Generalized-Hukuhara-Gradient Efficient-Direction Method to Solve Optimization Problems with Interval-valued Functions and its Application in Least Squares Problems

Debdas Ghosh[a,*], Amit Kumar Debnath[a], Ram Surat Chauhan[a], Oscar Castillo[b]

*[a]Department of Mathematical Sciences, Indian Institute of Technology (BHU) Varanasi
Uttar Pradesh–221005, India*
[b]Tijuana Institute of Technology, Tomas Aquino, Tijuana 22414, Mexico

Abstract

This article proposes a general *gH-gradient efficient-direction method* and a *$\mathcal{W}$-gH-gradient efficient method* for the optimization problems with interval-valued functions. The convergence analysis and the step-wise algorithms of both the methods are presented. It is observed that the $\mathcal{W}$-gH-gradient efficient method converges linearly for a strongly convex interval-valued objective function. To develop the proposed methods and to study their convergence, the idea of strong convexity and sequential criteria for gH-continuity of interval-valued function are illustrated. In the sequel, a new definition of gH-differentiability for interval-valued functions is also proposed. The new definition of gH-differentiability is described with the help of a newly defined concept of linear interval-valued function. It is noticed that the proposed gH-differentiability is superior to the existing ones. For a gH-differentiable interval-valued function, the relation of convexity with the gH-gradient of an interval-valued function and an optimality condition of an interval optimization problem are derived. For the derived optimality condition, a notion of *efficient direction* for interval-valued functions is introduced. The idea of efficient direction is used to develop the proposed gradient methods. As an application of the proposed methods, the least square problem for interval-valued data by $\mathcal{W}$-gH-gradient efficient method is solved. The proposed method for least square problems is illustrated by a polynomial fitting and a logistic curve fitting.

Keywords: Interval-valued functions, Convexity, Strong convexity, gH-continuity, gH-gradient, gH-differentiability, Efficient solution, Efficient direction, Least square problems.

AMS Mathematics Subject Classification (2010): 90C30 · 65K05

---

*Corresponding author
Email addresses:* debdas.mat@iitbhu.ac.in (Debdas Ghosh),
amitkdebnath.rs.mat18@itbhu.ac.in (Amit Kumar Debnath),
rschauhan.rs.mat16@itbhu.ac.in (Ram Surat Chauhan), ocastillo@tectijuana.mx (Oscar Castillo)

1. Introduction

Each area of science, engineering, management, economics, and other practices, uses optimization techniques extensively. Optimization techniques assist us to find the best under specified circumstances. The optimization problems with interval-valued functions (IVFs), known as interval optimization problems (IOPs), has become a significant research topic over the last two decades due to inherent imprecise and uncertain events in different real-world events. In this paper, we attempt to derive a technique for IOPs to capture its solution set. The proposed method reduces to the steepest descent method for the optimization problems with real-valued functions.

1.1. Literature Survey

The ordering and subtraction of intervals has always been a issue in pursuit of an optimal solution for IOPs [25]. In order to deal with interval-valued data, Moore introduced interval arithmetic [45, 46]. However, with the interval arithmetic in [45, 46], one cannot find the additive inverse of a nondegenerate interval (whose lower and upper limits are different), i.e., for a nondegenerate interval $\mathbf{A}$, there does not exist an interval $\mathbf{B}$ such that $\mathbf{A} \oplus \mathbf{B} = \mathbf{0}$. Due to this reason, Wu [58, 59, 60] used a new concept of difference of intervals, known as Hukuhara difference [31] for the difference of two nonempty, closed, bounded and convex subsets of a real linear space. In spite of the fact that the Hukuhara difference for intervals satisfies $\mathbf{A} \ominus_H \mathbf{A} = \mathbf{0}$, $\mathbf{A} \ominus_H \mathbf{B}$ can be calculated only when the width of $\mathbf{A}$ is greater than equal to that of $\mathbf{B}$. In order to overcome this inefficiency of Hukuhara difference of intervals, the 'nonstandard subtraction', introduced by Markov [43], has been used and named as generalized Hukuhara difference (gH-difference) by Stefanini [50, 52]. The generalized Hukuhara difference can be calculated for any pair of intervals and has the property that $\mathbf{A} \ominus_{gH} \mathbf{A} = \mathbf{0}$ [50].

In the ordering of intervals, as intervals are not linearly ordered in contrast to the real numbers, Ishibuchi and Tanaka [32] showed various partial ordering structures and solution concepts for IOPs. They suggested a method to solve a linear IOP by converting it to a bi-objective optimization problem, which is generalized by Chanas and Kuchuta [13]. For nonlinear IOPs, Ghosh studied a Newton method [22] and a quasi-Newton method [23]. Interestingly, many researchers proposed different types of algorithms to solve various types of practical IOPs, for instance, see [14, 15, 17, 20, 38, 56, 57]. Recently, Ghosh et al. [28] introduced variable ordering relations of intervals and proposed an algorithm to obtain the solutions to IOPs. However, research into the applicability of conventional optimization techniques for IOPs is still not concentrated. More surprisingly, although the interplay between geometry and calculus yield optimization techniques, the calculus for IVFs is not rigorously developed until now.

In the year of 2007, with the help of a Hausdorff metric between any two intervals, Wu [58] illustrated the concept of continuity of an IVF. In the same article [58], based on the Hukuhara difference, the concept of Hukuhara-differentiability (H-differentiability) of an IVF has been proposed. Accordingly, the KKT optimality conditions for IOPs have been given in [58]. Further, applying the concept of H-differentiability, Wu [58, 59, 60] studied various duality theories of IOPs. Thereafter, showing the restrictiveness of H-differentiability, Chalco-Cano et al. [11] developed the calculus of IVFs based on the modified concept of the gH-difference,

2

known as generalized-Hukuhara differentiability (gH-differentiability). Chalco-Cano et al. [10] and Ghosh et al. [26] also derived the KKT conditions and duality theories in the view of gH-differentiability.

In the development of interval calculus, calculus for fuzzy-valued functions plays an important role because intervals are particular fuzzy numbers with a special membership function. In connection with fuzzy calculus, Bede and Gal [5] introduced generalized (Hukuhara-based) differentiability; the paper motivated the search for a gH-difference for intervals and fuzzy numbers (see [50, 51, 52]) and applications to fuzzy generalized Hukuhara differentiability (see [6]). A recent contribution in this direction is the article by Stefanini and Arana-Jiménez [53] which contains definitions of total, directional and partial gH-derivatives for multi-variable interval- and fuzzy-valued functions.

In the existing literature on interval calculus, unlike the definition of differentiability of real-valued functions, none of the existing approaches used the concept of a *linear IVF* to define the differentiability of an IVF. Although similar to the definition of differentiability of real-valued functions, the authors of [22] and [53] introduced the new definitions of gH-differentiability for IVFs and studied the properties gH-differentiable IVFs. However, none of them also mentioned about the linear IVF and used the concept of linear IVF to define gH-differentiability for IVFs.

Since the last two decades, with the development of the calculus of IVFs and theories related to IOPs, many techniques, and their algorithmic implementations to obtain the efficient solutions of various types of practical IOPs have been appeared, for instance, see almost all the papers in the references. However, the majority of the methods are provided from the perspective of conventional bi-objective optimization. Thus, to apply those techniques one has to explicitly express an IVF **F** in terms of its real-valued lower $\underline{f}$ and upper $\overline{f}$ boundary functions, which is quite restrictive. For example, in a general least square problem for interval-valued data (see Section 6), one cannot easily express the interval-valued error function in terms of its lower and upper boundary functions. The authors of [22, 25] have studied a parametric form of IOPs and developed the theories and techniques to find efficient solutions to the IOPs with the objective functions that can be parametrically presented. However, for the parametric representation of an IVF one needs its explicit form which is often practically not possible, for instance, consider the function $\mathcal{E}(\beta)$ in (20).

1.2. *Motivation and Contribution of the Paper*

The literature on IOPs shows that there is still no emphasis on the study of conventional optimization strategies for IOPs. Surprisingly, the basic descent method is not yet developed for IOPs. Further, to derive a technique for IOPs which is similar to the standard descent method, we need to rigorously establish the notion of gH-differentiability concept for IVFs. More importantly, it must be kept in mind that the derived technique must be applicable to general IVFs regardless of whether or not

(i) the objective function can be expressed parametrically, or

(ii) the explicit form of the lower and upper function of the objective function can be found.

After illustrating the concept of a linear IVF, this paper proposes a new definition of gH-differentiability. It is shown that if an IVF is gH-differentiable at a point, its gH-gradient exists at that point. It is shown that the proposed definition of gH-differentiability is superior to the existing ones (see Remark 8 for details). With the help of gH-gradient, a few characterization results for a gH-differentiable convex IVF are derived. Also, several results related to the gH-gradient of a strong convex gH-differentiable IVF are studied.

Further, with the help of the proposed gH-differentiability for IVFs, this article develops a gradient descent method for interval optimization, namely a general gH-gradient efficient-direction method for IOPs. Similar to the steepest descent method, a method is also proposed, named $\mathcal{W}$-gradient efficient method, to obtain efficient solutions of IOPs. The main advantages of the proposed methods are that one needs neither the explicit forms of upper and lower functions of the objective function nor parametric forms of the corresponding IVFs of an IOP. It is shown that the $\mathcal{W}$-gradient efficient method for IOP converges linearly in the case of strong convexity of the interval-valued objective function. In order to develop these methods, the notion of efficient-direction for an IVF and its several characteristics are studied.

1.3. Delineation

The presentation sequence of the proposed work is the following. The next section covers some basic terminologies and notions of intervals analysis followed by the convexity and a few topics of differential calculus of IVFs. Also, the sequential criteria of gH-continuity of an IVF is discussed in Section 2. The concept of a linear IVF, a new concept of gH-differentiability of an IVF, and a few characterizations of a gH-differentiable convex IVF are given in Section 3. The concept of efficient solutions and an optimality condition of an IOP are discussed in Section 4. In Section 5, a general gH-gradient efficient-direction method for IOPs and a $\mathcal{W}$-gradient efficient method for IOP are proposed. Their algorithmic implementations and the convergence analysis are also studied in Section 5. The section 6 deals with the application of $\mathcal{W}$-gradient efficient method for IOPs in least square problems with interval data. Finally, in Section 7, a few future directions of this study are given.

2. Preliminaries and Terminologies

This section provides some basic terminologies and notions on intervals followed by the convexity and a few topics of differential calculus of IVFs.

2.1. Arithmetic of Intervals and their Dominance Relation

At first, this section describes the generalized concept of the difference of two intervals and the ordering concepts of intervals. Along with these definitions, we use Moore's interval addition ($\oplus$) multiplication ($\odot$) and division ($\oslash$) [45, 46] throughout the paper.

Let the set of real numbers be denoted by $\mathbb{R}$ and the set of all closed and bounded intervals be denoted by $I(\mathbb{R})$. Throughout the article, the elements of $I(\mathbb{R})$ are represented by bold

capital letters $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$.. To represent an element $\mathbf{A} \in I(\mathbb{R})$ in the interval form, the corresponding small letter is used in the following way: $\mathbf{A} = [\underline{a}, \overline{a}]$. If $\underline{a} = \overline{a}]$, then $\mathbf{A}$ is called a *degenerate interval*.

It is to be mentioned that any singleton $\{p\}$ of $\mathbb{R}$ can be represented by an interval $\mathbf{P} = [\underline{p}, \overline{p}]$, where $\underline{p} = p = \overline{p}$. In particular,

$$\mathbf{0} = \{0\} = [0, 0] \text{ and } \mathbf{1} = \{1\} = [1, 1].$$

Remark 1. It is easy to check that the addition and multiplication of intervals are commutative, the addition of intervals is associative, and

$$\mathbf{A} \ominus \mathbf{B} = \mathbf{A} \oplus (-1) \odot \mathbf{B}.$$

Since the property of subtraction of intervals cannot provide an additive inverse of a non-degenerate interval, in this article, we use the gH-difference of intervals, which is defined as follows.

Definition 2.1. (gH-*difference of intervals* [52]). Let $\mathbf{A}$ and $\mathbf{B}$ be two elements of $I(\mathbb{R})$. The gH-difference between $\mathbf{A}$ and $\mathbf{B}$, denoted $\mathbf{A} \ominus_{gH} \mathbf{B}$, is defined by an interval $\mathbf{C}$ such that

$$\mathbf{A} = \mathbf{B} \oplus \mathbf{C} \text{ or } \mathbf{B} = \mathbf{A} \ominus \mathbf{C}.$$

It is to be noted that for $\mathbf{A} = [\underline{a}, \overline{a}]$ and $\mathbf{B} = [\underline{b}, \overline{b}]$,

$$\mathbf{A} \ominus_{gH} \mathbf{B} = \left[\min\{\underline{a} - \underline{b}, \overline{a} - \overline{b}\}, \max\{\underline{a} - \underline{b}, \overline{a} - \overline{b}\} \right].$$

Thus,

$$\mathbf{A} \ominus_{gH} \mathbf{A} = \mathbf{0} \text{ and } \mathbf{0} \ominus_{gH} \mathbf{A} = (-1) \odot \mathbf{A}.$$

Definition 2.2. (*Algebraic operations on* $I(\mathbb{R})^n$). Let $\bar{\mathbf{A}} = (\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)^T$ and $\bar{\mathbf{B}} = (\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n)^T$ be two elements of $I(\mathbb{R})^n$. An algebraic operation '$\star$' between $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$, denoted $\bar{\mathbf{A}} \star \bar{\mathbf{B}}$, is defined by

$$\bar{\mathbf{A}} \star \bar{\mathbf{B}} = (\mathbf{A}_1 \star \mathbf{B}_1, \mathbf{A}_2 \star \mathbf{B}_2, \ldots, \mathbf{A}_n \star \mathbf{B}_n)^T,$$

where $\star \in \{\oplus, \ominus, \ominus_{gH}\}$.

Definition 2.3. (*Dominance relation of interval* [58]). For any two intervals $\mathbf{A}$ and $\mathbf{B}$ in $I(\mathbb{R})$,

(i) if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$, then $\mathbf{B}$ is said to be dominated by $\mathbf{A}$ and denoted by $\mathbf{A} \preceq \mathbf{B}$;

(ii) if either $\underline{a} \leq \underline{b}$ and $\overline{a} < \overline{b}$ or $\underline{a} < \underline{b}$ and $\overline{a} \leq \overline{b}$ hold, then $\mathbf{B}$ is said to be strictly dominated by $\mathbf{A}$ and denoted by $\mathbf{A} \prec \mathbf{B}$;

(iii) if $\mathbf{B}$ is not dominated by $\mathbf{A}$, then $\mathbf{A} \npreceq \mathbf{B}$ and if $\mathbf{B}$ is not strictly dominated by $\mathbf{A}$, then $\mathbf{A} \nprec \mathbf{B}$;

(iv) if $\mathbf{A} \npreceq \mathbf{B}$ and $\mathbf{B} \npreceq \mathbf{A}$, then it will be said that none of $\mathbf{A}$ and $\mathbf{B}$ dominates the other, or $\mathbf{A}$ and $\mathbf{B}$ are not comparable.

One can note that Wu [58, 60] used the term 'superior than' to describe the dominance relation between two intervals. However, in this article, we use the term 'dominated by' instead of 'superior than'.

Lemma 2.1. For two elements $\mathbf{A}$ and $\mathbf{B}$ of $I(\mathbb{R})$,

(i) $\mathbf{A} \preceq \mathbf{B} \Longleftrightarrow \mathbf{A} \ominus_{gH} \mathbf{B} \preceq \mathbf{0}$ and

(ii) $\mathbf{A} \not\prec \mathbf{B} \Longleftrightarrow \mathbf{A} \ominus_{gH} \mathbf{B} \not\prec \mathbf{0}$.

Proof. See Appendix A. □

Lemma 2.2. For an $\mathbf{A} \in I(\mathbb{R})$,

(i) $\mathbf{0} \preceq \mathbf{A} \Longleftrightarrow (-1) \odot \mathbf{A} \preceq \mathbf{0}$ and

(ii) $\mathbf{0} \not\prec \mathbf{A} \Longleftrightarrow (-1) \odot \mathbf{A} \not\prec \mathbf{0}$.

Proof. As $\mathbf{0} \ominus_{gH} \mathbf{A} = (-1) \odot \mathbf{A}$, replacing $\mathbf{A}$ by $\mathbf{0}$ and $\mathbf{B}$ by $\mathbf{A}$ in Lemma 2.1, we get the required results. □

Definition 2.4. (*Norm on $I(\mathbb{R})$* [45]). For an $\mathbf{A} = [\underline{a}, \bar{a}]$ in $I(\mathbb{R})$, the function $\|.\|_{I(\mathbb{R})} : I(\mathbb{R}) \to \mathbb{R}^+$, defined by

$$\|\mathbf{A}\|_{I(\mathbb{R})} = \max\{|\underline{a}|, |\bar{a}|\},$$

is a norm on $I(\mathbb{R})$.

