omega_{k_i} \to 0$ i.e., for every $\epsilon > 0$ there exists an index $k_0 \in \mathbb{N}$ such that $\Omega_{k_i} < \epsilon$ for all $i \geq k_0$. Then, k_0 is the first iteration with $\Omega_k < \epsilon$ and, since $\Omega_{k_i} \to 0$, $\Omega_{k_0} < \Omega_{k_0-1}$. Also, according to trust-region update of Case 3 in Step 10: of Algorithm 3, we have $\gamma_1 \Omega_{k_0-1} \leq \Omega_{k_0}$. Then, if we consider

$$\epsilon = \frac{\gamma_1 \sigma \beta (1 - \eta_2)}{2\mathcal{K}_1}$$

with the parameters $\gamma_1 \in (0, 1)$ from Algorithm 6.6.1, β from (6.45), and $\mathcal{K}_1$ from Assumption 6.6.2, we have, for each $j \in [\omega_{k_0-1}]$,

$$\Omega_{k_0-1} \leq \frac{\Omega_{k_0}}{\gamma_1} < \frac{\epsilon}{\gamma_1} = \frac{\sigma\beta(1-\eta_2)}{2\mathcal{K}_1} \leq \frac{\beta(1-\eta_2)}{2\mathcal{K}_1} \frac{|\Delta_{-K}(\nabla f^{a_j^{k_0-1}}(x_{k_0-1})^\top s_{k_0-1})|}{\|s_{k_0-1}\|},$$

where the rightmost inequality is due to the hypothesis of this lemma. So, from Theorem 6.6, $k_0 - 1$ iteration was very successful and the trust-region radius must increase, i.e., $\Omega_{k_0-1} \leq \Omega_{k_0}$. However, as per our assumption at the beginning of the proof $\Omega_{k_0} \leq \Omega_{k_0-1}$, which is a contradiction. Therefore, for all k, we must have

$$\Omega_k > \frac{\gamma_1 \sigma \beta (1 - \eta_2)}{2\mathcal{K}_1} = \Omega_{\min}.$$

This conclude the proof. $\square$

Now, the earlier two results (Theorem 6.6 and Lemma 6.2) will allow us to establish the criticality of the limiting point of the sequence of iterates when the number of successful iterations is finite. For this purpose, let us first denote the collection of all successful iterations for Algorithm 3 by

$$\mathcal{S} = \{k \geq 0 : \rho_k^{a_j^k} \geq \eta_1 \; \forall \; j \in [\omega_k] \text{ and } \exists \; l \; \in [\omega_k] \text{ s.t } \rho_k^{a_l^k} < \eta_2\}. \tag{6.55}$$

.

Theorem 6.7 *Suppose that Assumptions 6.6.1-6.6.6, Theorem 6.6 and Lemma 6.2 hold, and there are only finitely many successful iterations in $\mathcal{S}$. Then, $x_k = x^*$ for all sufficiently large k, and x^* is critical.*

Proof: We prove this by contradiction. Let us assume k_0 is the last successful iteration and every iteration after that is unsuccessful, i.e. for all $k > k_0$ and for all $j \in [\omega_k]$, $\rho_k^{a_j^k} < \eta_1$. Then, as per Step 9: of Algorithm 6.6.1 we get all the future unsuccessful iterates to be the same i.e. $x_{k_0+1} = x_{k_0+q}$ for all $q > 1$, and let us denote these iterates by x^*. Further, since all iterations are unsuccessful for sufficiently large k, as per the trust-region update of Case 3 (unsuccessful iteration) of Step 10: of Algorithm 6.6.1, we get $\{\Omega_k\}$ converging to zero i.e., for every $\epsilon > 0$, there exists k_0 such that $\Omega_k < \epsilon$ for all $k > k_0$.

Suppose x_{k_0+1} is non-critical. Then, by Definition 6.4, $|\Delta_{-K}(\nabla f^{a_j^{k_0+1}}(x_{k_0+1})^\top s_{k_0+1})|$ $\neq 0$ for all $j \in [\omega_{k_0+1}]$. Letting $\epsilon = \frac{\beta|\Delta_{-K}(\nabla f^{a_j^{k_0+1}}(x_{k_0+1})^\top s_{k_0+1})|(1-\eta_2)}{2\mathcal{K}_1\|s_{k_0+1}\|}$, we get, for each $j \in [\omega_{k_0+1}]$,

$$\Omega_{k_0+1} < \frac{\beta|\Delta_{-K}(\nabla f^{a_j^{k_0+1}}(x_{k_0+1})^\top s_{k_0+1})|(1-\eta_2)}{2\mathcal{K}_1\|s_{k_0+1}\|}.$$

As a consequence, we can apply Theorem 6.6 to show that k_0+1 is a successful iteration. However, this is in contradiction to x_{k_0} being the last successful iterate as we assumed in the beginning. This proves that x^* is critical point. $\square$

In the preceding theorem, we studied the case where Algorithm 3 produces a finite number of successful iterations and showed its convergence to a critical point. Moving forward, we now address the case of infinitely many successful iterations and show that in that case as well Algorithm 3 terminates to a critical point. To begin this, we present a key lemma as follows.

Lemma 6.6.1 *Suppose that Assumptions 6.6.1-6.6.6 and Corollary 6.6.1 hold, and Algorithm 3 generates infinitely many successful iterations $k \in \mathcal{S}$. Then, we have*

$$\liminf_{k\to\infty} \max_{j\in[\omega_k]} \left\{ \frac{|\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|}{\|s_k\|} \right\} = 0. \tag{6.56}$$

Proof: By the method of contradiction, let us assume that (6.56) is not true. This implies that there exists $\epsilon > 0$ such that $\forall k$ there exists $j \in [\omega_k]$ such that

$$\frac{|\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|}{\|s_k\|} \geq \epsilon. \tag{6.57}$$

Now, let k be a successful iteration and (t_k, s_k) be the solution of the subproblem (6.18). Then, for all $j \in [\omega_k]$, from (6.55), we have

$$\rho_k^{a_j^k} = \frac{-(\Delta_{-K}(-(f^{a_j^k}(x_k) - f^{a_j^k}(x_{k+1}))))}{(\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k)))} \geq \eta_1,$$

and using Lemma 6.4.1 (vii), we get

$$\frac{\Delta_{-K}(f^{a_j^k}(x_k)) - \Delta_{-K}(f^{a_j^k}(x_{k+1}))}{\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))} \geq \eta_1.$$

Then, from the sufficient decrease condition for the model m^{a^k} of f^{a^k} in Theorem 6.4, and using (6.46) and (6.57), we have

$$\begin{aligned}
\Delta_{-K}(f^{a_j^k}(x_k)) - \Delta_{-K}(f^{a_j^k}(x_{k+1})) \geq\ & \eta_1(\Delta_{-K}(m^{a_j^k}(0) - m^{a_j^k}(s_k))) \\
\geq\ & \eta_1\beta\Bigg(-\tfrac{1}{2}\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\|s_k\|} \\
& \min\left\{ -\frac{\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)}{\|s_k\|\mathcal{K}_1}, \Omega_k \right\} \Bigg) \\
\geq\ & \eta_1\epsilon\beta\tfrac{1}{2}\min\left\{ \frac{\epsilon}{\mathcal{K}_1}, \Omega_{\min} \right\}.
\end{aligned}$$

For every successful iteration, it holds $x_{k+1} = x_k + s_k$. Thus, summing over all the successful iterations k, we get

$$\begin{aligned}
\Delta_{-K}(f^{a_j^k}(x_0) - f^{a_j^k}(x_{k+1}))\ & \geq \Delta_{-K}(f^{a_j^k}(x_0)) - \Delta_{-K}(f^{a_j^k}(x_{k+1})) \text{ by Lemma 6.4.1 ((vii))} \\
& = \sum_{l=0, l\in\mathcal{S}}^{k} (\Delta_{-K}(f^{a_j^k}(x_l)) - \Delta_{-K}(f^{a_j^k}(x_{l+1})))
\end{aligned}$$

$$\geq \sum_{l=0, l\in\mathcal{S}}^{k} \eta_1 \epsilon \beta \tfrac{1}{2} \min\left\{\frac{\epsilon}{\mathcal{K}_1}, \Omega_{\min}\right\}$$
$$\geq \alpha_k \eta_1 \epsilon \beta \tfrac{1}{2} \min\left\{\frac{\epsilon}{\mathcal{K}_1}, \Omega_{\min}\right\},$$

where α_k is the number of successful iterations up to the k-th iteration. Now, for infinitely many successful iterations, we should have

$$\lim_{k\to\infty} \alpha_k = \infty,$$

which, from Definition 2.1 in [5], either means that $(f^{a_j^k}(x_0)) - (f^{a_j^k}(x_{k+1}))$ is infinite or $-K = \emptyset$. The former is in contradiction with Assumption 6.6.4, and the latter contradicts the fact that K is a solid cone. Therefore, there cannot exist any $\epsilon > 0$ satisfying (6.57), which concludes the proof.

$$\liminf_{k\to\infty} \max_{j\in[\omega_k]} \left\{ \frac{|\Delta_{-K}(\nabla f^{a_j^k}(x_k)^\top s_k)|}{\|s_k\|} \right\} = 0.$$

□

Finally, we present the global convergence proof of our Algorithm 3 when it generates infinitely many successful iterations.