Definition 2.5. (*Norm on $I(\mathbb{R})^n$* [46]). For an $\bar{\mathbf{A}} = (\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n)^T$ in $I(\mathbb{R})^n$, the function $\|\cdot\|_{I(\mathbb{R})^n} : I(\mathbb{R}) \to \mathbb{R}^+$, defined by

$$\|\bar{\mathbf{A}}\|_{I(\mathbb{R})^n} = \sum_{i=1}^n \|\mathbf{A}_i\|_{I(\mathbb{R})}$$

is a norm on $I(\mathbb{R})^n$.

In this article, although we use the notions '$\|\cdot\|_{I(\mathbb{R})}$' and '$\|\cdot\|_{I(\mathbb{R})^n}$' to denote the norms on $I(\mathbb{R})$ and $I(\mathbb{R})^n$, respectively, we simply use the notion '$\|\cdot\|$' to denote the usual Euclidean norm on $\mathbb{R}^n$.

2.2. Convexity and Basic Differential Calculus of Interval-valued Functions

Let $\mathcal{X}$ be a nonempty subset of $\mathbb{R}^n$. An IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$, for each argument point $x \in \mathcal{X}$, is presented by the ontic (see [19]) way:

$$\mathbf{F}(x) = \left[\underline{f}(x), \overline{f}(x)\right],$$

where $\underline{f}$ and $\overline{f}$ are real-valued functions on $\mathcal{X}$. The functions $\underline{f}$ and $\overline{f}$ are called the lower and the upper functions of $\mathbf{F}$, respectively.

Definition 2.6. (*Convex IVF* [58]). Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a convex set. An IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is said to be a convex function if for any two vectors x_1 and x_2 in $\mathcal{X}$,

$$\mathbf{F}(\lambda_1 x_1 + \lambda_2 x_2) \preceq \lambda_1 \odot \mathbf{F}(x_1) \oplus \lambda_2 \odot \mathbf{F}(x_2)$$

for all $\lambda_1, \lambda_2 \in [0, 1]$ with $\lambda_1 + \lambda_2 = 1$.

Remark 2. (See [58]). $\mathbf{F}$ is convex if and only if $\underline{f}$ and $\overline{f}$ are convex.

Definition 2.7. (*Strongly convex IVF*). Let $\mathcal{X}$ be a nonempty convex subset of $\mathbb{R}^n$. An IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is said to be strongly convex on $\mathcal{X}$ if there exists a convex IVF $\mathbf{G} : \mathcal{X} \to I(\mathbb{R})$ and a $\sigma > 0$ such that

$$\mathbf{F}(x) = \mathbf{G}(x) \oplus \frac{1}{2}\|x\|^2 \odot [\sigma, \sigma] \text{ for all } x \in \mathcal{X}.$$

Remark 3. It is to be observed that

$$\mathbf{F}(x) = \mathbf{G}(x) \oplus \frac{1}{2}\|x\|^2 \odot [\sigma, \sigma]$$

implies

$$\left[\underline{g}(x), \overline{g}(x)\right] \oplus \frac{1}{2}\|x\|^2 \odot [\sigma, \sigma] = \left[\underline{f}(x), \overline{f}(x)\right]$$

$$\text{or, } \left[\underline{g}(x) + \frac{\sigma}{2}\|x\|^2, \overline{g}(x) + \frac{\sigma}{2}\|x\|^2\right] = \left[\underline{f}(x), \overline{f}(x)\right]$$

$$\text{or, } \underline{g}(x) = \underline{f}(x) - \frac{\sigma}{2}\|x\|^2 \text{ and } \overline{g}(x) = \overline{f}(x) - \frac{\sigma}{2}\|x\|^2.$$

Therefore,

$$\mathbf{F} \text{ is strongly convex} \Longleftrightarrow \mathbf{G} \text{ is convex}$$
$$\Longleftrightarrow \underline{g} \text{ and } \overline{g} \text{ are convex, by Remark 2}$$
$$\Longleftrightarrow \underline{f} \text{ and } \overline{f} \text{ are strongly convex.}$$

Definition 2.8. (*gH-continuity* [22]). Let $\mathbf{F}$ be an IVF on a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$. Let $\bar{x}$ be an interior point of $\mathcal{X}$ and $d \in \mathbb{R}^n$ be such that $\bar{x} + d \in \mathcal{X}$. The function $\mathbf{F}$ is said to be continuous at $\bar{x}$ if

$$\lim_{\|d\| \to 0} \left(\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})\right) = \mathbf{0}.$$

Lemma 2.3. An IVF $\mathbf{F}$ on a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$ is gH-continuous if and only if $\underline{f}$ and $\overline{f}$ are continuous.

Proof. See Appendix B. □

Lemma 2.4. (*Sequential criteria of gH-continuity*). An IVF $\mathbf{F}$ on a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$ is gH-continuous at a point $\bar{x} \in \mathcal{X}$ if and only if for every sequence $\{x_n\}$ in $\mathcal{X}$ converging to $\bar{x}$, the sequence $\{\mathbf{F}(x_n)\}$ converges to $\mathbf{F}(\bar{x})$.

Proof. See Appendix C. □

Definition 2.9. (*gH-Lipschitz continuous IVF* [27]). Let $\mathcal{X} \subseteq \mathbb{R}^n$. An IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is said to be gH-Lipschitz continuous on $\mathcal{X}$ if there exists $L > 0$ such that

$$\|\mathbf{F}(x) \ominus_{gH} \mathbf{F}(y)\|_{I(\mathbb{R})} \leq L\|x - y\| \text{ for all } x, y \in \mathcal{X}.$$

The constant L is called a Lipschitz constant.

Definition 2.10. (*gH-derivative* [51]). Let $\mathcal{X} \subseteq \mathbb{R}$. The gH-derivative of an IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ at $\bar{x} \in \mathcal{X}$ is defined by

$$\mathbf{F}'(\bar{x}) = \lim_{d \to 0} \frac{\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})}{d}, \text{ provided the limit exists.}$$

Remark 4. (See [9]). Let $\mathcal{X}$ be a nonempty subset of $\mathbb{R}$. The gH-derivative of an IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ at $\bar{x} \in \mathcal{X}$ exists if the derivatives of $\underline{f}$ and $\overline{f}$ at $\bar{x}$ exist and

$$\mathbf{F}'(\bar{x}) = \left[\min\left\{\underline{f}'(\bar{x}), \overline{f}'(\bar{x})\right\}, \max\left\{\underline{f}'(\bar{x}), \overline{f}'(\bar{x})\right\}\right].$$

However, the converse is not true.

Definition 2.11. (*Partial gH-derivative* [10]). Let $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ be an IVF, where $\mathcal{X}$ is a nonempty subset of $\mathbb{R}^n$. Let $\mathbf{G}_i$ be defined by

$$\mathbf{G}_i(x_i) = \mathbf{F}(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \ldots, \bar{x}_n),$$

where $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)^T \in \mathcal{X}$. If the gH-derivative of $\mathbf{G}_i$ exists at $\bar{x}_i$, then the i-th partial gH-derivative of $\mathbf{F}$ at $\bar{x}$, denoted $D_i\mathbf{F}(\bar{x})$, is defined by

$$D_i\mathbf{F}(\bar{x}) = \mathbf{G}_i'(\bar{x}_i) \text{ for all } i = 1, 2, \ldots, n.$$

Definition 2.12. (*gH-gradient* [10]). Let $\mathcal{X}$ be a nonempty subset of $\mathbb{R}^n$. The gH-gradient of an IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ at a point $\bar{x} \in \mathcal{X}$, denoted $\nabla\mathbf{F}(\bar{x})$, is defined by

$$\nabla\mathbf{F}(\bar{x}) = (D_1\mathbf{F}(\bar{x}), D_2\mathbf{F}(\bar{x}), \ldots, D_n\mathbf{F}(\bar{x}))^T.$$

It is to be mentioned that the authors of [10] used the notations '$\left(\frac{\partial \mathbf{F}}{\partial x_i}\right)_g (\bar{x})$' and '$\nabla_g\mathbf{F}(\bar{x})$' for i-th partial gH-derivative and gH-gradient of $\mathbf{F}$ at $\bar{x}$, respectively. However, throughout the article we simply use the notations '$D_i\mathbf{F}(\bar{x})$' and '$\nabla\mathbf{F}(\bar{x})$' for i-th partial gH-derivative and gH-gradient of $\mathbf{F}$ at $\bar{x}$, respectively.

Definition 2.13. (*gH-Lipschitz gradient* [27]). An IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is said to have gH-Lipschitz gradient on $\mathcal{X} \subseteq \mathbb{R}^n$ if there exists $M > 0$ such that

$$\|\nabla\mathbf{F}(x) \ominus_{gH} \nabla\mathbf{F}(y)\|_{I(\mathbb{R})^n} \leq M\|x - y\| \text{ for all } x, y \in \mathcal{X}.$$

Until now, the concepts of gH-continuity, gH-derivative, partial gH-derivative, gH-gradient for IVF have been discussed. In the next section, we illustrate the idea of differentiability for

8

IVFs. This idea differentiability for IVFs is used in the rest of the paper to develop the gradient descent method of IOP.

3. Differentiability of Interval-valued Functions

==Behind the concept of differentiability of a function the concept of linearity plays an important role==. Thus, before exploring the concept of differentiability of an IVF, we discuss the concept of a linear IVF.

Definition 3.1. (*Linear IVF*). Let $\mathcal{X}$ be a linear subspace of $\mathbb{R}^n$. A function $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is said to be linear if

$$\mathbf{F}(x) = \bigoplus_{i=1}^{n} x_i \odot \mathbf{F}(e_i) \text{ for all } x = (x_1, x_2, \ldots, x_n)^T \in \mathcal{X},$$

where e_i is the i-th standard basis vector of $\mathbb{R}^n$, $i = 1, 2, \ldots, n$ and '$\bigoplus_{i=1}^{n}$' denotes successive addition of n number of intervals.

Remark 5. It is noteworthy that any IVF $\mathbf{F} : \mathbb{R}^n \to I(\mathbb{R})$ of the following form

$$\mathbf{F}(x) = \bigoplus_{i=1}^{n} x_i \odot \mathbf{A}_i = \bigoplus_{i=1}^{n} x_i \odot [\underline{a}_i, \overline{a}_i],$$

is a linear IVF.

Example 3.1. The IVF $\mathbf{F}(x) : \mathbb{R} \to I(\mathbb{R})$, which is defined by

$$\mathbf{F}(x) = [-3, 7] \odot x$$

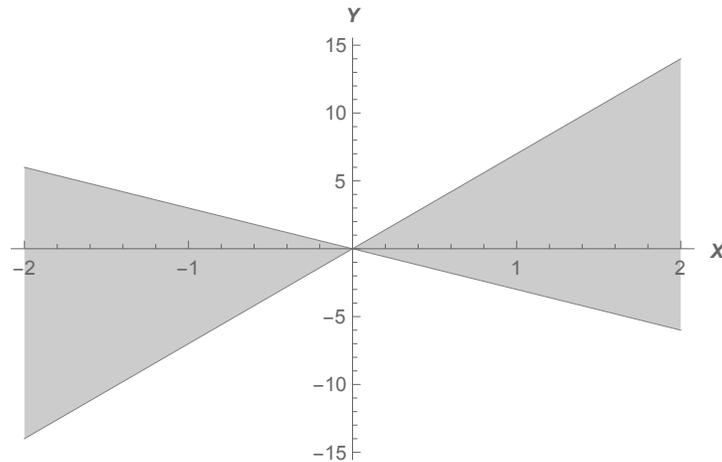is a linear IVF, which is depicted in Figure 1 by gray shaded region.

Figure 1: Interval-valued function of Example 3.1

9

Remark 6. A linear IVF $\mathbf{F}$ on a linear subspace $\mathcal{X}$ of $\mathbb{R}^n$ satisfies the following conditions:

(i) $\mathbf{F}(\lambda x) = \lambda \odot \mathbf{F}(x)$ for all $x \in \mathcal{X}$ and for all $\lambda \in \mathbb{R}$, and

(ii) for all $x,\ y \in \mathcal{X}$, either
$$\mathbf{F}(x + y) = \mathbf{F}(x) \oplus \mathbf{F}(y)$$
or none of $\mathbf{F}(x) \oplus \mathbf{F}(y)$ and $\mathbf{F}(x + y)$ dominates the other.

For the proof, see Appendix D.

Definition 3.2. (*gH-differentiability*). Let $\mathcal{X}$ be a nonempty subset of $\mathbb{R}^n$. An IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is said to be gH-differentiable at a point $\bar{x} \in \mathcal{X}$ if there exists a linear IVF $\mathbf{L}_{\bar{x}} : \mathbb{R}^n \to I(\mathbb{R})$, an IVF $\mathbf{E}(\mathbf{F}(\bar{x}); d)$ and a $\delta > 0$ such that

$$(\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \mathbf{L}_{\bar{x}}(d) = \|d\| \odot \mathbf{E}(\mathbf{F}(\bar{x}); d) \ \text{ for all } \ d \ \text{ such that } \ \|d\| < \delta,$$

where $\mathbf{E}(\mathbf{F}(\bar{x}); d) \to \mathbf{0}$ as $\|d\| \to 0$.

If $\mathbf{F}$ is gH-differentiable at each point $\bar{x} \in \mathcal{X}$, then $\mathbf{F}$ is said to be gH-differentiable on $\mathcal{X}$.

Remark 7. It is to note from Definition 3.2 that

$$\lim_{\|d\| \to 0} [(\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \mathbf{L}_{\bar{x}}(d)] = \lim_{\|d\| \to 0} \mathbf{E}(\mathbf{F}(\bar{x}); d)$$

$$\text{or, } \lim_{\|d\| \to 0} (\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \lim_{\|d\| \to 0} \mathbf{L}_{\bar{x}}(d) = \mathbf{0}$$

$$\text{or, } \lim_{\|d\| \to 0} (\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})) = \mathbf{0}.$$

Thus, every gH-differentiable IVF $\mathbf{F}$ is gH-continuous.

The following lemma is same as Proposition 7 in [53]. However, in [53], Proposition 7 is proved by expressing an IVF $\mathbf{F}$ in terms of its midpoint-radius representation, i.e., $\mathbf{F} = \left[\frac{\underline{f} + \overline{f}}{2}, \frac{\overline{f} - \underline{f}}{2}\right]$, but in this article, to prove the following lemma we do not use the midpoint-radius representation of an IVF.

Lemma 3.1. Let $\mathcal{X}$ be a nonempty subset of $\mathbb{R}^n$. If an IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is gH-differentiable at $\bar{x} \in \mathcal{X}$, then there exists a nonzero λ and $\delta > 0$ such that

$$\lim_{\lambda \to 0} \frac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda h) \ominus_{gH} \mathbf{F}(\bar{x})) = \mathbf{L}_{\bar{x}}(h) \ \text{ for all } \ h \in \mathbb{R}^n \ \text{ with } \ |\lambda| \|h\| < \delta,$$

where $\mathbf{L}_{\bar{x}}$ is the linear IVF in Definition 3.2.

Proof. See Appendix E. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Theorem 3.1. *Let an IVF $\mathbf{F}$ on a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$ be gH-differentiable at $\bar{x} \in \mathcal{X}$. Then, for each $d = (d_1, d_2, \ldots, d_n)^T \in \mathbb{R}^n$, the gH-gradient of $\mathbf{F}$ at $\bar{x}$ exists and the linear IVF $\mathbf{L}_{\bar{x}}$ in Definition 3.2 can be expressed by*

$$\boldsymbol{L}_{\bar{x}}(d) = d^T \odot \nabla \boldsymbol{F}(\bar{x}), \tag{1}$$

where $d^T \odot \nabla \boldsymbol{F}(\bar{x}) = \bigoplus_{i=1}^{n} d_i \odot D_i \boldsymbol{F}(\bar{x})$.

Proof. See Appendix F. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

The following example shows that the converse of Theorem 3.1 is not true.