Theorem 6.8 *Let $\{x_k\}$ be a sequence of regular iterative points for F generated by Algorithm 3. Further, assume that the level set of the function F is bounded. Then $\{x_k\}$ converges to a critical point for ($\mathcal{SOP}^l_K$).*

Proof: Recall that the level set of F, $\mathcal{L}_c = \{x \in \mathbb{R}^n : F(x) \preceq^l_K F(x_0)\}$ is bounded. Since Algorithm 3 is $\preceq^l_k$-decreasing method and from (6.26), we have $x_1 \in \mathcal{L}_c$. By induction, we prove $x_{k+1} \in \mathcal{L}_c$ if $x_k \in \mathcal{L}_c$ for all $k > 1$. Using (6.26), we have

$$k \in \mathbb{N} \cup \{0\} : F(x_{k+1}) = F(x_k + s_k) \preceq^l_k F(x_k) \preceq^l_K F(x_0).$$

From transitivity of the relation $\preceq^l_k$, we can write $F(x_{k+1}) \preceq^l_k F(x_0)$. Then, the sequence $\{x_k\}$ generated by $\mathcal{L}_c$, is contained in $\mathcal{L}_c$. Since $\mathcal{L}_c$ is bounded, $\{x_k\}$ is bounded sequence as well. By Bolzano-weiertrass theorem in $\mathbb{R}^n$, $\{x_k\}$ has a convergent subsequence $\{x_{k_r}\}$ which converges $\bar{x}$. Now, in order to show that the set $\mathcal{L}_c$ is closed, let us recall $\{x_{r_k}\} \subseteq \mathcal{A}$ such that $x_{r_k} \to \bar{x}$. Then. we have

$$F(x_{r_k}) \preceq^l_K F(x_0)$$

$$\begin{aligned}
&\Longrightarrow F(\bar{x}) \subseteq F(x_{r_k}) + K\\
&\Longrightarrow \{f^i(x_0)\}_{i\in[p]} \subseteq \{f^i(x_{r_k})\}_{i\in[p]} + K\\
&\Longrightarrow \forall i \in [p], \exists j \in [p] : f^j(x_{r_k}) \preceq_K f^i(x_0)\\
&\Longrightarrow \forall i \in [p], \exists j \in [p] : f^j(x_{r_k}) - f^i(x_0) \in -K\\
&\Longrightarrow \forall i \in [p], \exists j \in [p] : \Delta_{-K}(f^j(x_{r_k}) - f^i(x_0)) \leq 0. \qquad (6.58)
\end{aligned}$$

Since f^i, for all $i \in [p]$, are continuous and Δ_{-K} is Lipschitz continuous (see Lemma 6.4.1 (i), taking the limit in (6.58) as $k \to \infty$, we get

$$\begin{aligned}
&\forall i \in [p], \exists j \in [p] : \Delta_{-K}(f^j(\bar{x}) - f^i(x_0)) \leq 0\\
&\Longrightarrow \forall i \in [p], \exists j \in [p] : f^j(\bar{x}) \preceq_K f^i(x_0) \qquad (6.59)\\
&\Longrightarrow F(\bar{x}) \preceq^l_K F(x_0). \qquad (6.60)
\end{aligned}$$

Thus, $\bar{x} \in \mathcal{L}_c$. Hence, $\mathcal{L}_c$ is compact. Therefore, the sequence $\{x_k\}$ has a convergent subsequence $Q = \{x_{r_k}\}$ in $\mathcal{A}$ with limiting point $\bar{x}$, namely. $x_k \xrightarrow{\mathcal{Q}} \bar{x}$.

Next, we need to show that $\{s_k\}_{k\in\mathcal{Q}}$ is bounded. Since x_k is a successful iterate for $(\mathcal{SOP}^l_K)$, we get $x_{k+1} = x_k + s_k$ and $s_k \neq 0$ is a descent direction. Thus, for all $j \in [\omega_k]$,

$$\Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top s_k) < 0.$$

Since $\{x_k\}$ is bounded. So, $\{x_{k+1}\}$ is also bounded. Since $s_k = x_{k+1} - x_k$, we get $\{s_k\}$ to be bounded in $\mathcal{Q}$ as well. Again, by Bolzano-weierstrass theorem in $\mathbb{R}^n$, $\{s_k\}$ must have a convergent subsequence in Q, which converging to $\bar{s}$, namely $s_k \xrightarrow{\mathcal{Q}} \bar{s}$.
Further, since there are only a finite number of subsets of $[p]$ and $\{x_k\}$ is a sequence of regular points for F, we can apply Lemma 1.2.1 to obtain, without loss of generality, the existance of $\mathcal{Q}' \subseteq P_k$ and $\bar{a} \in \mathcal{Q}'$ such that

$$\forall k \in \mathcal{Q} : \omega_k = \bar{\omega}, a^k = \bar{a}.$$

Since $\{x_k\}$ and $\{s_k\}$ are generated by Algorithm 3, and also due to the fact $f^{a^k_j}, \forall j \in [\omega_k]$, are continuous, we can obtain, using Lemma 6.6.1 that

$$\begin{aligned}
\liminf_{k \xrightarrow{\mathcal{Q}'} \infty} \max_{j\in[\omega_k]} \left\{ \frac{|\Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top s_k)|}{\|s_k\|} \right\} &= \liminf_{k \xrightarrow{\mathcal{Q}'} \infty} \max_{j\in[\omega_k]} \left\{ \frac{\Delta_{-K}(\nabla f^{a^k_j}(x_k)^\top s_k)}{\|s_k\|} \right\}\\
&\leq \Delta_{-K} \left(\max_{j\in[\omega_k]} \left\{ \liminf_{k \xrightarrow{\mathcal{Q}'} \infty} \frac{(\nabla f^{a^k_j}(x_k)^\top s_k)}{\|s_k\|} \right\} \right)
\end{aligned}$$

$$= \Delta_{-K}\left(\max_{j\in[\bar{\omega}]}\left\{\frac{(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s})}{\|\bar{s}\|}\right\}\right),$$

where the first inequality follows from inequality 1(3) in [158]. Then, from Lemma 6.6.1, it follows that

$$\Delta_{-K}\left(\max_{j\in[\bar{\omega}]}\left\{\frac{(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s})}{\|\bar{s}\|}\right\}\right) \geq 0,$$

i.e., $\max_{j\in[\bar{\omega}]}\left\{\frac{(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s})}{\|\bar{s}\|}\right\} \notin -K$ by (iii) and (iv) of Lemma 6.4.1. As a result, for $\bar{s} \in \mathbb{R}^n$, there exists $j \in [\bar{\omega}]$ such that $\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s} \notin -K$ i.e, by (iii) and (iv) of Lemma 6.4.1

$$\Delta_{-K}(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s}) \geq 0. \tag{6.61}$$

On the other hand, since $x_k \to \bar{x}$ and θ is continuous map on the set of all regular points, we have

$$\theta(\bar{x}) \leq 0. \tag{6.62}$$

Thus,

$$\begin{aligned}
0 &\overset{(6.61)}{\leq} \Delta_{-K}(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s}) \\
&\leq \max_{j\in[\bar{\omega}]}\left\{\Delta_{-K}(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s}), \Delta_{-K}\left(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{s} + \frac{1}{2}\bar{s}^\top \nabla^2 f^{\bar{a}_j}(\bar{x})\bar{s}\right)\right\} \\
&= \Theta_{\bar{x}}(\bar{a}, \bar{s}) = \min_{(\bar{a},s)\in\mathcal{Q}'\times\mathbb{R}^n} \Theta_{\bar{x}}(a, s) \\
&= \theta(\bar{x}) \leq 0,
\end{aligned}$$

which implies $\theta(\bar{x}) = 0$. Therefore, from Theorem 6.3, $\bar{x}$ must be a critical point. This completes the proof. □

6.7 Numerical experiment

In this section, we analyze performance of Algorithm 3 using a few test instances. The experiments are conducted in Matlab on a system equipped with a 12 CPUs of Intel(R) Core(TM) i7-9850H CPU processor running at 2.60 GHz, with 64 GB RAM and Windows 10 OS. Each test case is executed multiple times, and we report the statistic including min, max, median, mean, variance (Var), and standard deviation (SD) for the number of iterations and the execution time taken by the algorithm to reach the stopping condition (see Step 6:). The parameters used in our experiments are detailed in Table 6.1. Throughout all experiments, except for Example 6.7.3(b), the cone remain consistent as $K = \mathbb{R}^m_+$. In Example 6.7.3(b), the cone is specified as $K = \{(y_1, y_2) \in \mathbb{R}^2 : y_2 \geq 2y_1 \text{ and } y_2 \leq 4y_1\}$.

Table 6.1: Used parameter values for Algorithm 3 on all the examples

Ω_0	$\Omega_{\max}$	ϵ	η_1	η_2	γ_1	γ_2
1	20	0.1	0.001	0.75	0.4	0.9

Example 6.7.1 *In this example, we consider a set optimization problem of two variables and 10000 functions, each having three components. The objective function of the SOP is the set-valued map* $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^3$ *defined by*

$$F(x_1, x_2) := \{f^1(x_1, x_2), f^2(x_1, x_2), \ldots, f^{10000}(x_1, x_2)\},$$

where for each $i \in [10000]$, $f^i : \mathbb{R}^2 \to \mathbb{R}^3$ *is given by*

$$f^i(x_1, x_2) = g(x_1, x_2) + (x_1^2 + x_2^4) \begin{pmatrix} (2.1 + x_1^2 + \cos\frac{\phi_{1i}}{2}\sin\phi_{2i} - \sin\frac{\phi_{1i}}{2}\sin 2\phi_{2i})\cos\phi_{1i} \\ (2.1 + x_1^2 + \cos\frac{\phi_{1i}}{2}\sin\phi_{2i} - \sin\frac{\phi_{1i}}{2}\sin 2\phi_{2i})\sin\phi_{1i} \\ \sin\frac{\phi_{1i}}{2}\sin 2\phi_{2i} + \cos\frac{\phi_{1i}}{2}\sin 2\phi_{2i} \end{pmatrix}$$

with

$$g(x_1, x_2) = 100 \left[\begin{pmatrix} -x_1 \\ x_1 + x_2^2 \\ -x_1 \end{pmatrix} + 100(x_1^2 + x_1 + x_2 - 3) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right]$$

and the set $\{(\phi_{1i}, \phi_{2i}) : i \in [10000]\}$ *is an enumeration of the set* $\{\frac{\pi}{50}(j-1) : j \in [100]\} \times \{\frac{\pi}{50}(\ell - 1) : \ell \in [100]\}$.

For this problem, we ran Algorithm 3 for randomly chosen 100 *initial points taken from* $[-25, 25] \times [-25, 25]$. *Algorithm successfully converged for* 71 *out of* 100 *initial points within fewer than* 50 *iterations. The number of iterations, amount of execution time are reported in Table 6.2.*

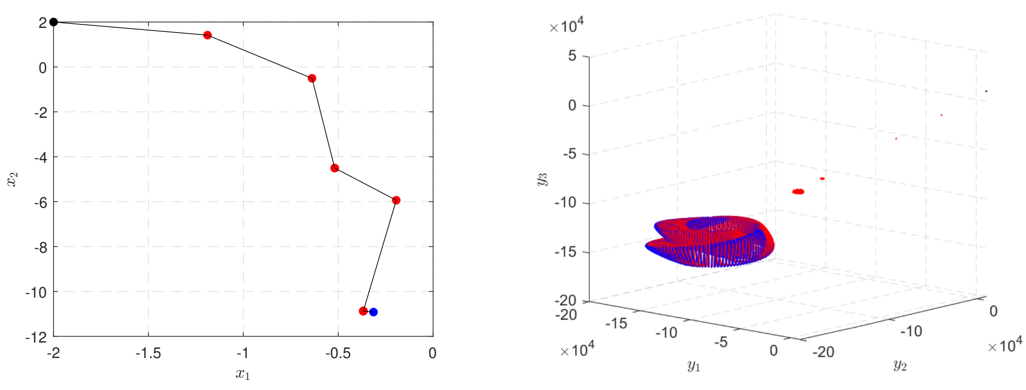
For the initial point $(-2, 2)$, *the intermediate iterates of Algorithm 3 in the variable space are shown in Figure 6.2(a) by red bullet points. Correspondingly, the values of* F *in the image space are represented by red regions in Figure 6.2(b). In Figure 6.2(a), the black bullet point signifies the initial point, while the blue bullet point denotes the termination point* $(-0.314, -10.920)$. *The black and blue regions in Figure 6.2(b) are* F*-values at the initial and terminal points, respectively. Zoomed-in version of the* F*-values at the initial point* $(-2, 2)$ *and at two intermediate iterates* $(-1.189, 1.415)$ *and*

$(-0.638, -0.507)$ *are presented in Figure 6.3.*

The iterative points $x^k = (x_1^k, x_2^k)$*, executed by Algorithm 3 for the initial point* $(-2, 2)$*, are explicitly provided in Table 6.3. Alongside, the function values of three sample functions* f^{10}*,* f^{100} *and* f^{300} *across all the iterative points are also reported in Table 6.3. It is to observe from Table 6.3 that although for the initial point* $(-2, 2)$*, Algorithm 3 takes* 15 *iterations before it terminates, there are just five distinct intermediate iterates in between the initial and terminal iterate. At the iteration counter* (k) *where no new iterate is generated (e.g.,* $k = 4, 7, 8, 9, 10, 11, 12, 13, 14$*), Algorithm 3 successively changes the trust-region radius without altering the current iterate. It is noteworthy from the last three columns of Table 6.3 that all the component functions of all the functions in* F *successively degrade across the iterates from the initial to the terminal iterate.*

Table 6.2: Performance of Algorithm 3 on Example 6.7.1

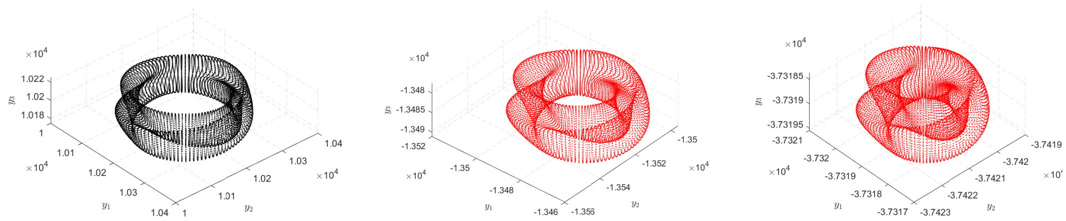
Number of initial points	Number of Iterations (Min, Max, Mean, Var, Median, SD)	CPU Time (in sec) (Min, Max, Mean, Median, Var, SD)
71	$(4, 50, 27.32, 227.79, 32, 15.09)$	$(194.38, 2554.63, 1358.47, 1535.18, 569051.42, 754.35)$



(a) Variable space

(b) Image space

Figure 6.2: Generated iterative points by Algorithm 3 on Example 6.7.1 with the initial point as $(-2, 2)$



(a) $F(-2,2)$ (b) $F(-1.189, 1.415)$ (c) $F(-0.638, -0.507)$

Figure 6.3: The initial point and next three consecutive points in the image space of Figure 6.2(b)

Table 6.3: Iterative points in the argument space for Example 6.7.1

Initial point	Iteration number (k)	(x_1^k, x_2^k)	$f^{10}(x_1^k, x_2^k)$	$f^{100}(x_1^k, x_2^k)$	$f^{300}(x_1^k, x_2^k)$
$(-2, 2)$	1	$(-1.189, \quad 1.415)$	$10^5 \times (-0.135, -0.135, -0.135)$	$10^5 \times (-0.135, -0.135, -0.135)$	$10^5 \times (-0.135, -0.135, -0.135)$
	2	$(-0.638, -0.507)$	$10^5 \times (-0.373, -0.374, -0.373)$	$10^5 \times (-0.373, -0.374, -0.373)$	$10^5 \times (-0.373, -0.374, -0.373)$
	3	$(-0.518, -4.506)$	$10^5 \times (-0.765, -0.756, -0.776)$	$10^5 \times (-0.773, -0.746, -0.775)$	$10^5 \times (-0.776, -0.765, -0.775)$
	4	$(-0.518, -4.506)$	$10^5 \times (-0.765, -0.756, -0.776)$	$10^5 \times (-0.773, -0.746, -0.775)$	$10^5 \times (-0.776, -0.765, -0.775)$
	5	$(-0.194, -5.937)$	$10^5 \times (-0.881, -0.874, -0.913)$	$10^5 \times (-0.903, -0.849, -0.909)$	$10^5 \times (-0.911, -0.901, -0.909)$
	6	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	7	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	8	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	9	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	10	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	11	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	12	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	13	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	14	$(-0.368, -10.873)$	$10^5 \times (-1.075, -1.293, -1.456)$	$10^5 \times (-1.334, -0.990, -1.410)$	$10^5 \times (-1.436, -1.604, -1.410)$
	15	$(-0.314, -10.920)$	$10^5 \times (-1.077, -1.295, -1.459)$	$10^5 \times (-1.336, -9.915, -1.413)$	$10^5 \times (-1.439, -1.606, -1.413)$