Example 3.2. On $\mathbb{R}^2$, consider the following IVF:

$$\mathbf{F}(x) = x_1 \odot \mathbf{A}_1 \oplus x_2 \odot \mathbf{A}_2 = x_1 \odot [-1, 1] \oplus x_2 \odot [0, 2].$$

At $\bar{x} = (0, 1)$,

$$\mathbf{F}(\bar{x}) = [0, 2], \ D_1\mathbf{F}(\bar{x}) = \mathbf{A}_1 = [-1, 1], \ D_2\mathbf{F}(\bar{x}) = \mathbf{A}_2 = [0, 2].$$

Therefore, the gH-gradient of $\mathbf{F}$ at $\bar{x} = (0, 1)$ exists, $\nabla \mathbf{F}(\bar{x}) = (\mathbf{A}_1, \mathbf{A}_2)^T$ and

$$d^T \odot \nabla \mathbf{F}(\bar{x}) = d_1 \odot \mathbf{A}_1 \oplus d_2 \odot \mathbf{A}_2 = \mathbf{F}(d) \text{ for any direction } d \in \mathbb{R}^2.$$

However, $\mathbf{F}$ is not gH-differentiable at $\bar{x}$ because at a direction $d = (t, -t)$ with $t > 0$ and $\|d\| < \frac{1}{2}$, we obtain

$$\mathbf{F}(\bar{x} + d) = \mathbf{F}(t, 1 - t) = [-t, 2 - t], \ \mathbf{F}(d) = \mathbf{F}(t, -t) = [-3t, t],$$

and

$$\begin{aligned}
&\lim_{t \to 0+} \frac{1}{\|d\|} \odot \left((\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} d^T \odot \nabla \mathbf{F}(\bar{x}) \right) \\
&= \lim_{t \to 0+} \frac{1}{\sqrt{2}t} \odot \left(([-t, 2 - t] \ominus_{gH} [0, 2]) \ominus_{gH} [-3t, t] \right) \\
&= \lim_{t \to 0+} \frac{1}{\sqrt{2}t} \odot \left([-t, -t] \ominus_{gH} [-3t, t] \right) \\
&= \lim_{t \to 0+} \frac{1}{\sqrt{2}t} \odot [-2t, 2t] \\
&= \sqrt{2} \odot [-1, 1] \\
&\neq \mathbf{0}.
\end{aligned}$$

Remark 8. By Theorem 3.1, one can notice that the proposed Definition 3.2 of gH-differentiability of this article implies the definition of gH-differentiability proposed in [50]. One may think that the definition of gH-differentiability of this article is same as that in [22]. However, it can be noted that the IVF $\mathbf{L}_{\bar{x}}$ in [22] has been considered with the following two properties:

(a) $\mathbf{L}_{\bar{x}}(\lambda x) = \lambda \odot \mathbf{L}_{\bar{x}}(x)$ for all $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^n$ and

(b) $\mathbf{L}_{\bar{x}}(x + y) = \mathbf{L}_{\bar{x}}(x) \oplus \mathbf{L}_{\bar{x}}(y)$ for all $x, \ y \in \mathbb{R}^n$.

Thus, the IVF $\mathbf{L}_{\bar{x}}$ in [22] is a particular case of the proposed $\mathbf{L}_{\bar{x}}$ (see Definition 3.1). Hence, the definition of gH-differentiability of this article is more general than the definition of [22].

In the following example, we provide an IVF, which is gH-differentiable in the sense of this article but not in the sense of [22].

Example 3.3. Consider the IVF $\mathbf{F} : \mathbb{R} \to I(\mathbb{R})$ which is defined by

$$\mathbf{F}(x) = [-1, 1] \odot x^2, \ x \in \mathbb{R}.$$

Thus,

$$\underline{f}(x) = -x^2 \ \text{and} \ \overline{f}(x) = x^2.$$

The gH-gradient of $\mathbf{F}$ is

$$\nabla \mathbf{F}(x) = [-2, 2] \odot x.$$

Since both the real-valued functions $\underline{f}$ and $\overline{f}$ are differentiable at $\bar{x} = 1$, according to Remark 4 the IVF $\mathbf{F}$ is gH-differentiable at $\bar{x} = 1$. Hence, due to Theorem 3.1 of this article and Theorem 1 of [22] there exists an IVF $\mathbf{L}_1$ such that

$$\begin{aligned} \mathbf{L}_1(h) &= h \odot \nabla \mathbf{F}(1) \\ &= h \odot [-2, 2] \odot (1) \\ &= [-2, 2] \odot h, \ \text{where} \ h \in \mathbb{R}. \end{aligned}$$

By Remark 5, $\mathbf{L}_1$ is a linear IVF. Hence, in the sense of the definition of gH-differentiability of this article, $\mathbf{F}$ is gH-differentiable at $\bar{x} = 1$.

However, $\mathbf{F}$ is not gH-differentiable at $\bar{x} = 1$ in the sense of [22] because there exist some p, $q \in \mathbb{R}$ such that

$$\mathbf{L}_1(p + q) \neq \mathbf{L}_1(p) \oplus \mathbf{L}_1(q).$$

For instance, consider $p = 3$ and $q = -2$. Then,

$$\mathbf{L}_1(p + q) = \mathbf{L}_1(1) = [-2, 2]$$

and

$$\mathbf{L}_1(p) \oplus \mathbf{L}_1(q) = [-9, 9] \oplus [-6, 6] = [-15, 15] \neq [-2, 2].$$

Remark 9. It may appear as if Theorem 3.1 of this article is same as Theorem 1 in [22] but we have seen that $\mathbf{L}_{\bar{x}}$ in [22] is a particular case of $\mathbf{L}_{\bar{x}}$ of this article. Thus, it is clear that Theorem 3.1 of this article is the generalized version of Theorem 1 in [22].

Remark 10. It is noteworthy that although each linear real-valued function is differentiable in its domain, Example 3.2 shows that there exists a few linear IVFs that are not gH-differentiable.

The following theorem provides a condition for a linear IVF to be gH-differentiable.

Theorem 3.2. *Let $\mathcal{X}$ be a linear subspace of $\mathbb{R}^n$ and $\mathbf{F}$ be an IVF on $\mathcal{X}$. For a given $\bar{x} \in \mathcal{X}$, if for any $d \in \mathcal{N}_\delta(\bar{x}) \cap \mathcal{X}$,*

$$\mathbf{F}(\bar{x} + d) = \mathbf{F}(\bar{x}) \oplus \mathbf{F}(d),$$

where $\mathcal{N}_\delta(\bar{x})$ is a δ-neighborhood of $\bar{x}$, then $\mathbf{F}$ is gH-differentiable at $\bar{x} \in \mathcal{X}$.

□

Theorem 3.3. *Let an IVF $\mathbf{F}$ on a nonempty open convex subset $\mathcal{X}$ of $\mathbb{R}^n$ be gH-differentiable. If the function $\mathbf{F}$ is convex on $\mathcal{X}$, then*

$$(y - x)^T \odot \nabla \mathbf{F}(x) \preceq \mathbf{F}(y) \ominus_{gH} \mathbf{F}(x) \text{ for all } x, y \in \mathcal{X}.$$

□

Theorem 3.4. *Let an IVF $\mathbf{F}$ on a nonempty open convex subset $\mathcal{X}$ of $\mathbb{R}^n$ be gH-differentiable on $\mathcal{X}$. If the function $\mathbf{F}$ is convex on $\mathcal{X}$, then*

$$\mathbf{0} \preceq (x - y)^T \odot \nabla \mathbf{F}(x) \ominus_{gH} (x - y)^T \odot \nabla \mathbf{F}(y) \text{ for all } x, y \in \mathcal{X}.$$

□

Remark 11. One may think that for a gH-differentiable IVF $\mathbf{F}$ on $\mathcal{X} \subseteq \mathbb{R}^n$,

$$(x-y)^T \odot \nabla \mathbf{F}(x) \ominus_{gH} (x-y)^T \odot \nabla \mathbf{F}(y) = (x-y)^T \odot (\nabla \mathbf{F}(x) \ominus_{gH} \nabla \mathbf{F}(y)) \text{ for all } x, y \in \mathcal{X}.$$

Unfortunately, it is not true in general even if $\mathbf{F}$ is convex on $\mathcal{X}$. For instance, consider the following IVF on $\mathbb{R}^2$:

$$\mathbf{F}(x_1, x_2) = [1, 3] \odot x_1^2 \oplus [1, 4] \odot x_2^2 = [x_1^2 + x_1^2, \, 3x_1^2 + 4x_2^2].$$

Since $\underline{f}(x_1, x_2) = x_1^2 + x_1^2$ and $\overline{f}(x_1, x_2) = 3x_1^2 + 4x_2^2$ are convex on $\mathbb{R}^2$, by Remark 2, $\mathbf{F}$ is convex on $\mathbb{R}^2$.

The gH-gradient of $\mathbf{F}$ is

$$\nabla \mathbf{F}(x) = ([2, 6] \odot x_1, [2, 8] \odot x_2)^T.$$

Considering $x = (2, 0)$ and $y = (1, 1)$ we have

$$\begin{aligned}
&(x - y)^T \odot \nabla \mathbf{F}(x) \ominus_{gH} (x - y)^T \odot \nabla \mathbf{F}(y) \\
&= (1, -1) \odot ([4, 12], [0, 0])^T \ominus_{gH} (1, -1) \odot ([2, 6], [2, 8])^T \\
&= [4, 12] \oplus [-6, 4] \\
&= [-2, 16]
\end{aligned}$$

and

$$\begin{aligned}
(x - y)^T (\nabla \mathbf{F}(x) \ominus_{gH} \nabla \mathbf{F}(y)) &= (1, -1) \odot \left(([4, 12], [0, 0])^T \ominus_{gH} ([2, 6], [2, 8])^T \right) \\
&= (1, -1) \odot ([2, 6], [-8, -2])^T \\
&= [4, 14] \\
&\neq [-2, 16].
\end{aligned}$$

13

4. Interval Optimization Problem and its Efficient Solutions

This section explores the connection between solutions and the gH-derivatives of the following IOP:

$$\min_{x \in \mathcal{X} \subseteq \mathbb{R}^n} \mathbf{F}(x), \tag{2}$$

where $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is a gH-differentiable function.

The concept of an efficient solution of the IOP (2) is defined below.

Definition 4.1. (*Efficient solution* [59]). A point $\bar{x} \in \mathcal{X}$ is called a global efficient solution of the IOP (2) if $\mathbf{F}(x) \not\prec \mathbf{F}(\bar{x})$ for all $x(\neq \bar{x}) \in \mathcal{X}$.

A point $\bar{x} \in \mathcal{X}$ is called a local efficient solution of IOP (2) if there exists a δ-neighborhood $N_\delta(\bar{x})$ of $\bar{x}$ such that

$$\mathbf{F}(x) \not\prec \mathbf{F}(\bar{x}) \text{ for all } x(\neq \bar{x}) \in N_\delta(\bar{x}) \cap \mathcal{X}.$$

It is to mention that here that Wu [59] named the efficient solution of this article as non-domiated solution. However, throughout this article, we follow Definition 4.1 for an efficient solution, and in the rest of the article, by an efficient solution we mean a global efficient solution.

Since an IOP is a special case of a fuzzy optimization problem [24], the following theorem can be considered as a corollary of Theorem 10 of [53]. Further, as differentiability of a function is a special case of Gâteaux differentiability, the following theorem can also be considered as a corollary of Theorem 4.2 of [27]. In this article, as we are dealing with interval optimization problems and gH-differentiability of interval-valued functions, we show the proof of the following theorem to enhance the readability. However, we prove the following theorem with a different approach than [27].

Theorem 4.1. (Optimality condition). *Let $\boldsymbol{F}$ be a gH-differentiable IVF on a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$. If $\bar{x} \in \mathcal{X}$ is an efficient solution of the IOP (2), then*

$$0 \in d^T \odot \nabla \boldsymbol{F}(\bar{x}) \text{ for all } d \in \mathbb{R}^n.$$

Proof. Let $\bar{x} \in \mathcal{X}$ be an efficient solution of the IOP (2). Therefore, for all $d \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ so that $\bar{x} + \lambda d \in \mathcal{X}$, we have

$$\mathbf{F}(\bar{x} + \lambda d) \not\prec \mathbf{F}(\bar{x})$$

or, $\mathbf{F}(\bar{x} + \lambda d) \ominus_{gH} \mathbf{F}(\bar{x}) \not\prec \mathbf{0}$, by Lemma 2.1

or, $\lim_{\lambda \to 0+} \frac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda d) \ominus_{gH} \mathbf{F}(\bar{x})) \not\prec \mathbf{0}$

or, $\lim_{\lambda \to 0} \frac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda d) \ominus_{gH} \mathbf{F}(\bar{x})) \not\prec \mathbf{0}$, since $\mathbf{F}$ is gH-differentiable on $\mathcal{X}$.

Therefore, by Lemma 3.1 and Theorem 3.1, we obtain

$$d^T \odot \nabla \mathbf{F}(\bar{x}) \nprec \mathbf{0} \text{ for all } d \in \mathbb{R}^n. \tag{3}$$

Replacing d by $-d$ in (3) we obtain

$$(-d)^T \odot \nabla \mathbf{F}(\bar{x}) \nprec \mathbf{0}$$
$$\text{or, } (-1) \odot (d^T \odot \nabla \mathbf{F}(\bar{x})) \nprec \mathbf{0}.$$

Thus, by Lemma 2.2,

$$\mathbf{0} \nprec d^T \odot \nabla \mathbf{F}(\bar{x}). \tag{4}$$

By (3) and (4), we get $0 \in d^T \odot \nabla \mathbf{F}(\bar{x})$ for all $d \in \mathbb{R}^n$. $\qquad\square$

Remark 12. It is noteworthy that the converse of Theorem 4.1 is not always true even if $\mathbf{F}$ is a convex IVF. For instance, let us consider the following IOP:

$$\min_{x \in \mathbb{R}} \mathbf{F}(x), \tag{5}$$

where

$$\mathbf{F}(x) = \begin{cases} [0,3] \ominus_{gH} [-1,0] \odot x^2, & \text{if } -1 \leq x \leq 1 \\ [0,2] \oplus [1,1] \odot x^2, & \text{otherwise.} \end{cases}$$

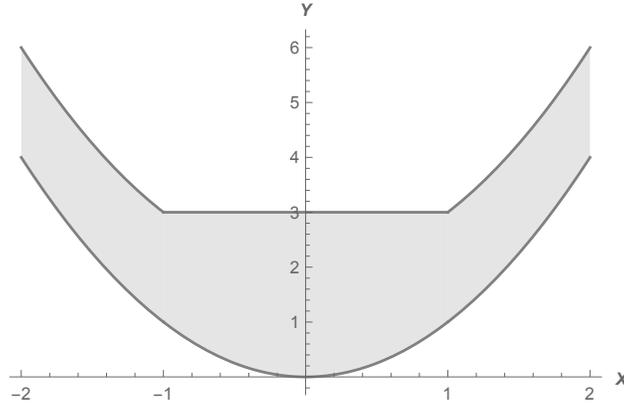Figure 2: Interval-valued function of the IOP (5)

The graph of the IVF $\mathbf{F}$ is depicted in Figure 2 by the shaded region. From Figure 2 it is clear that the IVF $\mathbf{F}$ is convex since its lower and upper functions are convex.

The gH-gradient of $\mathbf{F}$ is

$$\nabla \mathbf{F}(x) = \begin{cases} [0,2] \odot x, & \text{if } -1 \leq x \leq 1 \\ [2,2] \odot x, & \text{otherwise.} \end{cases}$$

15

Thus, at $x = -1$,

$$\nabla \mathbf{F}(-1) = [-2, 0] \implies 0 \in d \odot \nabla \mathbf{F}(-1) \text{ for all } d \in \mathbb{R}.$$

But it is notable that for $0 < h < 2$,

$$\mathbf{F}(-1 + h) = [(-1 + h)^2, 3] \prec [1, 3] = \mathbf{F}(-1).$$

Therefore, although $0 \in d \odot \nabla \mathbf{F}(-1)$ for all $d \in \mathbb{R}$, -1 is not an efficient solution of the IOP (5).

Corollary 4.1. Let $\mathbf{F}$ be a gH-differentiable IVF on a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$. If $\bar{x} \in \mathcal{X}$ is an efficient solution of the IOP (2), then

$$0 \in D_i \mathbf{F}(\bar{x}) \text{ for each } i \in \{1, 2, \ldots, n\}.$$

Proof. Let $\bar{x} \in \mathcal{X}$ be an efficient solution of the IOP (2). According to Theorem 4.1, for all $d \in \mathbb{R}^n$, we have

$$0 \in d^T \odot \nabla \mathbf{F}(\bar{x}) = \bigoplus_{i=1}^{n} d_i \odot D_i \mathbf{F}(\bar{x}).$$

For each $i \in \{1, 2, \ldots, n\}$, by considering $d = e_i$, we obtain

$$0 \in D_i \mathbf{F}(\bar{x}).$$

$\square$

5. gH-gradient Efficient Methods for Interval Optimization Problem

This section develops gH-gradient efficient techniques to obtain the efficient solutions of the IOP (2). In the conventional gradient descent technique, to find a minimizer, we move sequentially along descent directions. Likewise, for IOP, to find an efficient solution we attempt to move sequentially along efficient-directions, defined below.