Example 6.7.2 *In this example, we consider the Test Instance 5.2 in [18], which has come from the robust counterpart of a vector-valued facility location problem with the points in the mesh* $A = P \times P$ *of* 100 *points being the uncertainty set with respect to the facility sites*

$$b_1 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad b_2 := \begin{pmatrix} 8 \\ 0 \end{pmatrix}, b_3 := \begin{pmatrix} 0 \\ 8 \end{pmatrix},$$

where P *is a uniform partition of* 10 *points of the interval* $[-1, 1]$*, i.e.,*

$$P = \{-1, -0.7778, -0.5556, -0.3333, -0.1111, 0.1111, 0.3333, 0.5556, 0.7778, 1\}.$$

Let $\{a_1, a_2, \ldots, a_{100}\}$ *be an enumeration of* A*, and consider the* 100 *functions as follows.*

For $i \in [100]$, define $f^i : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$f^i(x) := \frac{1}{2}\begin{pmatrix} \|x - a_i - b_1\|^2 \\ \|x - a_i - b_2\|^2 \\ \|x - a_i - b_3\|^2. \end{pmatrix}$$

Finally, we consider the SOP with the objective function as the set-valued mapping $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^3$, defined by

$$F(x) := \{f^1(x), f^2(x), \ldots, f^{100}(x)\}.$$

The local weakly-minimal solutions of this SOP (see [18]) lie in the convex hull of the set

$$(b_1 + P) \cup (b_2 + P) \cup (b_3 + P).$$

For this problem, we ran Algorithm 3 for randomly chosen 100 initial points taken from $[-50, 50] \times [-50, 50]$. The algorithm converged for 60 out of 100 initial points in less than 50 iterations. The number of iterations, amount of execution time are given in Table 6.4.

Solution found for 8 initial points (e.g., $x_0 = (-6, 8.95), (-3, 12), (0, 13), (4, 4), (-6, -2), (-2, -6), (12, -4), (9.5, -3)$) are shown in Figure 6.4. For these initial points, the intermediate iterates of Algorithm 3 in the variable space are shown by red bullet points. In Figure 6.4, the black bullet points are the initial points, denoted by $I_i, i \in [8]$, and the blue bullet points are the termination points, denoted by $T_i, i \in [8]$, for their corresponding initial points. The locations of b_1, b_2, b_3 are shown by relatively large black bullet points compared to initial black bullet points and the elements of the set $(b_1 + P) \cup (b_2 + P) \cup (b_3 + P)$ are depicted by gray bullet points. We can observe that all termination points are contained in the convex hull of the set $(b_1 + P) \cup (b_2 + P) \cup (b_3 + P)$. The convex hull is shaded by sky blue.

Among the 8 initial points, just for two initial points $(12, -4), (-2, -6)$, the iterative points $x^k = (x_1^k, x_2^k)$ is executed separately Algorithm 3 are explicitly provided by Table 6.7.2. Alongside, the function values of three sample functions f^5, f^{10} and f^{50} across all the iterative points are also reported in Table 6.5. It is noticeable that from the last three columns for each row corresponding to $(12, -4)$ and $(-2, -6)$ in Table 6.5 that all the component functions of all functions in F successively degrade across the iterates from the initial to the terminal iterate.

Table 6.4: Performance of Algorithm 3 on Test instance 6.7.2

Number of initial points	Iterations (Min, Max, Mean, Var, Median, SD)	CPU time (in sec) (Min, Max, Mean, Median, Var, SD)
60	$(1, 40, 11.1, 52.87, 8, 7.27)$	$(5.98, 406.32, 84.23, 57.70, 4827.49, 69.48)$

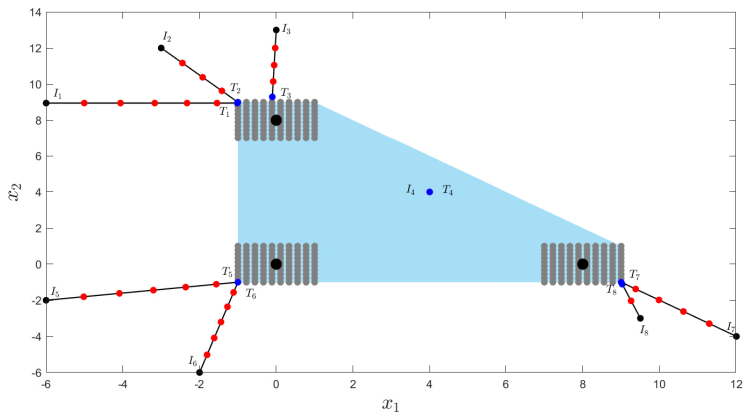


Figure 6.4: Solution found (in blue) for 8 initial points in the variable space for the Example 6.7.2

Table 6.5: Solution found in the argument space for Test instance 6.7.2

Initial point	Iteration number (k)	(x_1^k, x_2^k)	$f^5(x_1^k, x_2^k)$	$f^{10}(x_1^k, x_2^k)$	$f^{50}(x_1^k, x_2^k)$
$(12, -4)$	1	$(11.293, -3.293)$	$(80.620, 14.276, 138.074)$	$(84.772, 18.429, 151.115)$	$(74.240, 15.008, 140.583)$
	2	$(10.621, -2.621)$	$(70.676, 9.706, 122.756)$	$(74.082, 13.113, 135.051)$	$(64.147, 10.289, 125.116)$
	3	$(9.983, -1.983)$	$(62.065, 6.201, 109.040)$	$(64.762, 8.898, 120.626)$	$(55.394, 6.642, 111.258)$
	4	$(9.377, -1.377)$	$(54.639, 3.625, 96.764)$	$(56.663, 5.649, 107.676)$	$(47.834, 3.931, 98.848)$
	5	$(9.000, -1.000)$	$(50.395, 2.395, 89.506)$	$(52.000, 4.000, 100.000)$	$(43.506, 2.617, 91.506)$
$(-2, -6)$	1	$(-1.804, 5.019)$	$(12.369, 50.800, 83.635)$	$(18.440, 56.871, 98.595)$	$(19.549, 65.092, 99.705)$
	2	$(-1.618, 4.088)$	$(8.098, 45.039, 71.912)$	$(13.134, 50.074, 85.837)$	$(14.078, 58.130, 86.781)$
	3	$(-1.441, 3.203)$	$(4.877, 40.401, 61.611)$	$(8.929, 44.454, 74.552)$	$(9.716, 52.352, 75.339)$
	4	$(-1.272, 2.362)$	$(2.571, 36.750, 52.579)$	$(5.689, 39.869, 64.587)$	$(6.326, 47.617, 65.224)$
	5	$(-1.113, 1.563)$	$(1.061, 33.963, 44.680)$	$(3.292, 36.194, 55.800)$	$(3.787, 43.800, 56.295)$
	6	$(-1.000, 1.000)$	$(0.395, 32.395, 39.506)$	$(2.000, 34.000, 50.000)$	$(2.395, 41.506, 50.395)$

Example 6.7.3 *(a) In this case, we consider Test Instance 5.3 in [18] with a slight alteration, which is a two-variables 100-functions with two-components problem. For $i \in [100]$, the function $f^i : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by*

$$f^i(x_1, x_2) := \begin{pmatrix} e^{\frac{x_1}{2}} \cos x_2 + x_1 \cos x_2 \sin \frac{\pi(i-1)}{50} - x_2 \sin x_2 \cos^3 \frac{\pi(i-1)}{50} \\ e^{\frac{x_2}{20}} \sin x_1 + x_1 \sin x_2 \sin^3 \frac{\pi(i-1)}{100} + x_2 \cos x_2 \cos \frac{\pi(i-1)}{100}) \end{pmatrix}$$

with $\mathcal{S} = [-20, 20] \times [-20, 20]$.

For this problem, we ran Algorithm 3 *for randomly chosen* 100 *initial points taken from* $[-20, 20] \times [-20, 20]$. *Algorithm converged for* 85 *out of* 100 *initial points in less than* 50 *iterations. The number of iterations, amount of execution time are presented in Table 6.6.*

For the initial points $(7, 6)$ *and* $(9, 8)$, *the intermediate iterates in the variable space are shown in Figure 6.5(a) and Figure 6.6(a), respectively by red bullet points. The corresponding* F*-values in the image space are shown by red colored sets in the Figure 6.5(b) and Figure 6.6(b), respectively. In Figure 6.5(a) and 6.6(a), the black bullet points are the initial points, and the blue bullet points are the termination points* $(4.687, 6.456)$ *and* $(11.415, 8.238)$, *respectively. The black and blue colored set in Figure 6.5(b) and Figure 6.6(b) are* F*-values at the initial and terminal points, respectively.*

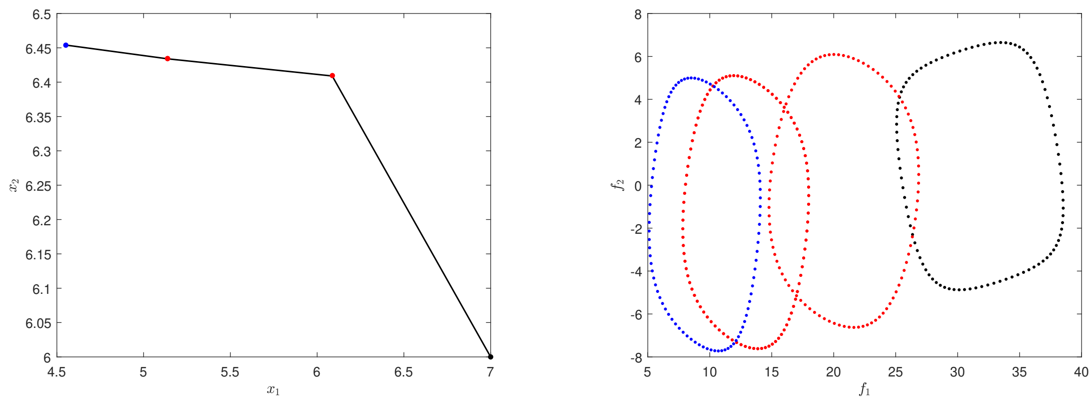
Iterative points $x^k = (x_1^k, x_2^k)$ *for both the initial points* $(7, 6)$ *and* $(9, 8)$, *are explicitly provided in Table 6.7. Alongside, the function values of three sample functions* f^1, f^{10} *and* f^{100} *across all the iterative points are reported in Table 6.7. From the last three columns for each row corresponding to* $(7, 6)$ *and* $(9, 8)$, *we note that all the component functions of all functions in* F *successively decrease across the iterates.*

Table 6.6: Performance of Algorithm 3 on Example 6.7.3 (a)

Number of initial points	Iterations (Min, Max, Mean, Var, Median, SD)	CPU time (in sec) (Min, Max, Mean, Median, Var, SD)
85	$(1, 17, 4.28, 14.84, 3, 3.85)$	$(4.08, 281.55, 46.72, 30.35, 2711.85, 52.07)$

Table 6.7: Solution found in the argument space for Example 6.7.3 (a)

Initial point	Iteration number (k)	(x_1^k, x_2^k)	$f^1(x_1^k, x_2^k)$	$f^{10}(x_1^k, x_2^k)$	$f^{100}(x_1^k, x_2^k)$
$(7, 6)$	1	$(6.079, 6.390)$	$(20.096, 6.075)$	$(23.606, 5.186)$	$(19.720, 6.062)$
	2	$(5.131, 6.441)$	$(11.833, 5.100)$	$(14.950, 4.234)$	$(11.521, 5.087)$
	3	$(4.687, 6.456)$	$(9.156, 4.979)$	$(12.071, 4.113)$	$(8.872, 4.967)$
$(9, 8)$	1	$(9.994, 8.113)$	$(-45.714, -2.885)$	$(-43.963, -1.075)$	$(-45.507, -2.883)$
	2	$(10.940, 8.199)$	$(-87.964, -4.275)$	$(-86.874, -2.260)$	$(-87.687, -4.273)$
	3	$(11.415, 8.238)$	$(-120.415, -4.464)$	$(-119.66, -2.356)$	$(-120.101, -4.461)$



(a) Variable space

(b) Image space

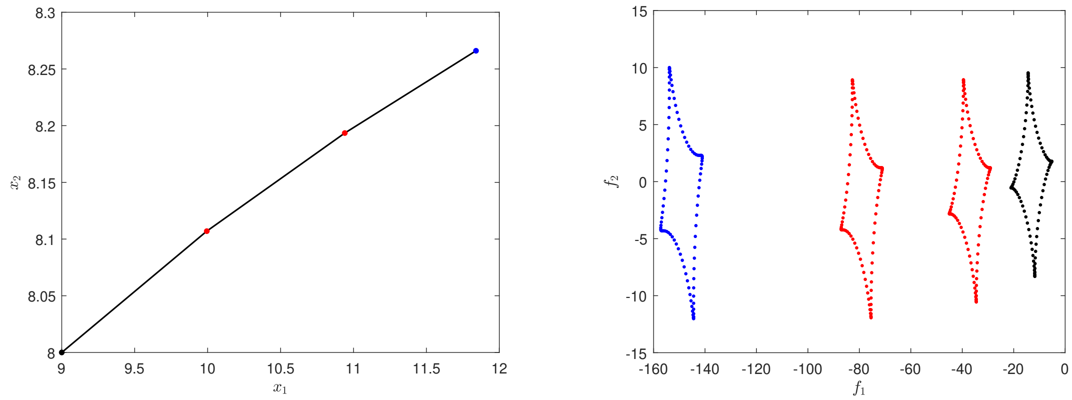
Figure 6.5: The generated iterative points by Algorithm 3 on Example 6.7.3 (a) with the initial point as $(7, 6)$

(b) *This case is exactly same as Example 6.7.3 (a) but with a different cone* $K = \{(y_1, y_2) \in \mathbb{R}^2 : y_2 \geq 2y_1$ *and* $y_2 \leq 4y_1\}$. *In this case, Algorithm converged for* 84 *out of* 100 *initial points in less than* 50 *iterations. The number of iterations, amount of execution time are presented in Table 6.8.*

For the initial points $(7, 6)$ *and* $(9, 8)$, *the intermediate iterates in the variable space are shown in Figure 6.7(a) and Figure 6.8(a), respectively by red bullet points. The corresponding* F*-values in the image space are shown by red colored sets in Figure 6.7(b) and Figure 6.8(b), respectively. In Figure 6.7(a) and Figure 6.8(a), the black bullet points are the initial points, and the blue bullet points are the termination points* $(5.051, 5.926)$ *and* $(10.950, 8.017)$, *respectively. The black and blue colored set in Figure 6.7(b) and Figure 6.8(b) are* F*-values at the initial and terminal points, respectively. Iterative points* $x^k = (x_1^k, x_2^k)$ *for both the initial points* $(7, 6)$ *and* $(9, 8)$ *are explicitly provided in Table 6.9. Alongside, the function values of three sample functions* f^1, f^{10} *and* f^{100} *across all the iterative points are reported in Table 6.9. From the last three columns for each row corresponding to* $(7, 6)$ *and* $(9, 8)$, *we note that all the component functions of all functions in* F *successively decrease across the iterates.*

Table 6.8: Performance of Algorithm 3 on Example 6.7.3 ((b))

Number of initial points	Iterations (Min, Max, Mean, Var, Median, SD)	CPU time (in sec) (Min, Max, Mean, Median, Var, SD)
84	$(1, 46, 4.94, 54.65, 1, 7.39)$	$(9.44, 1012.83, 100.13, 16.61, 26819.82, 163.76)$



(a) Variable space (b) Image space

Figure 6.6: The generated iterative points by Algorithm 3 on Example 6.7.3 (a) with the initial point as $(9, 8)$

Table 6.9: Solution found in the argument space for Example 6.7.3 ((b))

Initial point	Iteration number (k)	(x_1^k, x_2^k)	$f^1(x_1^k, x_2^k)$	$f^{10}(x_1^k, x_2^k)$	$f^{100}(x_1^k, x_2^k)$
$(7, 6)$	1	$(6.001, 5.968)$	$(20.951, 5.298)$	$(23.272, 4.129)$	$(20.582, 5.288)$
	2	$(5.051, 5.926)$	$(13.783, 4.284)$	$(15.494, 3.148)$	$(13.473, 4.273)$
$(9, 8)$	1	$(10.000, 8.010)$	$(-30.939, -2.055)$	$(-28.621, -0.341)$	$(-30.795, -2.055)$
	2	$(10.950, 8.017)$	$(-46.528, -2.789)$	$(-44.329, -0.924)$	$(-46.370, -2.789)$