Definition 5.1. (*Efficient-direction*). Let $\mathcal{X} \subseteq \mathbb{R}^n$. A direction $d \in \mathbb{R}^n$ is said to be an efficient-direction of an IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ at $\bar{x} \in \mathcal{X}$ if there exists a $\delta > 0$ such that

(i) $\mathbf{F}(\bar{x}) \not\preceq \mathbf{F}(\bar{x} + \lambda d)$ for all $\lambda \in (0, \delta)$,

(ii) there also exists a point $x' = \bar{x} + \alpha d$ with $\alpha \in (0, \delta)$ and a positive real number $\delta' \leq \alpha$ such that

$$\mathbf{F}(x' + \lambda d) \not\succ \mathbf{F}(x') \text{ for all } \lambda \in (-\delta', \delta').$$

The point x' is known as an efficient point of $\mathbf{F}$ in the direction d.

In Figure 3, the points $\bar{x}$ and x', the direction d, and the nonnegative real numbers δ and δ' of Definition 5.1 are illustrated on $\mathbb{R}^2$ plane ($n = 2$).
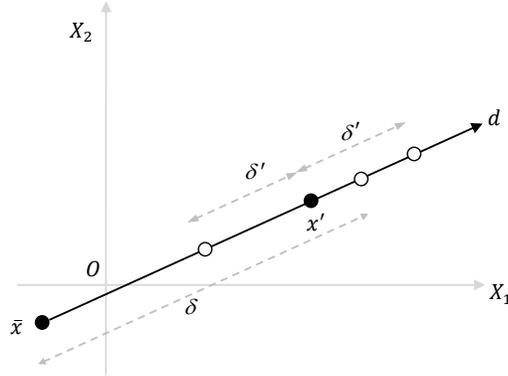
Figure 3: Locations of $\bar{x}$ and x' in Definition 5.1

Remark 13. One may think that only the condition (i) of Definition 5.1 is sufficient to define an efficient-direction. However, it is not true in general. Because, for **A**, **B** and $\mathbf{C} \in I(\mathbb{R})$,

$$\mathbf{A} \not\preceq \mathbf{B} \text{ and } \mathbf{B} \not\preceq \mathbf{C} \implies \mathbf{A} \not\preceq \mathbf{C} \text{ in general.}$$

For instance, consider

$$\mathbf{A} = [4, 6], \quad \mathbf{B} = [2, 10] \text{ and } \mathbf{C} = [5, 7].$$

We, then, see that

$$\mathbf{A} \not\preceq \mathbf{B} \text{ and } \mathbf{B} \not\preceq \mathbf{C} \text{ but } \mathbf{A} \prec \mathbf{C}.$$

That is why the condition (ii) of Definition 5.1 is necessary to define an efficient direction.

Remark 14. For the degenerate case of the IVF **F**, i.e., for $\underline{f}(x) = \overline{f}(x) = f(x)$ for all $x \in \mathcal{X}$, Definition 5.1 reduces to the following. A direction $d \in \mathbb{R}^n$ is said to be an efficient-direction of f if there exists a $\delta > 0$ such that

$$f(\bar{x}) > f(\bar{x} + \lambda d) \text{ for all } \lambda \in (0, \delta).$$

Thus, an efficient direction for a degenerate IVF is a descent direction.

Theorem 5.1. *Let* **F** *be a* gH-*differentiable IVF on a nonempty subset* $\mathcal{X}$ *of* $\mathbb{R}^n$. *Then, every direction* $d \in \mathbb{R}^n$ *that satisfies*

$$\mathbf{0} \not\preceq d^T \odot \nabla\mathbf{F}(\bar{x}), \tag{6}$$

is an efficient-direction of **F** *at* $\bar{x} \in \mathcal{X}$, *where the corresponding efficient point* x' *is provided by*

$$x' = \bar{x} + \alpha'd \text{ with } \alpha' = \underset{\alpha \in \mathbb{R}_+}{\operatorname{argeff}} \mathbf{F}(\bar{x} + \alpha d),$$

*where by '*argeff*' of* $\mathbf{F}(\bar{x} + \alpha d)$, *we mean a point* α' *such that*

$$\mathbf{F}(\bar{x} + \alpha d) \not\prec \mathbf{F}(\bar{x} + \alpha'd) \text{ for all } \alpha \in \mathbb{R}.$$

17

Proof. Let $d \in \mathbb{R}^n$ be a direction that satisfies the relation (6). Since $\mathbf{F}$ is gH-differentiable at $\bar{x} \in \mathcal{X}$, by Lemma 3.1 and Theorem 3.1, we have

$$\lim_{\lambda \to 0} \frac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda d) \ominus_{gH} \mathbf{F}(\bar{x})) = d^T \odot \nabla \mathbf{F}(\bar{x}),$$

which implies

$$\lim_{\lambda \to 0+} \frac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda d) \ominus_{gH} \mathbf{F}(\bar{x})) = d^T \odot \nabla \mathbf{F}(\bar{x}).$$

Due to the relation (6), the last equation yields

$$\mathbf{0} \not\preceq \lim_{\lambda \to 0+} \frac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda d) \ominus_{gH} \mathbf{F}(\bar{x})). \tag{7}$$

Therefore,

$$[0,0] \not\preceq \lim_{\lambda \to 0+} \frac{1}{\lambda} \odot \left[\min \left\{ \underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}), \overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right\}, \right.$$
$$\left. \max \left\{ \underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}), \overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right\} \right]$$

$$\text{or, } [0,0] \not\preceq \left[\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right) \right\}, \right.$$
$$\left. \max \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right) \right\} \right],$$

which implies

$$\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right) \right\} < 0. \tag{8}$$

Thus, we have following two cases.

• **Case** 1. Let

$$\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right) \right\}$$
$$= \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right).$$

Then, by the equation (8), we obtain $\lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right) < 0$. Therefore, there exists a $\delta_1 > 0$ such that for all $\lambda \in (0, \delta_1)$,

$$\frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right) < 0$$
$$\text{or, } \underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) < 0$$
$$\text{or, } \underline{f}(\bar{x} + \lambda d) < \underline{f}(\bar{x})$$
$$\text{or, } \left[\underline{f}(\bar{x}), \overline{f}(\bar{x}) \right] \not\preceq \left[\underline{f}(\bar{x} + \lambda d), \overline{f}(\bar{x} + \lambda d) \right].$$

- **Case** 2. Let

$$\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\underline{f}(\bar{x} + \lambda d) - \underline{f}(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right) \right\}$$

$$= \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right).$$

Then, by the equation (8), we get

$$\lim_{\lambda \to 0+} \frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right) < 0.$$

Thus, there exists a $\delta_2 > 0$ such that for $\lambda \in (0, \delta_2)$,

$$\frac{1}{\lambda} \left(\overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) \right) < 0$$
$$\text{or, } \overline{f}(\bar{x} + \lambda d) - \overline{f}(\bar{x}) < 0$$
$$\text{or, } \overline{f}(\bar{x} + \lambda d) < \overline{f}(\bar{x})$$
$$\text{or, } \left[\underline{f}(\bar{x}), \overline{f}(\bar{x}) \right] \not\preceq \left[\underline{f}(\bar{x} + \lambda d), \overline{f}(\bar{x} + \lambda d) \right]$$

Choosing $\delta = \min\{\delta_1, \delta_2\}$, from Case 1 and Case 2, we see that for all $\lambda \in (0, \delta)$,

$$\left[\underline{f}(\bar{x}), \overline{f}(\bar{x}) \right] \not\preceq \left[\underline{f}(\bar{x} + \lambda d), \overline{f}(\bar{x} + \lambda d) \right]$$
$$\text{or, } \mathbf{F}(\bar{x}) \not\preceq \mathbf{F}(\bar{x} + \lambda d).$$

Hence, d satisfies the condition (i) of Definition 5.1 at $\bar{x}$.

Further, let us choose an α' such that

$$\alpha' = \underset{\alpha \in \mathbb{R}_+}{\text{argeff}} \ \mathbf{F}(\bar{x} + \alpha d).$$

Therefore, there exists a $\bar{\delta} > 0$ such that for all $\alpha \in (\alpha' - \bar{\delta}, \alpha' + \bar{\delta})$,

$$\mathbf{F}(\bar{x} + \alpha d) \not\prec \mathbf{F}(\bar{x} + \alpha' d).$$

Considering $\lambda = \alpha - \alpha'$, we have

$$\mathbf{F}(\bar{x} + \alpha' d + \lambda d) \not\prec \mathbf{F}(\bar{x} + \alpha' d)$$
$$\text{or, } \mathbf{F}(x' + \lambda d) \not\prec \mathbf{F}(x'),$$

where $x' = \bar{x} + \alpha' d$. Choosing $\delta' = \min\{\alpha', \bar{\delta}\}$, we have

$$\mathbf{F}(x' + \lambda d) \not\prec \mathbf{F}(x') \ \text{ for all } \ \lambda \in (-\delta', \delta').$$

Therefore, d satisfies the condition (ii) of Definition 5.1 at $\bar{x}$. Hence, d is an efficient-direction at $\bar{x}$. $\qquad\square$

Remark 15. A question may arise here: in the definition of efficient-direction (Definition 5.1), whether or not the condition (i) can be replaced by

$$\mathbf{F}(\bar{x} + \lambda d) \prec \mathbf{F}(\bar{x}) \text{ for all } \lambda \in (0, \delta)? \tag{9}$$

To answer, we note that if we choose the relation (9) in place of the condition (i) of Definition 5.1, then in the same way of proving Theorem 5.1 it can be proved that any $d \in \mathbb{R}^n$ that satisfies

$$d^T \odot \nabla \mathbf{F}(\bar{x}) \prec \mathbf{0} \tag{10}$$

also holds the relation (9) and vice versa for a gH-differentiable IVF F. However, the relation (6) is more general than the relation (10). Because there are some directions, along which there exists an efficient solution of an IOP, satisfy the relation (6) but do not satisfy the relation (10).

For instance, consider the IOP:

$$\min_{x \in \mathcal{X}} \mathbf{F}(x) = [-1, 2] \odot x_1^2 \oplus [3, 4] \odot x_2^2, \tag{11}$$

where $\mathcal{X} = [-10, 10] \times [-10, 10] \subseteq \mathbb{R}^2$.

In what follows, we show that $(\bar{x}_1, 0)^T \in \mathcal{X}$ is an efficient solution of the IOP (11) for any $\bar{x}_1 \in [-10, 10]$. On contrary, let there exist two nonzero numbers h_1 and h_2 with $(\bar{x}_1 + h_1, h_2)^T \in \mathcal{X}$ such that

$$\mathbf{F}(\bar{x}_1 + h_1, h_2) \prec \mathbf{F}(\bar{x}_1, 0)$$
$$\text{or, } [-1, 2] \odot (\bar{x}_1 + h_1)^2 \oplus [3, 4] h_2^2 \prec [-1, 2] \odot \bar{x}_1^2$$
$$\text{or, } [-\bar{x}_1^2 - 2\bar{x}_1 h_1 - h_1^2 + 3h_2^2, 2x_1^2 + 4\bar{x}_1 h + 2h_1^2 + 4h_2^2] \prec [-\bar{x}_1^2, 2\bar{x}_1^2].$$

This implies

$$-2x_1 h_1 - h_1^2 + 3h_2^2 \leq 0 \text{ and } 4x_1 h_1 + 2h_1^2 + 4h_2^2 \leq 0.$$

Hence, $h_2^2 \leq 0$, which is not possible as $h_2 \neq 0$. So, there does not exist any $x \in \mathcal{X}$, which strictly dominates any $(\bar{x}_1, 0)^T \in \mathcal{X}$. Thus, $(\bar{x}_1, 0)^T$ is an efficient solution of the IOP (11).

Now we choose a point $\hat{x} = (3, 2)^T$ and two directions $d' = (1, -2)^T$ and $d'' = (5, -2)^T$. We also choose an $\alpha = 1$. As

$$\mathbf{F}(3, 0) = [-1, 2] \prec [5, 10] = \mathbf{F}(3, 2) = \mathbf{F}(\hat{x}),$$

$\hat{x}$ is not an efficient solution of the IOP (11). But the points $x' = \hat{x} + \alpha d' = (4, 0)^T$ and $x'' = \hat{x} + \alpha d'' = (8, 0)^T$ both are efficient solutions of the IOP (11). Therefore, both d' and d'' are efficient-directions of $\mathbf{F}$ at $\hat{x}$. Further, as the gH-gradient of $\mathbf{F}$ is

$$\nabla \mathbf{F}(x) = (D_1 \mathbf{F}(x), D_2 \mathbf{F}(x))^T = ([-2, 4] \odot x_1, [6, 8] \odot x_2)^T,$$

we have $\nabla\mathbf{F}(\hat{x}) = ([-6, 12], [12, 16])^T$. Therefore,

$$d' \odot \nabla\mathbf{F}(\hat{x}) = [-6, 12] \oplus [-32, -24] = [-38, -12]$$
$$\implies d' \odot \nabla\mathbf{F}(\hat{x}) \prec \mathbf{0} \text{ and also, } \mathbf{0} \not\preceq d' \odot \nabla\mathbf{F}(\hat{x}).$$

Again,

$$d'' \odot \nabla\mathbf{F}(\hat{x}) = [-30, 60] \oplus [-32, -24] = [-38, 36]$$
$$\implies \mathbf{0} \not\preceq d'' \odot \nabla\mathbf{F}(\hat{x})$$

Hence, it is clear that although both d' and d'' are efficient-directions of the IVF $\mathbf{F}$ at the point $\hat{x}$ and satisfy the relation (6) but only d' satisfies the relation (10). So, the condition (i) of Definition 5.1 is more general than the condition (10).

5.1. General gH-gradient Efficient-Direction Method for Interval Optimization Problems
To produce the efficient solutions of the IOP (2) we provide Algorithm 1. As Algorithm 1

(i) uses gH-gradient at every iterative step and

(ii) endeavors to find an efficient solution by sequentially moving along efficient-directions,

we name the method as general gH-gradient efficient-direction method. The term 'general' is due to the reason that we do not choose a special or a particular efficient-direction d_k; any general d_k that satisfies $\mathbf{0} \not\preceq d_k^T \odot \nabla\mathbf{F}(x_k)$ will lead to reaching at an efficient point.

Algorithm 1 General gH-gradient efficient-direction method for IOP

Require: Given the initial point x_0 and the IVF $\mathbf{F} : \mathcal{X}(\subseteq \mathbb{R}^n) \to I(\mathbb{R})$.
1: Set $k = 0$.
2: If $0 \in d^T \odot \nabla\mathbf{F}(x_k)$ for all $d \in \mathbb{R}^n$, then Return x_k as an efficient-solution and Stop. Otherwise go to Step 3.
3: Find a d_k such that $\mathbf{0} \not\preceq d_k^T \odot \nabla\mathbf{F}(x_k)$ and an α_k such that

$$\alpha_k = \underset{\alpha \in \mathbb{R}_+}{\text{argeff}} \ \mathbf{F}(x_k + \alpha d_k)$$

4: Calculate

$$x_{k+1} = x_k + \alpha_k d_k.$$

5: Set $k \leftarrow k + 1$ and go to Step 2.

In the next, we give the convergence analysis of the Algorithm 1. Towards the convergent analysis, we need the following notions of the algorithmic map and the closed map regarding IVFs.

Definition 5.2. (*Algorithmic map* [4]). Let $\mathcal{X}$ be a nonempty subset of $\mathbb{R}^n$. An algorithmic map $\mathcal{A}$ of an algorithm is a point-to-set map on its domain $\mathcal{X}$ which describes the iterating process of the algorithm such that if the sequence $\{x_k\}$ is generated by the algorithm then $x_{k+1} \in \mathcal{A}(x_k)$.

Remark 16. The map $\mathcal{A} : \mathcal{X} \to \mathcal{X}$ that generates the sequence $\{x_k\}$ in Algorithm 1 with $\mathbf{F}(x_k) \not\preceq \mathbf{F}(x_{k+1})$ is an algorithmic map.

Definition 5.3. (*Closed map* [4]). A point-to-set map $\mathcal{A}$ from a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$ to a subset $\mathcal{Y}$ of $\mathbb{R}^m$ is said to be closed at $x \in \mathcal{X}$ if for any sequences $\{x_k\}$ and $\{y_k\}$ such that

$$x_k \to x \text{ and } y_k \to y, \text{ where } x_k \in \mathcal{X} \text{ and } y_k \in \mathcal{A}(x_k),$$

we have $y \in \mathcal{A}(x)$.

Theorem 5.2. *Let $\mathbf{F} : \mathbb{R}^n \to I(\mathbb{R})$ be an IVF on a nonempty open subset $\mathcal{X}$ of $\mathbb{R}^n$ and $\mathbf{D}$ be an element of $I(\mathbb{R}_+)$. Define a point-to-set map $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by $\mathcal{L}(x, d) = \{y : y = x + \bar{\lambda}d\}$ for some $\bar{\lambda} \in \mathbf{D}$. Let $\mathbf{F}(x + \lambda d) \not\prec \mathbf{F}(y)$ for each $\lambda \in \mathbf{D}$ and $d \neq 0$. If $\mathbf{F}$ is gH-continuous at x and $d \neq 0$, then $\mathcal{L}$ is closed at (x, d).*

Proof. Let $\{(x_k, d_k)\}$ be a sequence such that $(x_k, d_k) \to (x, d)$. Let $\{y_k\}$, $y_k \in \mathcal{L}(x_k, d_k)$, be a sequence such that $y_k \to y$. To prove the theorem, we have to show that $y \in \mathcal{L}(x, d)$.

It is to note that there exists $\lambda_k \in \mathbf{D}$ such that $y_k = x_k + \lambda_k d_k$ for $k = 1, 2, \ldots, n$. Since $d \neq 0$, for large enough k we have $d_k \neq 0$. Then, $\lambda_k = \frac{\|y_k - x_k\|}{\|d_k\|}$.

Taking the limit as $k \to \infty$ we have $\lambda_k \to \frac{\|y - x\|}{\|d\|} = \bar{\lambda}$, say. Hence, $y = x + \bar{\lambda}d$. Furthermore, since $\lambda_k \in \mathbf{D}$ for each k and $\mathbf{D}$ is closed, $\bar{\lambda} \in \mathbf{D}$.

Therefore, as $\lambda \in \mathbf{D}$, $\mathbf{F}(x_k + \lambda d_k) \not\prec \mathbf{F}(y_k)$ for all k, and $\mathbf{F}$ is gH-continuous, by Lemma 2.4, evidently, we have $\mathbf{F}(x + \lambda d) \not\prec \mathbf{F}(y)$. Hence, $y \in \mathcal{L}(x, d)$ and so, $\mathcal{L}$ is closed. $\qquad\square$

Remark 17. As according to Theorem 5.2, the map $\mathcal{L}$ is closed, the composite map $\mathcal{A} = \mathcal{L} \circ \mathcal{D}$ will be closed if the direction generating map $\mathcal{D} : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$, defined by $\mathcal{D}(x) = (x, d)$, is also closed.

Theorem 5.3. (Convergence of general gH-gradient efficient method for IOP). *Let $\mathbf{F}$ be a gH-differentiable IVF on a nonempty open subset $\mathcal{X}$ of $\mathbb{R}^n$. Suppose $\Omega \subseteq \mathcal{X}$ be the set of all efficient points of $\mathbf{F}$ and $\mathcal{A} : \mathcal{X} \to \mathcal{X}$ be an algorithmic map of Algorithm 1. Suppose that the algorithm map $\mathcal{A}$ produces the sequence $\{x_k\}$, which converges at $\bar{x} \in \mathcal{X}$. Also, assume that there exists $M > 0$ such that $\|d_k\| < M$ for all k. If $\bar{d}$ is an accumulation point of d_k, then we have*

$$0 \in \bar{d}^T \odot \nabla \mathbf{F}(\bar{x}). \tag{12}$$

Proof. Since $\{d_k\}$ is bounded, there exists an index set K_1 such that $\lim\limits_{k \in K_1} d_k = \bar{d}$ and $\lim\limits_{k \in K_1} x_k = \bar{x}$. We have the following two cases.

- **Case** 1. *If $\bar{d} = 0$. Then, (12) is trivial.*

- **Case** 2. *If $\bar{d} \neq 0$.* If possible, let us assume that the conclusion (12) is not true. Hence,

$$0 \notin \bar{d}^T \odot \nabla \mathbf{F}(\bar{x}) \Longrightarrow \bar{x} \notin \Omega.$$

Since, for $k \in K_1, x_k \to \bar{x}$, also $x_{k+1} \in \mathcal{A}(x_k)$ and $x_{k+1} \to \bar{x}$, therefore, $\bar{x} \in \mathcal{A}(\bar{x})$. As $\mathcal{A}$ is closed at $\bar{x}$ due to Remark 17. Thus, $\mathbf{F}(\bar{x}) \npreceq \mathbf{F}(\bar{x})$, which is a contradiction. Hence, $0 \in \bar{d}^T \odot \nabla \mathbf{F}(\bar{x})$.

$\square$

5.2. $\mathcal{W}$-gradient Method for Interval Optimization Problems

In this section, we develop a particular type of efficient-direction. Towards this, for two given numbers $w, w' \in [0, 1]$ with $w + w' = 1$, we define a mapping $\mathcal{W} : I(\mathbb{R})^n \to \mathbb{R}^n$ by

$$\mathcal{W}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n) = (w\underline{a}_1 + w'\overline{a}_1, w\underline{a}_2 + w'\overline{a}_2, \ldots, w\underline{a}_n + w'\overline{a}_n)^T.$$

Remark 18. It is to observe that for any two elements $\bar{\mathbf{A}}, \bar{\mathbf{B}}$ in $I(\mathbb{R})^n$,

$$\mathcal{W}(\bar{\mathbf{A}} \oplus \bar{\mathbf{B}}) = \mathcal{W}(\bar{\mathbf{A}}) + \mathcal{W}(\bar{\mathbf{B}}).$$

Lemma 5.1. For an interval $\mathbf{A} = [\underline{a}, \overline{a}]$, $(w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] \nprec \mathbf{0}$, where $w, w' \in [0, 1]$ with $w + w' = 1$.

Proof. For an interval $\mathbf{A} = [\underline{a}, \overline{a}]$, we have the following three cases.

- **Case** 1. If $\underline{a} \geq 0$ *and* $\overline{a} \geq 0$. Then, $\underline{a}(w\underline{a} + w'\overline{a}) \geq 0$ and $\overline{a}(w\underline{a} + w'\overline{a}) \geq 0$. Hence,

$$\mathbf{0} \preceq (w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] \implies (w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] \nprec \mathbf{0}.$$

- **Case** 2. If $\underline{a} < 0$ *and* $\overline{a} \leq 0$. Then, $\underline{a}(w\underline{a} + w'\overline{a}) \geq 0$ and $\overline{a}(w\underline{a} + w'\overline{a}) \geq 0$. Hence,

$$\mathbf{0} \preceq (w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] \implies (w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] \nprec \mathbf{0}.$$

- **Case** 3. If $\underline{a} \leq 0$ *and* $\overline{a} > 0$. Then, either $(w\underline{a} + w'\overline{a}) = 0$ or $(w\underline{a} + w'\overline{a}) \neq 0$. If $(w\underline{a} + w'\overline{a}) = 0$, we have

$$(w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] = \mathbf{0} \implies (w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] \nprec \mathbf{0}.$$

Further, if $(w\underline{a} + w'\overline{a}) \neq 0$, the terms $\underline{a}(w\underline{a} + w'\overline{a})$ and $\overline{a}(w\underline{a} + w'\overline{a})$ are alternative in sign. So,

$$(w\underline{a} + w'\overline{a}) \odot [\underline{a}, \overline{a}] \nprec \mathbf{0}.$$

$\square$

Theorem 5.4. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a nonempty set and $\mathbf{F}(x) : \mathcal{X} \to I(\mathbb{R})$ be gH-differentiable at a point $\bar{x} \in \mathcal{X}$. Then, the direction $-\mathcal{W}(\nabla \mathbf{F}(\bar{x}))$ is an efficient-direction of $\mathbf{F}$ at $\bar{x}$, provided $0 \notin D_i \mathbf{F}(\bar{x})$ for at least one $i \in \{1, 2, \ldots, n\}$.*

23

Proof. Let $D_i\mathbf{F}(\bar{x}) = \mathbf{A}_i$ for each $i \in \{1, 2, \ldots, n\}$ and $0 \notin \mathbf{A}_i$ for at least one i. Thus,

$$(\mathcal{W}(\nabla\mathbf{F}(\bar{x})))^T \odot \nabla\mathbf{F}(\bar{x}) = \bigoplus_{i=1}^{n}(w\underline{a}_i + w'\overline{a}_i) \odot [\underline{a}_i, \overline{a}_i].$$

Let $0 \notin \mathbf{A}_i$ for $i = j$. Therefore, $\underline{a}_j$ and $\overline{a}_j$ both are either positive or negative. Thus, $(w\underline{a}_j + w'\overline{a}_j) \neq 0$, and hence $(w\underline{a}_j + w'\overline{a}_j) \odot [\underline{a}_j, \overline{a}_j] \neq \mathbf{0}$. Therefore, according to the property of interval addition we get

$$(\mathcal{W}(\nabla\mathbf{F}(\bar{x})))^T \odot \nabla\mathbf{F}(\bar{x}) \neq \mathbf{0}. \tag{13}$$

If possible, let $(\mathcal{W}(\nabla\mathbf{F}(\bar{x})))^T \odot \nabla\mathbf{F}(\bar{x}) \prec \mathbf{0}$. Hence, $(w\underline{a}_i + w'\overline{a}_i) \odot [\underline{a}_i, \overline{a}_i] \prec \mathbf{0}$ for at least one $i \in \{1, 2, \ldots, n\}$, which is not possible due to Lemma 5.1. Therefore,

$$(\mathcal{W}(\nabla\mathbf{F}(\bar{x})))^T \odot \nabla\mathbf{F}(\bar{x}) \nprec \mathbf{0} \implies \mathbf{0} \nsucc -(\mathcal{W}(\nabla\mathbf{F}(\bar{x})))^T \odot \nabla\mathbf{F}(\bar{x}). \tag{14}$$

By relations (13) and (14), we obtain

$$\mathbf{0} \npreceq -(\mathcal{W}(\nabla\mathbf{F}(\bar{x})))^T \odot \nabla\mathbf{F}(\bar{x}).$$

Hence, by Theorem 5.1, $-\mathcal{W}(\nabla\mathbf{F}(\bar{x}))$ is an efficient-direction of $\mathbf{F}$ at $\bar{x}$, where the corresponding efficient point x', given by

$$x' = \bar{x} - \alpha'\mathcal{W}(\nabla\mathbf{F}(\bar{x})) \text{ with } \alpha' = \underset{\alpha \in \mathbb{R}_+}{\mathrm{argeff}} \ \mathbf{F}(\bar{x} - \alpha\mathcal{W}(\nabla\mathbf{F}(\bar{x}))),$$

satisfies the condition (i) of Definition 5.1 at $\bar{x}$. $\qquad\square$

Based on Theorem 5.4 and Corollary 4.1, Algorithm 1 is reduced to the following Algorithm 2. As Algorithm 2 is a particular case of Algorithm 1, we name the method as $\mathcal{W}$-gH-gradient efficient-direction method. The letter '$\mathcal{W}$' is due to the reason that we use the mapping $\mathcal{W} : I(\mathbb{R})^n \to \mathbb{R}^n$ to generate efficient-direction at each iteration in Algorithm 2.

Algorithm 2 $\mathcal{W}$-gH-Gradient Method for IOP

Require: Give $w \in [0, 1]$, the initial point x_0 and the IVF $\mathbf{F} : \mathcal{X}(\subseteq \mathbb{R}^n) \to I(\mathbb{R})$.
 1: Set $k = 0$.
 2: If $0 \in D_i\mathbf{F}(x_k)$ for all $i \in \{1, 2, \ldots, n\}$, then Return x_k as an efficient-solution and Stop. Otherwise go to Step 3.
 3: Set $d_k = -\mathcal{W}(\nabla\mathbf{F}(x_k))$ and find an α_k, where

$$\alpha_k = \underset{\alpha \in \mathbb{R}_+}{\mathrm{argeff}} \ \mathbf{F}(x_k - \alpha d_k).$$

 4: Calculate

$$x_{k+1} = x_k + \alpha_k d_k.$$

 5: Set $k \leftarrow k + 1$ and go to Step 2.

Remark 19. It is to be mentioned that in Algorithm 2, for the degenerate case of the IVF $\mathbf{F}$,

i.e., for $\underline{f}(x) = \overline{f}(x) = f(x)$ for all $x \in \mathcal{X}$, the direction d_k will be $-\nabla f(x_k)$ and step length α_k will be $\operatorname{argmin} f(x_k + \alpha d_k)$ at each iteration k, for any values of $w, w' \geq 0$ with $w + w' = 1$ in the mapping $\mathcal{W}$. Thus, in that case, $\mathcal{W}$-gH-gradient efficient method is same as steepest descent method.

Remark 20. At each iteration k in Algorithm 2, one may choose α_k such that

$$\alpha_k \in \{\alpha \mid w\underline{a} + w'\overline{a} = 0, \text{ where } [\underline{a}, \overline{a}] = \mathbf{F}'(x_k + \alpha d_k)\}, \tag{15}$$

where $\mathbf{F}'(x_k + \alpha d_k)$ is the gH-derivative of $\mathbf{F}(x_k + \alpha d_k)$ with respect to α. But it is to be kept in mind that (15) is a necessary condition to be α_k an argeff of $\mathbf{F}(x_k - \alpha d_k)$, not sufficient.

Remark 21. One question may arise: for a given pair of non-negative w and w' with $w + w' = 1$, if we consider the real-valued function $w\underline{f}(x) + w'\overline{f}(x)$ corresponding to the IOP (2) and apply the conventional steepest descent method, whether the obtained direction and step length in each iteration are identical with those obtained in each iteration of $\mathcal{W}$-gH-gradient efficient method?

The answer is *yes* only when

$$D_i \mathbf{F}(x) = \left[\frac{\partial \underline{f}(x)}{\partial x_i}, \frac{\partial \overline{f}(x)}{\partial x_i} \right] \text{ for all } i \in \{1, 2, \ldots, n\} \text{ at each } x \in \mathcal{X},$$

which is not true in general (see [11] for details).

Lemma 5.2. For two elements $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ in $I(\mathbb{R})^n$ and a vector $h \in \mathbb{R}^n$,

$$\mathbf{0} \preceq h^T \odot \left(\bar{\mathbf{A}} \ominus_{gH} \bar{\mathbf{B}} \right) \implies \left(\mathcal{W}\left(\bar{\mathbf{A}}\right) - \mathcal{W}\left(\bar{\mathbf{B}}\right) \right)^T h \geq 0.$$

Proof.

$$h^T \odot \left(\bar{\mathbf{A}} \ominus_{gH} \bar{\mathbf{B}} \right) = h^T \odot \left((\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)^T \ominus_{gH} (\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n)^T \right)$$
$$= h^T \odot \left((\mathbf{A}_1 \ominus_{gH} \mathbf{B}_1), (\mathbf{A}_2 \ominus_{gH} \mathbf{B}_2), \ldots, (\mathbf{A}_n \ominus_{gH} \mathbf{B}_n) \right)^T$$
$$= \bigoplus_{i=1}^{n} h_i \odot (\mathbf{A}_i \ominus_{gH} \mathbf{B}_i).$$

According to Remark 1, without loss of generality, let us assume

$$\bigoplus_{i=1}^{n} h_i \odot (\mathbf{A}_i \ominus_{gH} \mathbf{B}_i) = \bigoplus_{k=1}^{p} h_k \odot (\mathbf{A}_k \ominus_{gH} \mathbf{B}_k) \oplus \bigoplus_{l=p+1}^{n} h_l \odot (\mathbf{A}_l \ominus_{gH} \mathbf{B}_l)$$
$$= \bigoplus_{k=1}^{p} \mathbf{C}_k \oplus \bigoplus_{l=p+1}^{n} \mathbf{C}_l,$$

25

where $[\underline{c}_k, \overline{c}_k] = h_k \odot (\mathbf{A}_k \ominus_{gH} \mathbf{B}_k) = \left[(\underline{a}_k - \underline{b}_k) h_k, (\overline{a}_k - \overline{b}_k) h_k \right]$ for all k's and $[\underline{c}_l, \overline{c}_l] = h_l \odot (\mathbf{A}_l \ominus_{gH} \mathbf{B}_l) = \left[(\overline{a}_l - \overline{b}_l) h_l, (\underline{a}_l - \underline{b}_l) h_l \right]$ for all l's. Thus,

$$\mathbf{0} \preceq h^T \odot \left(\bar{\mathbf{A}} \ominus_{gH} \bar{\mathbf{B}} \right) \implies \mathbf{0} \preceq \bigoplus_{k=1}^{p} [\underline{c}_k, \overline{c}_k] \oplus \bigoplus_{l=p+1}^{n} [\underline{c}_l, \overline{c}_l]$$

$$\implies \mathbf{0} \preceq \left[\sum_{k=1}^{p} \underline{c}_k + \sum_{l=p+1}^{n} \underline{c}_l, \sum_{k=1}^{p} \overline{c}_k + \sum_{l=p+1}^{n} \overline{c}_l \right]$$

$$\implies 0 \leq \sum_{k=1}^{p} \underline{c}_k + \sum_{l=p+1}^{n} \underline{c}_l \text{ and } 0 \leq \sum_{k=1}^{p} \overline{c}_k + \sum_{l=p+1}^{n} \overline{c}_l.$$