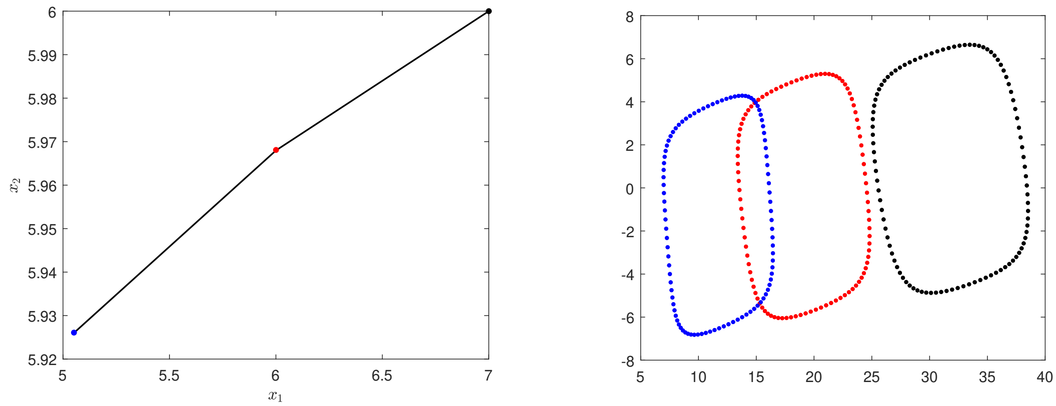
Example 6.7.4 *This example is an one-variable 50-functions with two-components problem. For $i \in [50]$, we consider the function $f^i : \mathbb{R} \to \mathbb{R}^2$ as*

$$f^i(x) := \begin{pmatrix} x + \frac{1}{128}(9 + e^{\sin\left(\frac{4\pi i}{50}\right)} - \sin\left(\frac{4\pi i}{50}\right) + 2(\cos^2(\frac{8\pi i}{50}))) \cos\left(\frac{2\pi i}{50}\right) \\ \cos(2x) + \frac{1}{1+e^{2x}} + \frac{1}{128}(9 + e^{\sin(\frac{4\pi i}{50})} - \sin\left(\frac{4\pi i}{50}\right) + 2\left(\cos^2(\frac{8\pi i}{50})\right)) \sin(\frac{2\pi i}{50}) \end{pmatrix}$$

with $\mathcal{S} = [-3, 6]$.

For this problem, Algorithm 3 is executed with 100 *randomly chosen initial points from* $[-3, 6]$. *The number of iterations and amount of execution time are reported in Table 6.10.*

Solution found for 8 *initial points (e.g., $x_0 = 0.5, 0.7, 3.5, 3.8, 4, 5.6, -0.5, 2.5, -0.09$) are shown in Figure 6.9. The F-values at initial points are successively denoted by $I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8$ and represented by black, blue, orange, purple, sky blue, yellow, cyan, green, magenta colored sets, respectively. In Figure 6.9, we can see that for*



(a) Variable space

(b) Image space

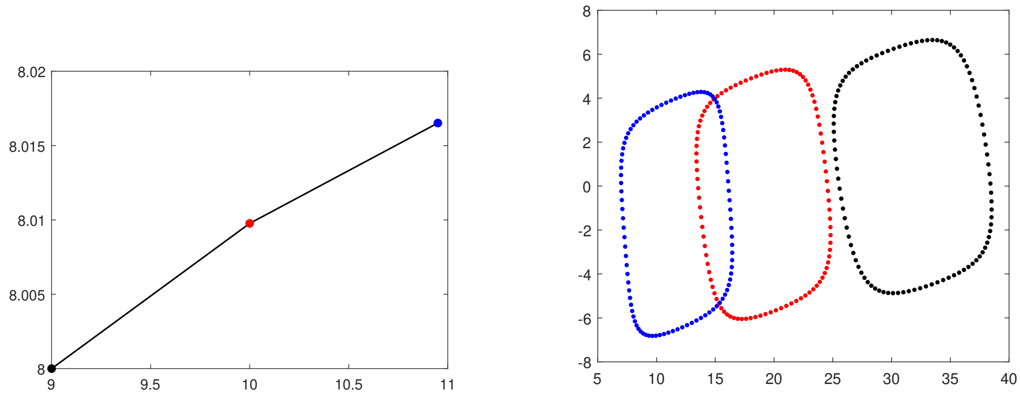
Figure 6.7: The generated iterative points by Algorithm 3 on Example 6.7.3 ((b)) with the initial point as (7, 6)

I_6, I_7, I_8, 5 corresponding F-values at intermediate iterates $\{I_6^i\}_{i\in[5]}, \{I_7^i\}_{i\in[5]}, \{I_8^i\}_{i\in[5]}$ are generated and terminated in T_6, T_7, T_8. For I_1, I_2, I_3, I_4, I_5, F values at initial points are the F-values T_1, T_2, T_3, T_4, T_5 at terminal points. The blue colored sets in Figure 6.9 are F-values $T_i, i \in [8]$. Zoomed-in version of the F-values I_6, I_7 and I_8 at the initial points are shown in Figure 6.10(a), 6.10(b) and 6.10(c).

The iterative points $x^k = (x_1^k, x_2^k)$, executed by Algorithm 3 for the initial point (5, 6), are explicitly provided in Table 6.11. Alongside, the function values of three sample functions f^1, f^{25} and f^{50} across all the iterative points are also reported in Table 6.11. It is noteworthy from the last three columns of Table 6.11 that all the component functions of all functions in F successively degrade across the iterates from the initial to the terminal iterate. It is to observe from Table 6.11 that for the initial point (2, 5), Algorithm 3 takes 6 iterations before it terminates. It is noteworthy from the last three columns of Table 6.11 that all component functions of all functions in F successively degrade across the iterates from the initial to the terminal iterate.

Table 6.10: Performance of Algorithm 3 on Example 6.7.4

Number of initial points	Iterations (Min, Max, Mean, Var, Median, SD)	CPU time (in sec) (Min, Max, Mean, Median, Var, SD)
100	(2, 26, 22.47, 6.29, 23, 2.50)	(21.96, 748.81, 251.58, 237.43, 7242.41, 85.10)



(a) Variable space (b) Image space

Figure 6.8: The generated iterative points by Algorithm 3 on Example 6.7.3 ((b)) with the initial point as $(9, 8)$

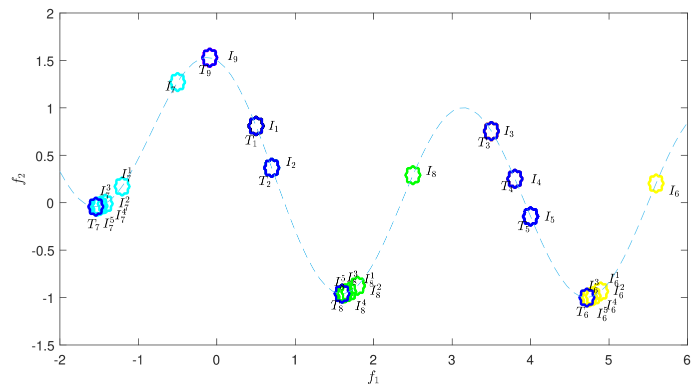
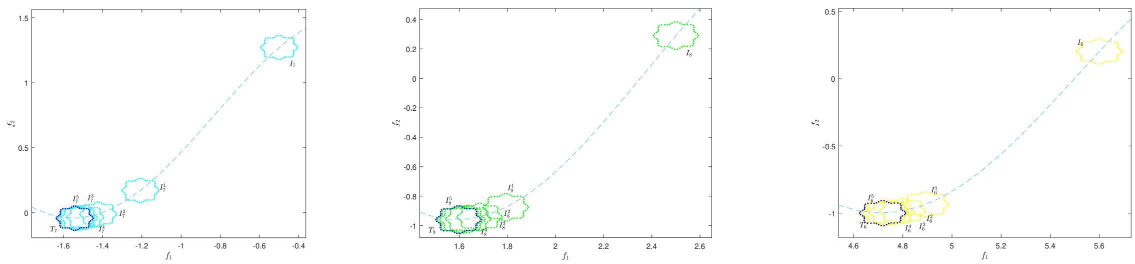


Figure 6.9: Performance of Algorithm 3 for 8 selected initial points in Example 6.7.4

Table 6.11: Iterative points in the argument space for Test instance 6.7.4

Initial point	Iteration number (k)	(x_1^k, x_2^k)	$f^1(x_1^k, x_2^k)$	$f^{25}(x_1^k, x_2^k)$	$f^{50}(x_1^k, x_2^k)$
(5.6)	1	4.893	(4.983, −0.924)	(4.799, −0.935)	(4.986, −0.935)
	2	4.799	(4.888, −0.974)	(4.704, −0.985)	(4.892, −0.985)
	3	4.755	(4.845, −0.985)	(4.661, −0.996)	(4.849, −0.996)
	4	4.734	(4.823, −0.988)	(4.640, −0.999)	(4.827, −0.999)
	5	4.723	(4.813, −0.988)	(4.629, −1.000)	(4.816, −1.000)
	6	4.718	(4.807, −0.989)	(4.624, −1.000)	(4.811, −1.000)

Example 6.7.5 *This example is a two-variables 100-functions with two-components*



(a) For I_1 to T_1 (b) For I_8 to T_8 (c) For I_7 to T_7

Figure 6.10: Intermediate points for the initial points I_6, I_8, I_7 in Figure 6.9

problem. For $i \in [100]$, we consider the function $f^i : \mathbb{R}^2 \to \mathbb{R}^2$ as

$$f^i(x) = \begin{pmatrix} (x_1 + x_2) + (1 + (\cos^{16}(\frac{4\pi i}{100}))) \cos(\frac{2\pi i}{100}) \\ (x_1 + x_2)^2 + (1 + (\cos^{16}(\frac{4\pi i}{100}))) \sin(\frac{2\pi i}{100}) \end{pmatrix}$$

with $\mathcal{S} = [-9, 11] \times [-9, 11]$.