Therefore,

$$\sum_{k=1}^{p} \underline{c}_k + \sum_{l=p+1}^{n} \overline{c}_l \geq \sum_{k=1}^{p} \underline{c}_k + \sum_{l=p+1}^{n} \underline{c}_l \geq 0 \text{ and } \sum_{k=1}^{p} \overline{c}_k + \sum_{l=p+1}^{n} \underline{c}_l \geq \sum_{k=1}^{p} \underline{c}_k + \sum_{l=p+1}^{n} \underline{c}_l \geq 0$$

$$\implies \sum_{k=1}^{p} (\underline{a}_k - \underline{b}_k) h_k + \sum_{l=p+1}^{n} (\underline{a}_l - \underline{b}_l) h_l \geq 0 \text{ and } \sum_{k=1}^{p} (\overline{a}_k - \overline{b}_k) h_k + \sum_{l=p+1}^{n} (\overline{a}_l - \overline{b}_l) h_l \geq 0$$

$$\implies \sum_{i=1}^{n} (\underline{a}_i - \underline{b}_i) h_i \geq 0 \text{ and } \sum_{i=1}^{n} (\overline{a}_i - \overline{b}_i) h_i \geq 0$$

$$\implies w \sum_{i=1}^{n} (\underline{a}_i - \underline{b}_i) h_i \geq 0 \text{ and } w' \sum_{i=1}^{n} (\overline{a}_i - \overline{b}_i) h_i \geq 0$$

$$\implies \sum_{i=1}^{n} w (\underline{a}_i - \underline{b}_i) h_i + w' (\overline{a}_i - \overline{b}_i) h_i \geq 0$$

$$\implies \left(\mathcal{W} (\bar{\mathbf{A}}) - \mathcal{W} (\bar{\mathbf{B}}) \right)^T h \geq 0.$$

$\square$

Lemma 5.3. If $\mathbf{F}$ is gH-differentiable IVF, then

$$\| \mathcal{W}(\nabla \mathbf{F}(x)) - \mathcal{W}(\nabla \mathbf{F}(y)) \|_{\mathcal{X}} \leq \| \nabla \mathbf{F}(x) \ominus_{gH} \nabla \mathbf{F}(y) \|_{I(\mathbb{R})^n} \text{ for all } x, y \in \mathcal{X}. \quad (16)$$

Proof. Let

$$\nabla \mathbf{F}(x) = \bar{\mathbf{A}} = (\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)^T = ([\underline{a}_1, \overline{a}_1], [\underline{a}_2, \overline{a}_2], \ldots, [\underline{a}_n, \overline{a}_n])^T$$

and

$$\nabla \mathbf{F}(y) = \bar{\mathbf{B}} = (\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n)^T = ([\underline{b}_1, \overline{b}_1], [\underline{b}_2, \overline{b}_2], \ldots, [\underline{b}_n, \overline{b}_n])^T.$$

Therefore,

$$\| \nabla \mathbf{F}(x) \ominus_{gH} \nabla \mathbf{F}(y) \|_{I(\mathbb{R})^n} = \| (\mathbf{A}_1 \ominus_{gH} \mathbf{B}_1, \mathbf{A}_2 \ominus_{gH} \mathbf{B}_2, \ldots, \mathbf{A}_n \ominus_{gH} \mathbf{B}_n) \|_{I(\mathbb{R})^n}$$

$$= \ \|\mathbf{A}_1 \ominus_{gH} \mathbf{B}_1\|_{I(\mathbb{R})} + \|\mathbf{A}_2 \ominus_{gH} \mathbf{B}_2\|_{I(\mathbb{R})} + \cdots + \|\mathbf{A}_n \ominus_{gH} \mathbf{B}_n\|_{I(\mathbb{R})}.$$

We note that

$$\mathcal{W}(\nabla \mathbf{F}(x)) = (w\underline{a}_1 + w'\overline{a}_1, w\underline{a}_2 + w'\overline{a}_2, \dots, w\underline{a}_n + w'\overline{a}_n)^T$$

and

$$\mathcal{W}(\nabla \mathbf{F}(y)) = (w\underline{b}_1 + w'\overline{b}_1, w\underline{b}_2 + w'\overline{b}_2, \dots, w\underline{b}_n + w'\overline{b}_n)^T.$$

So,

$$\begin{aligned}
&\|\mathcal{W}(\nabla \mathbf{F}(x)) - \mathcal{W}(\nabla \mathbf{F}(y))\|_{\mathcal{X}} \\
=\ & \|(w(\underline{a}_1 - \underline{b}_1) + w'(\overline{a}_1 - \overline{b}_1), w(\underline{a}_2 - \underline{b}_2) + w'(\overline{a}_2 - \overline{b}_2), \\
& \qquad \cdots, w(\underline{a}_n - \underline{b}_n) + w'(\overline{a}_n - \overline{b}_n))\|_{\mathcal{X}} \\
\leq\ & |w(\underline{a}_1 - \underline{b}_1) + w'(\overline{a}_1 - \overline{b}_1)| + |w(\underline{a}_2 - \underline{b}_2) + w'(\overline{a}_2 - \overline{b}_2)| \\
& \qquad + \cdots + |w(\underline{a}_n - \underline{b}_n) + w'(\overline{a}_n - \overline{b}_n)| \\
\leq\ & \max\{|\underline{a}_1 - \underline{b}_1|, |\overline{a}_1 - \overline{b}_1|\} + \max\{|\underline{a}_2 - \underline{b}_2|, |\overline{a}_2 - \overline{b}_2|\} \\
& \qquad + \cdots + \max\{|\underline{a}_n - \underline{b}_n|, |\overline{a}_n - \overline{b}_n|\} \\
=\ & \|\mathbf{A}_1 \ominus_{gH} \mathbf{B}_1\|_{I(\mathbb{R})} + \|\mathbf{A}_2 \ominus_{gH} \mathbf{B}_2\|_{I(\mathbb{R})} + \cdots + \|\mathbf{A}_n \ominus_{gH} \mathbf{B}_n\|_{I(\mathbb{R})} \\
=\ & \|\nabla \mathbf{F}(x) \ominus_{gH} \nabla \mathbf{F}(y)\|_{I(\mathbb{R})^n}.
\end{aligned}$$

$\square$

Lemma 5.4. Let $\mathbf{F}$ be gH-differentiable and a strongly convex IVF on a nonempty convex subset $\mathcal{X}$ of $\mathbb{R}^n$. Then, there exists a $\sigma > 0$ such that

$$(\mathcal{W}(\nabla \mathbf{F}(x)) - \mathcal{W}(\nabla \mathbf{F}(y)))^T (x - y) \ \geq \ \sigma \|x - y\|^2 \ \text{ for all } \ x,\, y \in \mathcal{X}.$$

Proof. Since $\mathbf{F}$ is strong convex on $\mathcal{X}$, then there exists a convex IVF $\mathbf{G}$ on $\mathcal{X}$ such that

$$\mathbf{G}(x) \oplus \frac{1}{2}\|x\|^2 \odot [\sigma, \sigma] = \mathbf{F}(x) \ \text{ for all } \ x \in \mathcal{X}.$$

where $\sigma > 0$. Therefore,

$$\begin{aligned}
& \nabla \mathbf{G}(x) \oplus ([\sigma x_1, \sigma x_1], [\sigma x_2, \sigma x_2], \dots, [\sigma x_n, \sigma x_n])^T = \nabla \mathbf{F}(x) \\
\iff\ & \mathcal{W}(\nabla \mathbf{G}(x)) + \sigma x = \mathcal{W}(\nabla \mathbf{F}(x)), \ \text{ by Remark } 18.
\end{aligned}$$

This implies

$$\mathcal{W}(\nabla \mathbf{G}(x)) = \mathcal{W}(\nabla \mathbf{F}(x)) - \sigma x. \tag{17}$$

Further, since $\mathbf{F}$ is gH-differentiable, $\mathbf{G}$ is gH-differentiable; as $\mathbf{G}$ is also convex IVF, by Lemma 5.2 and Theorem 3.4, we have

$$(\mathcal{W}(\nabla \mathbf{G}(x)) - \mathcal{W}(\nabla \mathbf{G}(y)))^T (x - y) \ \geq \ 0 \ \text{ for all } \ x,\, y \in \mathcal{X}.$$

Then, from (17) we obtain

$$(\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y)) - \sigma(x-y))^T(x-y) \geq 0$$
$$\implies (\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y)))^T(x-y) - \sigma\|x-y\|^2 \geq 0$$
$$\implies (\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y)))^T(x-y) \geq \sigma\|x-y\|^2.$$

$\square$

Lemma 5.5. If a gH-differentiable IVF $\mathbf{F}$ on a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$ has gH-Lipschitz gradient, then for some $L > 0$ we have

$$\|\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y))\| \leq L\|x-y\| \text{ for all } x, y \in \mathcal{X}$$

Proof. Since $\mathbf{F}$ has gH-Lipschitz gradient, for some $L > 0$, we have

$$\|\nabla\mathbf{F}(x) \ominus_{gH} \nabla\mathbf{F}(y)\|_{I(\mathbb{R})} \leq L\|x-y\| \text{ for all } x, y \in \mathcal{X}.$$

By Lemma 5.3 and the last relation we obtain

$$\|\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y))\| \leq L\|x-y\| \text{ for all } x, y \in \mathcal{X}.$$

$\square$

Lemma 5.6. If $\mathbf{F}$ is a strong convex and gH-differentiable IVF on a nonempty convex subset $\mathcal{X}$ of $\mathbb{R}^n$ with gH-Lipschitz gradient, then for all $x, y \in \mathcal{X}$, there exists a $\sigma > 0$ and an $L > 0$ such that

$$(\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y))^T(y-x) \geq \frac{\sigma}{L^2}\|\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y)\|^2 \text{ for all } x, y \in \mathcal{X}.$$

Proof. Since $\mathbf{F}$ has gH-Lipschitz gradient, due to Lemma 5.5 we get

$$\|\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y))\| \leq L\|x-y\| \text{ for all } x, y \in \mathcal{X}.$$

By Lemma 5.4 and the last inequality, we obtain

$$(\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y))^T(y-x) \geq \frac{\sigma}{L^2}\|\mathcal{W}(\nabla\mathbf{F}(x)) - \mathcal{W}(\nabla\mathbf{F}(y)\|^2 \text{ for all } x, y \in \mathcal{X}.$$

$\square$

Theorem 5.5. (Linear convergence under strong convexity). *Let $\mathcal{X}$ be a nonempty convex subset of $\mathbb{R}^n$. If $\mathbf{F}$ is a strong convex and gH-differentiable IVF with gH-Lipschitz gradient on $\mathcal{X}$. Then, there exists a $\sigma > 0$ and an $L > 0$ such that if $\bar{x} \in \mathcal{X}$ be an efficient solution of the IOP (2), the mapping $H_\alpha(x) = x - \alpha\mathcal{W}(\nabla\mathbf{F}(x))$ with constant step size $\alpha \in \left[0, \frac{2\sigma}{L^2}\right]$ satisfies*

$$\|H_\alpha(x) - H_\alpha(\bar{x})\| \leq \|x - \bar{x}\| \text{ for all } x \in \mathcal{X}.$$

Proof. Let $\mathbf{F}$ be strong convex and gH-differentiable IVF with gH-Lipschitz gradient on $\mathcal{X}$.

Therefore, for all $x, y \in \mathcal{X}$, we have

$$
\begin{aligned}
& \|H_\alpha(x) - H_\alpha(y)\|^2 \\
&= \left\|\big(x - \alpha \mathcal{W}(\nabla \mathbf{F}(x))\big) - \big(y - \alpha \mathcal{W}(\nabla \mathbf{F}(y))\big)\right\|^2 \\
&= \|x - y\|^2 - 2\alpha(\mathcal{W}(\nabla \mathbf{F}(x)) - \mathcal{W}(\nabla \mathbf{F}(y))^T(x - y) + \alpha^2 \|\mathcal{W}(\nabla \mathbf{F}(x)) - \mathcal{W}(\nabla \mathbf{F}(y))\|^2 \\
&\leq \|x - y\|^2 - 2\alpha\sigma\|x - y\|^2 + \alpha^2 L^2 \|x - y\|^2 \text{ by Lemma 5.4 and Lemma 5.6} \\
&\leq \|x - y\|^2 \text{ since } \alpha \in \left[0, \frac{2\sigma}{L^2}\right].
\end{aligned}
$$

Taking $y = \bar{x}$, we have $\|H_\alpha(x) - H_\alpha(\bar{x})\| \leq \|x - \bar{x}\|$ for all $x \in \mathcal{X}$. Hence, the $\mathcal{W}$-gH-gradient efficient method converges linearly. $\qquad\square$

5.3. Numerical Examples

Here we apply the proposed algorithm of $\mathcal{W}$-gH-gradient efficient method on the IOPs of the following two examples and capture the efficient solutions of the IOPs.

Example 5.1. Consider the following IOP:

$$
\min_{x \in [-3,7]} \mathbf{F}(x) = [1, 2] \odot x^2 \oplus [-8, 0] \odot x \oplus [3, 25]. \tag{18}
$$

The gH-gradient of $\mathbf{F}$ is

$$
\nabla \mathbf{F}(x) = [2, 4] \odot x \oplus [-8, 0] \text{ for all } x \in [-3, 7].
$$

The graph of the IVF $\mathbf{F}$ is illustrated in the Figure 4 by the gray shaded region and the region of the efficient solutions is marked by bold black line on x-axis. From Figure 4 it is clear that each $x \in [0, 4]$ is an efficient solution of the IOP (18).
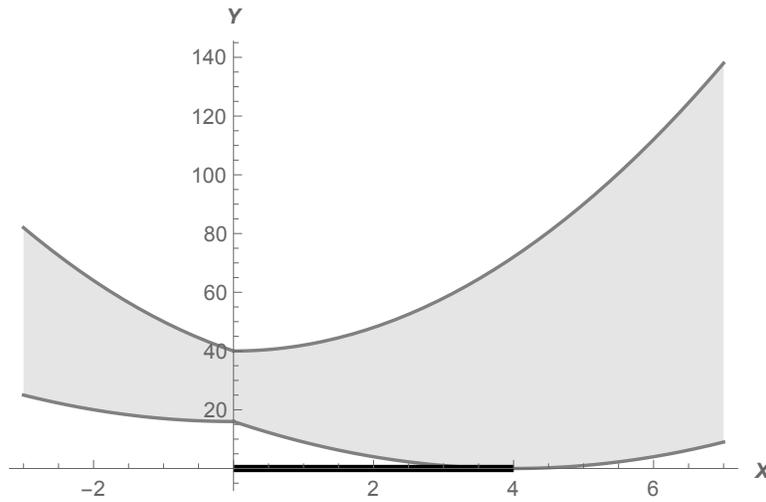

Figure 4: Interval-valued function and efficient solution of the IOP (18) in Example 5.1

Since the IOP (18) has infinite number of efficient solutions, applying $\mathcal{W}$-gH-gradient efficient method for various w and w', we obtain few of them, which are presented in Table 1.

Table 1: Output of Algorithm 2 to find efficient solutions of IOP (18)

w	w'	Initial point	Number of iterations	Efficient solution
0	1	-2	22	0
		-0.5	20	0
		6	1	3
0.2	0.8	-2	3	0.24
		-0.5	2	0.2719
		6	1	1.9778
0.4	0.6	-2	2	0.2621
		-0.5	1	0.1549
		6	1	0.6875
0.5	0.5	-2	1	0.1333
		-0.5	1	0.2424
		6	2	0.3111
0.7	0.3	-2	1	0.7891
		-0.5	1	0.4489
		6	2	0.8051
0.9	0.1	-2	1	2.25
		-0.5	1	0.7233
		6	2	3.6968

As the IOP (18) in Example 5.1 has infinite number of efficient solutions, Table 1 shows that the algorithm of $\mathcal{W}$-gH-gradient efficient method has stopped with different efficient solutions for different combinations of w, w' and initial points.

In the next example, we consider an IOP which has only one efficient solution and we show that the algorithm of $\mathcal{W}$-gH-gradient efficient method will stop at the efficient solution for different combinations of w, w' and initial points.

Example 5.2. Consider the following IOP:

$$\min_{x \in \mathcal{X} \subseteq \mathbb{R}^2} \mathbf{F}(x) = [2,6] \odot (x_1 - 2)^2 \oplus [5,7] \odot (x_2 - 3)^2 \oplus [5,12], \tag{19}$$

30

where $\mathcal{X} = [0, 6] \times [0, 6]$. The gH-gradient of $\mathbf{F}$ is

$$\nabla \mathbf{F}(x) = (D_1\mathbf{F}(x), D_2\mathbf{F}(x))^T$$
$$= ([4, 12] \odot (x_1 - 2), \; [10, 14] \odot (x_2 - 3))^T.$$

We show that $\bar{x} = (2, 3)^T$ is an efficient solution to IOP (19). On contrary, if possible, let there exist two nonzero numbers h_1 and h_2 with $0 \leq 2 + h_1 \leq 6$ and $0 \leq 3 + h_2 \leq 6$ such that

$$\mathbf{F}(2 + h_1, 3 + h_2) \prec \mathbf{F}(2, 3)$$
$$\text{or, } [2, 6] \odot h_1^2 \oplus [5, 7] \odot h_2^2 \oplus [5, 12] \prec [5, 12]$$
$$\text{or, } \left([2, 6] \odot h_1^2 \oplus [5, 7] \odot h_2^2 \oplus [5, 12]\right) \ominus_{gH} [5, 12] \prec [0, 0]$$
$$\text{or, } [2h_1^2 + 5h_2^2, 6h_1^2 + 7h_2^2] \prec [0, 0],$$

which is not possible. Thus, there does not exist any $x(\neq \bar{x}) \in \mathcal{X}$ such that $\mathbf{F}(x) \prec \mathbf{F}(\bar{x})$. Hence, $\bar{x}$ is an efficient solution of the IOP (19). Also, one can easily check that $\mathbf{F}(\bar{x}) \prec \mathbf{F}(x)$. Thus, $\bar{x}$ is the only efficient solution of the IOP (19).

The IVF $\mathbf{F}$ is depicted in the Figure 5 by the gray shaded surface and the efficient solution is pointed by black dot on xy-plane. The Figure 5 shows that $\bar{x}$ is the only efficient solution of $\mathbf{F}$ on $\mathcal{X}$.
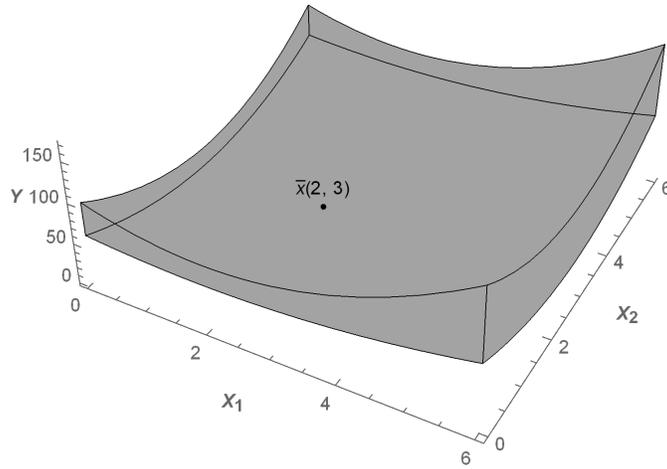

Figure 5: Interval-valued function and efficient solution of the IOP (19) in Example 5.2

From Table 2, we see that for every combination of w, w' and initial point, the $\mathcal{W}$-gH-gradient efficient method has stopped at the efficient solution $\bar{x} = (2, 3)^T$.

Table 2: Output of Algorithm 2 to find efficient solutions of IOP (19)

w	w'	Initial point	Number of iterations	Efficient solution
0.1	0.9	$(0,6)^T$	30	
		$(5,2)^T$	21	$(2,3)^T$
		$(2.5, 2.5)^T$	13	
0.3	0.7	$(0,6)^T$	16	
		$(5,2)^T$	15	$(2,3)^T$
		$(2.5, 2.5)^T$	13	
0.4	0.6	$(0,6)^T$	12	
		$(5,2)^T$	10	$(2,3)^T$
		$(2.5, 2.5)^T$	9	
0.6	0.4	$(0,6)^T$	10	
		$(5,2)^T$	9	$(2,3)^T$
		$(2.5, 2.5)^T$	11	
0.9	0.1	$(0,6)^T$	20	
		$(5,2)^T$	19	$(2,3)^T$
		$(2.5, 2.5)^T$	18	
1	0	$(0,6)^T$	25	
		$(5,2)^T$	21	$(2,3)^T$
		$(2.5, 2.5)^T$	22	

6. Application

<mark>This section applies the $\mathcal{W}$-gH-gradient efficient method in solving the *least square problems* for interval-valued data.</mark>

Suppose a set of n pairs of data $(\mathbf{X}_k, \mathbf{Y}_k)$ is given, where $\mathbf{Y}_k \in I(\mathbb{R})$ is the corresponding interval-valued output of $\mathbf{X}_k \in I(\mathbb{R})^p$ for all $k \in \{1, 2, \ldots, n\}$. We attempt to fit a function $\mathcal{H}(\cdot; \beta) : I(\mathbb{R})^p \to I(\mathbb{R})$, where $\beta \in \mathbb{R}^l$ is a parameter vector such that $\mathcal{H}\left(\mathbf{X}_k; \hat{\beta}\right)$ will be one of the best approximations of $\mathbf{Y}_k$ for all $k \in \{1, 2, \ldots, n\}$. By 'one of the best approximations' we mean that $\mathcal{H}(\mathbf{X}; \beta)$ gives a sum square error that is nondominated. Evidently, if $\hat{\beta}$ is an efficient solution of the following IOP:

$$\min_{\beta \in \mathbb{R}^l} \mathcal{E}(\beta) = \bigoplus_{k=1}^{n} \left(\mathcal{H}(\mathbf{X}_k; \beta) \ominus_{gH} \mathbf{Y}_k\right) \odot \left(\mathcal{H}(\mathbf{X}_k; \beta) \ominus_{gH} \mathbf{Y}_k\right), \quad (20)$$

then $\mathcal{H}\left(\mathbf{X}; \hat{\beta}\right)$ can be considered as an efficient choice of the approximating function $\mathcal{H}(\mathbf{X}; \beta)$.

It is noteworthy that the error function $\mathcal{E}$ and the function $\mathcal{H}(\mathbf{X}_k; \cdot)$ are IVFs from $\mathbb{R}^l$ to $I(\mathbb{R})$ for all $\mathbf{X}_k$. The partial gH-derivative of $\mathcal{E}$ with respect to β_i is

$$D_i \mathcal{E}(\beta) = 2 \odot \bigoplus_{k=1}^{n} (\mathcal{H}(\mathbf{X}_k; \beta) \ominus_{gH} \mathbf{Y}_k) \odot D_i \mathcal{H}(\mathbf{X}_k; \beta) \quad \text{for all } i \in \{1, 2, \ldots, l\}.$$

Hence, by applying the $\mathcal{W}$-gH-gradient efficient method on the IOP (20) one can easily obtain an efficient parameter vector $\hat{\beta}$ for the function $\mathcal{H}(\cdot; \beta)$. For examples, we consider the following two types of fitting with interval-valued data.

6.1. Polynomial Fitting

Let us consider a set of 21 pairs of interval-valued data that are displayed in Table 3. We attempt fit a polynomial function $\mathcal{H}^1(\cdot; \beta) : I(\mathbb{R}) \to I(\mathbb{R})$, defined by

$$\mathcal{H}^1(\mathbf{X}; \beta) = \beta_1 \odot \mathbf{C} \oplus \beta_2 \odot \mathbf{X} \oplus \beta_3 \odot \mathbf{X}^2,$$

where $\mathbf{C}$ is a constant interval and $\beta = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$. Therefore, for each $k = 1, 2, \ldots, 21$, the partial gH-derivative s of $\mathcal{H}^1(\mathbf{X}_k; \cdot)$ with respect to β_1, β_2 and β_3 are

$$D_1 \mathcal{H}^1(\mathbf{X}_k; \beta) = \mathbf{C}, \quad D_2 \mathcal{H}^1(\mathbf{X}_k; \beta) = \mathbf{X}_k \text{ and } D_3 \mathcal{H}(\mathbf{X}_k; \beta) = \mathbf{X}_k^2, \text{ respectively.}$$

Table 3: Data for polynomial fitting

$\mathbf{X}_k$	$\mathbf{Y}_k$	$\mathbf{X}_k$	$\mathbf{Y}_k$
$[-1.99, -1.86]$	$[1.21, 2.60]$	$[-1.80, -1.66]$	$[0.72, 2.00]$
$[-1.61, -1.54]$	$[0.43, 1.67]$	$[-1.48, -1.33]$	$[0.20, 1.39]$
$[-1.29, -1.14]$	$[0.01, 1.22]$	$[-1.08, -0.96]$	$[-0.28, 1.00]$
$[-0.89, -0.73]$	$[-0.60, 0.60]$	$[-0.68, -0.51]$	$[-0.8, 0.34]$
$[-0.45, -0.29]$	$[-0.95, 0.12]$	$[-0.23, -0.03]$	$[-1.12, -0.10]$
$[0.01, 0.15]$	$[-1.24, 0.00]$	$[0.19, 0.33]$	$[-1.3, 0.01]$
$[0.39, 0.54]$	$[-1.32, 0.02]$	$[0.60, 0.74]$	$[-1.25, 0.10]$
$[0.79, 0.93]$	$[-1.05, 0.13]$	$[0.98, 1.13]$	$[-0.96, 0.20]$
$[1.19, 1.33]$	$[-0.69, 0.40]$	$[1.39, 1.54]$	$[-0.34, 0.78]$
$[1.60, 1.74]$	$[-0.03, 1.09]$	$[1.79, 1.95]$	$[0.24, 1.47]$
$[2.00, 2.15]$	$[0.55, 1.80]$		

Considering $\mathbf{C} = [1.70, 12.00]$ and an initial $\beta = (6, -8, 9)^T$, and applying Algorithm 2 on the IOP (20) corresponding to the function $\mathcal{H}^1(\cdot; \beta)$ with $w = 0.3$, we obtain the value of $\hat{\beta}$ equal to $(-0.0876, -0.2974, 0.5458)^T$ in 24 iterations. With $w = 0.5$, we obtain the value of $\hat{\beta}$ equal to $(-0.0896, -0.2777, 0.5352)^T$ in 10 iterations.

In both the figures of Figure 6 show the comparison of the actual interval-valued output $\mathbf{Y}_k$ with the estimated output $\mathcal{H}^1\left(\mathbf{X}_k;\hat{\beta}\right)$ of the interval-valued data $\mathbf{X}_k$ in polynomial fitting for the values of w equal to 0.3 and 0.5, respectively, for $k = 1,\ 2,\ \ldots,\ 21$. The common portions of $\mathbf{Y}_k$ with $\mathcal{H}^1\left(\mathbf{X}_k;\hat{\beta}\right)$ are depicted by orange regions, where as the extended portions of $\mathbf{Y}_k$ and $\mathcal{H}^1\left(\mathbf{X}_k;\hat{\beta}\right)$ are illustrated by red and yellow regions, respectively.
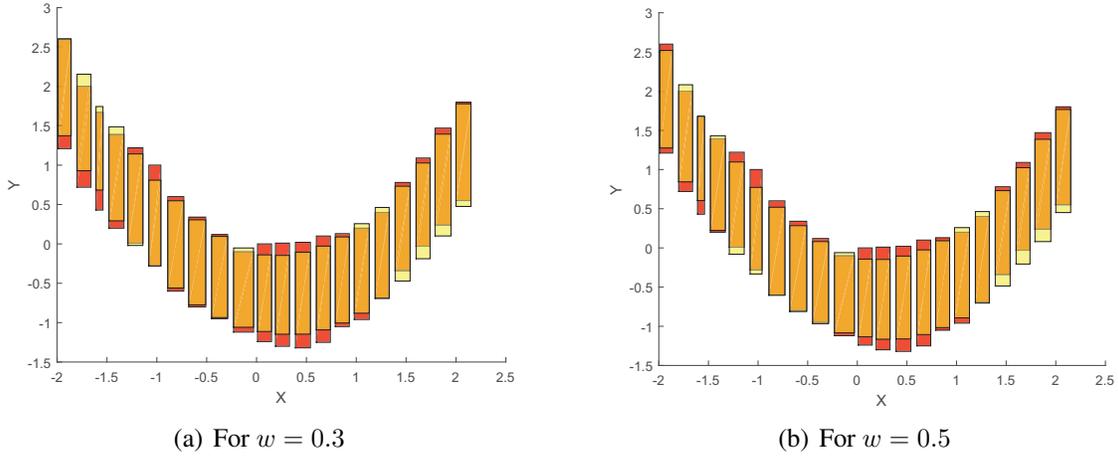

(a) For $w = 0.3$ (b) For $w = 0.5$

Figure 6: Comparison of the actual $\mathbf{Y}_k$ with the estimated output $\mathcal{H}^1\left(\mathbf{X}_k;\hat{\beta}\right)$, by the polynomial fitting, of the interval-valued data $\mathbf{X}_k$ in Table 3, $k = 1, 2, \ldots, 21$

6.2. Logistic Fitting

Let us consider a set of 15 pairs of interval-valued data that are displayed in Table 4 and we fit a logistic function $\mathcal{H}^2\left(\cdot;\beta\right) : I(\mathbb{R}) \to I(\mathbb{R})$ defined by

$$\mathcal{H}^2\left(\mathbf{X};\beta\right) = \frac{\mathbf{1}}{\mathbf{1} \oplus e^{-(\beta_1 \odot \mathbf{C} \oplus \beta_2 \odot \mathbf{X})}},$$

where $\mathbf{C}$ is a constant interval and $\beta = (\beta_1, \beta_2) \in \mathbb{R}^2$. Thus, the partial gH-derivative s of $\mathcal{H}^2\left(\mathbf{X}_k;\cdot\right)$ with respect to β_1 and β_2 are

$$D_1\mathcal{H}^2\left(\mathbf{X}_k;\beta\right) = \frac{\mathbf{1}}{\left[\mathbf{1} \oplus e^{-(\beta_1 \odot \mathbf{C} \oplus \beta_2 \odot \mathbf{X}_k)}\right]^2} \odot \mathbf{C} \text{ and } D_2\mathcal{H}\left(\mathbf{X}_k;\beta\right) = \frac{\mathbf{1}}{\left[\mathbf{1} \oplus e^{-(\beta_1 \odot \mathbf{C} \oplus \beta_2 \odot \mathbf{X}_k)}\right]^2} \odot \mathbf{X}_k,$$

respectively, for all $k = 1, 2, \ldots, 15$.

Table 4: Data for logistic fitting

$\mathbf{X}_k$	$\mathbf{Y}_k$	$\mathbf{X}_k$	$\mathbf{Y}_k$
$[-2.70, -2.55]$	$[10^{-8}, 13 \times 10^{-5}]$	$[-2.51, -2.35]$	$[6 \times 10^{-8}, 10^{-3}]$
$[-2.32, -2.23]$	$[3 \times 10^{-7}, 0.003]$	$[-2.19, -2.02]$	$[10^{-6}, 0.02]$
$[-2.00, -1.83]$	$[6 \times 10^{-6}, 0.078]$	$[-1.79, -1.65]$	$[3 \times 10^{-5}, 0.301]$
$[-1.60, -1.42]$	$[25 \times 10^{-5}, 0.760]$	$[-1.39, -1.20]$	$[15 \times 10^{-4}, 0.930]$
$[-1.16, -0.98]$	$[0.012, 0.980]$	$[-0.94, -0.72]$	$[0.080, 0.990]$
$[-0.69, -0.50]$	$[0.440, 0.998]$	$[-0.46, -0.28]$	$[0.880, 0.999]$
$[-0.24, -0.02]$	$[0.981, 1.000]$	$[0.02, 0.21]$	$[0.997, 0.999]$
$[0.26, 0.44]$	$[0.998, 1.000]$		

Taking $\mathbf{C} = [1.30, 3.40]$ and the initial value of β as $(7, -4)^T$, and applying Algorithm 2 with $w = 0.7$ on the IOP (20) corresponding to the function $\mathcal{H}^2(\cdot; \beta)$, we obtain the value of $\hat{\beta}$ as $(3.3940, 8.5835)^T$ in 10 iterations.