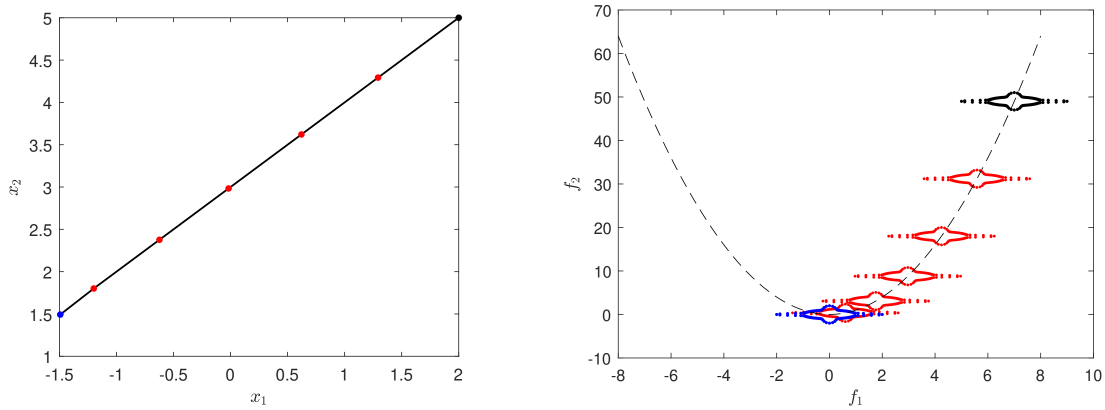
For this problem, we ran Algorithm 3 for randomly chosen 100 initial points taken from $[-9, 11] \times [-9, 11]$. In this test instance, for 85 out of 100 initial points, Algorithm 3 converged successfully in less than 50 iterations. The number of iterations and execution time taken by Algorithm 3 before reaching the stopping condition (see Step 6:) is reported in Table 6.12; SD is the abbreviation of 'standard deviation'.

For the initial point $(2, 5)$, the intermediate iterates of Algorithm 3 in the variable space are shown in Figure 6.11(a) by red bullet points. The corresponding F-values in the image space are shown by red-colored sets in Figure 6.11(b). In Figure 6.11(b), the black bullet point is the initial point, and the blue bullet point is termination point $(-1.493, 1.493)$. The black and blue colored sets in Figure 6.11(b) are F-values at the initial and terminal points, respectively.

The iterative points $x^k = (x_1^k, x_2^k)$, executed by Algorithm 3 for the initial point $(2, 5)$, are explicitly provided in Table 6.13. Alongside, the function values of three sample functions f^5, f^{10} and f^{50} across all the iterative points are also reported in Table 6.13.

Table 6.12: Performance of Algorithm 3 on Example 6.7.5

Number of initial points	Iterations (Min, Max, Mean, Var, Median, SD)	CPU time (in sec) (Min, Max, Mean, Median, Var, SD)
59	$(1, 22, 9.01, 34.91, 8, 5.90)$	$(7.20, 256.01, 109.45, 95.67, 5501.06, 74.16)$



(a) Variable space (b) Image space

Figure 6.11: The generated iterative points by Algorithm 3 on Example 6.7.4 with the initial point as $(2, 5)$

Table 6.13: Solution found in the argument space for Example 6.7.5

Initial point	Iteration number (k)	(x_1^k, x_2^k)	$f^5(x_1^k, x_2^k)$	$f^{10}(x_1^k, x_2^k)$	$f^{50}(x_1^k, x_2^k)$
$(2, 5)$	1	$(1.293, 4.293)$	$(6.569, 31.520)$	$(6.395, 31.789)$	$(3.586, 31.201)$
	2	$(0.621, 3.621)$	$(5.225, 18.316)$	$(5.051, 18.585)$	$(2.242, 17.997)$
	3	$(-0.017, 2.982)$	$(3.949, 9.116)$	$(3.775, 9.385)$	$(0.966, 8.797)$
	4	$(-0.623, 2.376)$	$(2.737, 3.394)$	$(2.562, 3.662)$	$(-0.247, 3.075)$
	5	$(-1.199, 1.801)$	$(1.585, 0.681)$	$(1.411, 0.950)$	$(-1.398, 0.362)$
	6	$(-1.493, 1.493)$	$(0.983, 0.319)$	$(0.809, 0.588)$	$(-2.000, 0.000)$

Example 6.7.6 *In this example, we consider three-variables 100-functions with three-components problem. The objective functions of the SOP is the set-valued map* $F : \mathbb{R}^3 \rightrightarrows \mathbb{R}^3$ *defined by*

$$F(x_1, x_2, x_3) := \{f^1(x_1, x_2, x_3), f^2(x_1, x_2, x_3), \ldots, f^{100}(x_1, x_2, x_3)\},$$

where for each $i \in [100]$, $f^i : \mathbb{R}^2 \to \mathbb{R}^3$ *is given by*

$$f^i(x_1, x_2, x_3) = h(x_1, x_2, x_3) + \tfrac{1}{16} \begin{pmatrix} \cos\phi_i \\ \cos\psi_i \sin\phi_i \\ \sin\psi_i \sin\phi_i \end{pmatrix}$$

with

$$h(x_1, x_2, x_3) = \begin{bmatrix} (1 + g(x_3)) \cos u(x_1) \cos v(x_1, x_2, x_3) \\ (1 + g(x_3)) \cos u(x_1) \sin v(x_1, x_2, x_3) \\ (1 + g(x_3)) \sin u(x_1) \end{bmatrix},$$

$$g(x_3) = (x_3 - \tfrac{1}{2})^2,\ u(x_1) = \frac{\pi x_1}{2},\ v(x_1, x_2, x_3) = \frac{\pi(1 + 2g(x_3)x_2)}{4\left(1 + g\left(\sqrt{x_1^2 + x_2^2 + x_3^2}\right)\right)}$$

and the set $\{(\phi_i, \psi_i) : i \in [100]\}$ *is an enumeration of the set*

$$\{\tfrac{\pi}{10}(j-1) : j \in [10]\} \times \{\tfrac{\pi}{5}(\ell - 1) : \ell \in [10]\},$$

with $\mathcal{S} = [0,1] \times [0,1] \times [0,1]$.

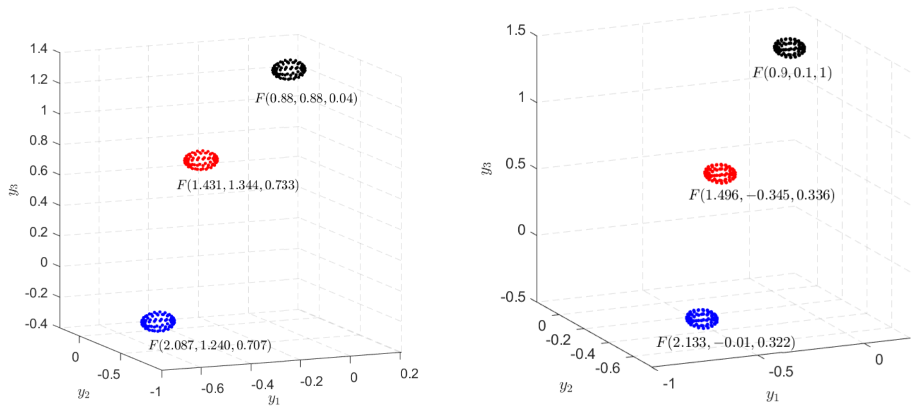
For this problem, we ran Algorithm 3 for randomly chosen 100 *initial points taken from* $[0,1] \times [0,1] \times [0,1]$. *The algorithm successfully converged for* 81 *out of* 100 *initial points in less than* 50 *iterations. The number of iterations and amount of execution time are reported in Table 6.14.*

For the initial point $(0.75, 0.09, 0.02)$, *the intermediate iterates of Algorithm 3 in the variable space are shown by red bullet points in Figure 6.13(a). The corresponding* F*-values in the image space are shown by red spheres in Figure 6.13(b). Similarly,* F*-values at the iterative points* $x^k = (x_1^k, x_2^k)$, *executed by Algorithm 3 for the both initial points* $(0.88, 0.88, 0.04)$ *and* $(0.9, 0.1, 1)$ *in the image space, are shown in Figure 6.12(a) and Figure 6.12(b). In Figure 6.13(a), the black bullet point is the initial point, and the blue bullet point is the termination point* $(0.75, 0.09, 0.02)$. *The black and blue spheres in Figure 6.13(b), 6.12(a) and 6.12(b), are* F*-values at initial and terminal points, respectively.*

The iterative points x^k *for the both initial points* $(0.9, 0.1, 1)$ *and* $(0.88, 0.88, 0.04)$, *are explicitly provided in Table 6.15. Alongside, the function values of three sample values* f^{10}, f^{50}, f^{100} *across all the iterative points are also reported in Table 6.15. It is noteworthy from the last three columns of Table 6.15 that all component functions of all functions in* F *successively degrade across the iterates from the initial to the terminal iterate.*

Table 6.14: Performance of Algorithm 3 on Example 6.7.6

Number of initial points	Iterations (Min, Max, Mean, Var, Median, SD)	CPU time (in sec) (Min, Max, Mean, Median, Var, SD)
81	$(1, 47, 2.01, 26.98, 1, 5.19)$	$(7.39, 573.11, 20.41, 8.23, 4071.34, 63.80)$



(a) Image space (b) Image space

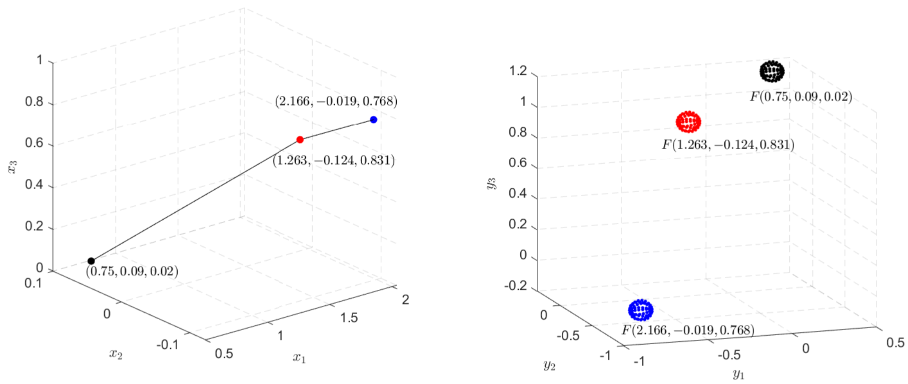
Figure 6.12: The generated iterative points by Algorithm 3 on Example 6.7.6 with the initial points as $(0.88, 0.88, 0.04)$ and $(0.9, 0.1, 1)$, respectively

Table 6.15: Solution found in the argument space for Example 6.7.5

Initial point	Iteration number (k)	(x_1^k, x_2^k)	$f^{10}(x_1^k, x_2^k)$	$f^{50}(x_1^k, x_2^k)$	$f^{100}(x_1^k, x_2^k)$
$(0.9, 0.1, 1)$	1	$(1.496, -0.345, 0.336)$	$(-0.496, -0.492, 0.731)$	$(-0.522, -0.461, 0.731)$	$(-0.558, -0.492, 0.731)$
	2	$(2.133, -0.010, 0.322)$	$(-0.699, -0.696, -0.214)$	$(-0.725, -0.665, -0.214)$	$(-0.762, -0.696, -0.214)$
	3	$(2.133, -0.010, 0.322)$	$(-0.699, -0.696, -0.214)$	$(-0.725, -0.665, -0.214)$	$(-0.762, -0.696, -0.214)$
$(0.88, 0.88, 0.04)$	1	$(1.431, 1.344, 0.733)$	$(-0.403, -0.498, 0.822)$	$(-0.429, -0.467, 0.822)$	$(-0.465, -0.498, 0.822)$
	2	$(2.087, 1.240, 0.707)$	$(-0.664, -0.765, -0.143)$	$(-0.689, -0.734, -0.143)$	$(-0.726, -0.765, -0.143)$
	3	$(2.087, 1.240, 0.707)$	$(-0.664, -0.765, -0.143)$	$(-0.689, -0.734, -0.143)$	$(-0.726, -0.765, -0.143)$
	4	$(2.087, 1.240, 0.707)$	$(-0.664, -0.765, -0.143)$	$(-0.689, -0.734, -0.143)$	$(-0.726, -0.765, -0.143)$
	5	$(2.087, 1.240, 0.707)$	$(-0.664, -0.765, -0.143)$	$(-0.689, -0.734, -0.143)$	$(-0.726, -0.765, -0.143)$
	6	$(2.087, 1.240, 0.707)$	$(-0.664, -0.765, -0.143)$	$(-0.689, -0.734, -0.143)$	$(-0.726, -0.765, -0.143)$
	7	$(2.087, 1.240, 0.707)$	$(-0.664, -0.765, -0.143)$	$(-0.689, -0.734, -0.143)$	$(-0.726, -0.765, -0.143)$

6.8 Conclusion and future directions

In this article, we studied the notion of critical point (Definition 6.1) for set optimization problems ($\mathcal{SOP}_K^l$) with respect to lower set less order relation, where the set-valued objective mapping is given by finitely many twice continuously differentiable vector-valued functions. In order to comprehend the concept of criticality for ($\mathcal{SOP}_K^l$), we have employed the vectorization scheme to examine the criticality concept for the family of vector optimization problems (Lemma 6.1). We have proposed trust-region algorithm (Algorithm 3) to generate a sequence of noncritical iterates that converges to a critical point of the problem ($\mathcal{SOP}_K^l$). To generate the sequence, at each itera-



(a) Variable space (b) Image space

Figure 6.13: The generated iterative points by Algorithm 3 on Example 6.7.6 with the initial points as $(0.75, 0.09, 0.02)$

tion k, at first, a specific element a_k has been chosen from the parition set P_k of the current iterate x_k, which is based on identifying a necessary optimality condition of a weakly-minimal elements of ($\mathcal{SOP}^l_K$). Then, we figured out the trust-region step s_k by solving the subproblem (6.18) whose objective is a model function of the objective function ($\mathcal{VOP}_{a^k}(x_k)$) corresponding to a^k. To decide if or not the trust-region step is descent direction of ($\mathcal{VOP}_{a^k}(x_k)$), the rule of reduction ratio has been studied by examining whether or not it exceeds a signficant threshold positive value (η_1) (Proposition 6.5.2). We have shown that for the chosen a^k, the descent direction of the objective ($\mathcal{VOP}_{a^k}(x_k)$) is the descent of the objective ($\mathcal{SOP}^l_K$) at the iterate x_k. The obtained direction of descent for ($\mathcal{SOP}^l_K$) has been used to find the next iterate x_{k+1}. The iterative process of algorithm 3 kept continuing until the stopping condition (step 6:) was satisfied. The employed stopping condition is a necessary (Proposition 6.5.1) and sufficient (Theorem 6.3) condition of critical points for ($\mathcal{SOP}^l_K$).

The welldefinedness (Subsection 6.5.6) and convergence analysis (Section 6.6) of the proposed Algorithm for the derived Trust-region have been discussed in detail. In case of certifying the well-definedness, we have assured the existence of (a^k, s_k) in Step 3: and existence of a solution of subproblem 6.18 in Step 5: (Theorem 6.3). Also, for any $j \in [w_k]$, reduction ratios $\rho_k^{a_j^k}$ used in Step 7: are well defined (Proposition 6.5.2 and Note 6.2). In the convergence analysis of Algorithm 3, we derived

(i) the predicted value of reduction of the objective ($\mathcal{VOP}_{a^k}(x_k)$) due to movement of the trust-region step is measured by positive value (Theorem 6.4),

(ii) the reduction of model will decrease the objective within sufficiently small trust-region radius (Theorem 6.5),

(iii) at the noncritical iterate, if trust-region radius is small enough, then the trust-region step is accepted and the current iteration is successful iteration (Theorem 6.51). So Algorithm 3 generates sequence of noncritical iterates in finitely many or infinitely many successful iterations,

(iv) global convergence to a critical point (Theorem 6.8) of generated sequence under regularity condition (Definition 1.2.1) and bounded level set of objective ($\mathcal{VOP}_{a^k}(x_k)$) .

Finally, we examined the performance of the proposed trust-region method on some known and newly introduced numerical test problems in Section 6.7.

****