Figure 7 shows the comparison of the actual interval-valued output $\mathbf{Y}_k$ with the estimated output $\mathcal{H}^2\left(\mathbf{X}_k; \hat{\beta}\right)$ of the interval-valued data $\mathbf{X}_k$ in logistic fitting for the value of $w = 0.7$, $k = 1, 2, \ldots, 12$. The common portions of $\mathbf{Y}_k$ with $\mathcal{H}^2\left(\mathbf{X}_k; \hat{\beta}\right)$ are illustrated by orange regions. The extended portions of $\mathbf{Y}_k$ and $\mathcal{H}^2\left(\mathbf{X}_k; \hat{\beta}\right)$ are depicted by red and yellow regions, respectively.
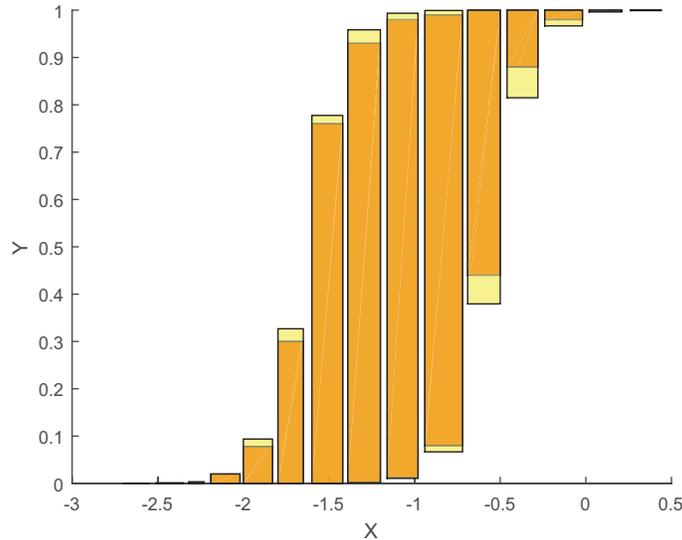

Figure 7: Comparison of the actual $\mathbf{Y}_k$ with the estimated output $\mathcal{H}^2(\mathbf{X}_k; \beta)$ of the interval-valued data $\mathbf{X}_k$ in logistic fitting for the data in Table 4, $k = 1, 2, \ldots, 15$

7. Conclusion and Future Directions

In this article, a general gH-gradient efficient-direction method and a $\mathcal{W}$-gradient efficient method for IOPs have been developed. The convergence analysis and the algorithmic implementations of both the methods have been presented. It has been shown that the $\mathcal{W}$-gradient efficient method converges linearly for a strongly convex interval-valued objective function. In the sequel, a few topics of calculus and convexity of IVFs have been proposed which were needed for the development of the methods. For a gH-differentiable IVFs, the relation between convexity and the gradient of a gH-differentiable IVF and an optimality condition of an IOP have been derived. Further, a notion of efficient-direction for IVFs has been introduced which is used to develop the general gH-gradient efficient and $\mathcal{W}$-gradient efficient methods. The proposed $\mathcal{W}$-gradient efficient method has been used to solve the least square problems with interval-valued data. The application has been exemplified by a polynomial fitting and a logistic curve fitting.

In connection with the proposed research, future research can evolve in several directions as follows.

- One may attempt to develop an appropriate approach to find the step lengths α_k's in Algorithms 1 and 2. Towards this, one can also concentrate on developing exact or inexact line search techniques for IOPs.

- It is notable that in the definition of strongly convex IVFs (Definition 2.7), we have taken a degenerate interval $[\sigma, \sigma]$ for some $\sigma > 0$. One can attempt to make a generalization of the used $[\sigma, \sigma]$ to $[\sigma_1, \sigma_2]$.

- Analysis of the proposed method for more flexible IVFs, especially for nonconvex IVFs can be performed in the future.

- One may attempt to apply the proposed logistic regression on the classification problems with interval-valued data.

- In the future, a rigorous error analysis of the red and yellow portions of Figure 6 and Figure 7 can be performed.

- Future research can be made on applying the proposed methods in constrained least square problems with interval-valued data.

- Applications of least square technique in practical problems can be performed.

- Also in the future, one can try to develop the proposed methods of this article based on constrained interval analysis [40].

Appendix A. Proof of Lemma 2.1

Proof. (i)

$$
\begin{aligned}
\mathbf{A} \preceq \mathbf{B} &\Longleftrightarrow [\underline{a}, \overline{a}] \preceq [\underline{b}, \overline{b}] \\
&\Longleftrightarrow \underline{a} \leq \underline{b} \text{ and } \overline{a} \leq \overline{b} \\
&\Longleftrightarrow \underline{a} - \underline{b} \leq 0 \text{ and } \overline{a} - \overline{b} \leq 0 \\
&\Longleftrightarrow \left[\min\left\{ \underline{a} - \underline{b}, \overline{a} - \overline{b} \right\}, \ \max\left\{ \underline{a} - \underline{b}, \overline{a} - \overline{b} \right\} \right] \preceq \mathbf{0} \\
&\Longleftrightarrow \mathbf{A} \ominus_{gH} \mathbf{B} \preceq \mathbf{0}.
\end{aligned}
$$

(ii) According to the Definition 3.2, $\mathbf{A} = \mathbf{B} \Longleftrightarrow \mathbf{A} \ominus_{gH} \mathbf{B} = \mathbf{0}$.
   If $\mathbf{A} \neq \mathbf{B}$, then

$$
\begin{aligned}
\mathbf{A} \not\preceq \mathbf{B} &\Longleftrightarrow [\underline{a}, \overline{a}] \not\preceq [\underline{b}, \overline{b}] \\
&\Longleftrightarrow \text{at least one of the inequalities } \underline{a} > \underline{b} \text{ and } \overline{a} > \overline{b} \text{ is true} \\
&\Longleftrightarrow \max\left\{ \underline{a} - \underline{b}, \overline{a} - \overline{b} \right\} > 0 \\
&\Longleftrightarrow \mathbf{A} \ominus_{gH} \mathbf{B} \not\preceq \mathbf{0}.
\end{aligned}
$$

$\square$

Appendix B. Proof of Lemma 2.3

Proof. Let $\mathbf{F}$ be gH-continuous at a point $\bar{x}$ of the set $\mathcal{X}$. Thus, for any $d \in \mathbb{R}^n$ such that $\bar{x} + d \in \mathcal{X}$,
$$
\lim_{\|d\| \to 0} \left(\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x}) \right) = \mathbf{0},
$$

which implies

$$
\lim_{\|d\| \to 0} \left([\underline{f}(\bar{x} + d), \ \overline{f}(\bar{x} + d)] \ominus_{gH} [\underline{f}(\bar{x}), \ \overline{f}(\bar{x})] \right) = [0, \ 0].
$$

Hence, by the definition of gH-difference we have

$$
\lim_{\|d\| \to 0} \left(\underline{f}(\bar{x} + d) - \underline{f}(\bar{x}) \right) \to 0 \text{ and } \lim_{\|d\| \to 0} \left(\overline{f}(\bar{x} + d) - \overline{f}(\bar{x}) \right) \to 0,
$$

i.e., $\underline{f}$ and $\overline{f}$ are continuous at $\bar{x} \in \mathcal{X}$.

   Conversely, let the functions $\underline{f}$ and $\overline{f}$ be continuous at $\bar{x} \in \mathcal{X}$. If possible, let $\mathbf{F}$ be not gH-continuous at $\bar{x}$. Then, as $\|d\| \to 0$, $\left(\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x}) \right) \not\to \mathbf{0}$. Therefore, as $\|d\| \to 0$ at least one of the functions $\left(\underline{f}(\bar{x} + d) - \underline{f}(\bar{x}) \right)$ and $\left(\overline{f}(\bar{x} + d) - \overline{f}(\bar{x}) \right)$ does not tend to 0. So it is clear that at least one of the functions $\underline{f}$ and $\overline{f}$ is not continuous at $\bar{x}$. This contradicts the assumption that the functions $\underline{f}$ and $\overline{f}$ both are continuous at $\bar{x}$. Hence, $\mathbf{F}$ is gH-continuous at $\bar{x}$.

$\square$

Appendix C. Proof of Lemma 2.4

Proof. The result is followed by Lemma 2.3 and the fact that the real-valued functions $\underline{f}$ and $\overline{f}$ are continuous at the point $\bar{x} \in \mathcal{X}$ if and only if for every sequence $\{x_n\}$ in $\mathcal{X}$ converging to $\bar{x}$, the sequences $\{\underline{f}(x_n)\}$ and $\{\overline{f}(x_n)\}$ converge to $\underline{f}(\bar{x})$ and $\overline{f}(\bar{x})$, respectively.

$\square$

Appendix D. Proof for the problem in Remark 6

Proof. Due to Definition 3.1,

$$\mathbf{F}(y) = \bigoplus_{i=1}^{n} y_i \odot \mathbf{F}(e_i) \text{ for all } y = (y_1, y_2, \ldots, y_n)^T \in \mathcal{X}.$$

Replacing y by λx, where $\lambda \in \mathbb{R}$ and $x = (x_1, x_2, \ldots, x_n)^T \in \mathcal{X}$, we have

$$\mathbf{F}(\lambda x) = \bigoplus_{i=1}^{n} \lambda x_i \odot \mathbf{F}(e_i) = \lambda \odot \bigoplus_{i=1}^{n} x_i \odot \mathbf{F}(e_i) = \lambda \odot \mathbf{F}(x)$$

Let us consider a pair of elements $x = (x_1, x_2, \ldots, x_n)^T$ and $y = (y_1, y_2, \ldots, y_n)^T$ of $\mathcal{X}$. If the corresponding x_i and y_i are of same sign for all $i \in \{1, 2, \ldots, n\}$, then

$$\begin{aligned}
\mathbf{F}(x) \oplus \mathbf{F}(y) &= \bigoplus_{i=1}^{n} x_i \odot \mathbf{F}(e_i) \oplus \bigoplus_{i=1}^{n} y_i \odot \mathbf{F}(e_i) \\
&= \bigoplus_{i=1}^{n} (x_i + y_i) \odot \mathbf{F}(e_i), \text{ by Remark 1} \\
&= \mathbf{F}(x + y).
\end{aligned}$$

Let a few of the x_i's and the corresponding y_i's are of different signs. Without loss of generality, let the first p number of x_i's are of same signs with their corresponding y_i's, thereafter consecutive $q - p$ numbers of x_i's are nonnegative but corresponding y_i's are nonpositive, and the last $n - q$ numbers of x_i's are nonpositive but corresponding y_i's are nonnegative, where $p \leq q \leq n$. Also, let

$$\mathbf{F}(e_i) = \mathbf{A}_i = [\underline{a}_i, \overline{a}_i], \text{ for all } e_i, \ i \in \{1, 2, \ldots, n\}$$

Then, we have

$$\begin{aligned}
\mathbf{F}(x) \oplus \mathbf{F}(y) = & \left(\bigoplus_{k=1}^{p} x_k \odot \mathbf{F}(e_k) \oplus \bigoplus_{l=p+1}^{q} x_l \odot \mathbf{F}(e_l) \oplus \bigoplus_{m=q+1}^{n} x_m \odot \mathbf{F}(e_m) \right) \\
& \oplus \left(\bigoplus_{k=1}^{p} y_k \odot \mathbf{F}(e_k) \oplus \bigoplus_{l=p+1}^{q} y_l \odot \mathbf{F}(e_l) \oplus \bigoplus_{m=q+1}^{n} y_m \odot \mathbf{F}(e_m) \right)
\end{aligned}$$

38

$$= \bigoplus_{k=1}^{p}(x_k + y_k) \odot \mathbf{F}(e_k) \oplus \bigoplus_{l=p+1}^{q} (x_l \odot \mathbf{F}(e_l) \oplus y_l \odot \mathbf{F}(e_l))$$

$$\oplus \bigoplus_{m=q+1}^{n} (x_m \odot \mathbf{F}(e_m) \oplus y_m \odot \mathbf{F}(e_m))$$

$$= \mathbf{B} \oplus \bigoplus_{l=p+1}^{q} (x_l \odot \mathbf{F}(e_l) \ominus v_l \odot \mathbf{F}(e_l)) \oplus \bigoplus_{m=q+1}^{n} (y_m \odot \mathbf{F}(e_m) \ominus u_m \odot \mathbf{F}(e_m)),$$

$$\text{where } \mathbf{B} = \bigoplus_{k=1}^{p}(x_k + y_k) \odot \mathbf{F}(e_k), \ v_l = -y_l \text{ and } u_m = -x_m$$

$$\text{for all } l = p+1, \ p+2, \ \ldots, \ q \text{ and } m = q+1, \ q+2, \ \ldots, \ r$$

$$= \mathbf{B} \oplus \left[\sum_{l=p+1}^{q} (\underline{a}_l x_l - \overline{a}_l v_l) + \sum_{m=q+1}^{n} (\underline{a}_m y_m - \overline{a}_m u_m), \right.$$

$$\left. \sum_{l=p+1}^{q} (\overline{a}_l x_l - \underline{a}_l v_l) + \sum_{m=q+1}^{n} (\overline{a}_m y_m - \underline{a}_m u_m) \right]$$

$$= \mathbf{B} \oplus \mathbf{C}, \quad \text{where } \underline{c} = \sum_{l=p+1}^{q} (\underline{a}_l x_l - \overline{a}_l v_l) + \sum_{m=q+1}^{n} (\underline{a}_m y_m - \overline{a}_m u_m)$$

$$\text{and } \overline{c} = \sum_{l=p+1}^{q} (\overline{a}_l x_l - \underline{a}_l v_l) + \sum_{m=q+1}^{n} (\overline{a}_m y_m - \underline{a}_m u_m).$$

Further, we note that

$$\mathbf{F}(x + y) = \bigoplus_{k=1}^{p}(x_k + y_k) \odot \mathbf{F}(e_k) \oplus \bigoplus_{l=p+1}^{q} (x_l + y_l) \odot \mathbf{F}(e_l) \oplus \bigoplus_{m=q+1}^{n} (x_m + y_m) \odot \mathbf{F}(e_m)$$

$$= \mathbf{B} \oplus \bigoplus_{l=p+1}^{q} (x_l - v_l) \odot \mathbf{F}(e_l) \oplus \bigoplus_{m=q+1}^{n} (y_m - u_m) \odot \mathbf{F}(e_m)$$

$$= \mathbf{B} \oplus \bigoplus_{l=p+1}^{q} (x_l - v_l) \odot [\underline{a}_l, \overline{a}_l] \oplus \bigoplus_{m=q+1}^{n} (y_m - u_m) \odot [\underline{a}_m, \overline{a}_m]$$

Again, without loss of generality, we let that among $q - p$ numbers of $x_l - v_l$ the first $r - p$ elements are nonnegative and rest are nonpositive, where $p \leq r \leq q$. Similarly, also we let that among $n - q$ numbers of $y_m - u_m$ the first $s - q$ elements are nonnegative and rest are nonpositive, where $q \leq s \leq n$. Then, we have $\mathbf{F}(x + y) = \mathbf{B} \oplus \mathbf{D}$, where

$$\underline{d} = \sum_{l=p+1}^{r} \underline{a}_l (x_l - v_l) + \sum_{l=r+1}^{q} \overline{a}_l (x_l - v_l) + \sum_{m=q+1}^{s} \underline{a}_l (y_m - u_m) + \sum_{m=s+1}^{n} \overline{a}_l (y_m - u_m)$$

and

$$\overline{d} = \sum_{l=p+1}^{r} \overline{a}_l \left(x_l - v_l\right) + \sum_{l=r+1}^{q} \underline{a}_l \left(x_l - v_l\right) + \sum_{m=q+1}^{s} \underline{a}_l \left(y_m - u_m\right) + \sum_{m=s+1}^{n} \overline{a}_l \left(y_m - u_m\right).$$

Since all x_l, v_l, y_m, u_m are positive for all $p+1 \leq l \leq q$ and $q+1 \leq m \leq n$, we obtain

$$\underline{a}_l x_l - \overline{a}_l x_l \leq \underline{a}_l(x_l - v_l), \ \ \underline{a}_l x_l - \overline{a}_l x_l \leq \overline{a}_l(x_l - v_l),$$
$$\underline{a}_m y_m - \overline{a}_m u_m \leq \underline{a}_m(y_m - u_m), \ \ \underline{a}_m y_m - \overline{a}_m u_m \leq \overline{a}_m(y_m - u_m),$$
$$\overline{a}_l x_l - \underline{a}_l x_l \geq \overline{a}_l(x_l - v_l), \ \ \overline{a}_l x_l - \underline{a}_l x_l \geq \underline{a}_l(x_l - v_l),$$
$$\overline{a}_m y_m - \underline{a}_m u_m \geq \overline{a}_m(y_m - u_m), \ \text{and} \ \overline{a}_m y_m - \underline{a}_m u_m \geq \underline{a}_m(y_m - u_m).$$

Thus, we get

$$\underline{c} \leq \underline{d} \ \text{ and } \ \overline{c} \geq \overline{d},$$
$$\underline{b} + \underline{c} \leq \underline{b} + \underline{d} \ \text{ and } \ \overline{b} + \overline{c} \geq \overline{b} + \overline{d}.$$

Therefore, either $\mathbf{F}(x) \oplus \mathbf{F}(y) = \mathbf{B} \oplus \mathbf{C}$ and $\mathbf{F}(x + y) = \mathbf{B} \oplus \mathbf{D}$ are equal or none of them dominates the other for all x and y in $\mathcal{X}$. □

Appendix E. Proof of Lemma 3.1

Proof. Let $\mathbf{F}$ be gH-differentiable at $\bar{x} \in \mathcal{X}$. By Definition 3.2, there exists a $\delta > 0$ such that

$$(\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \mathbf{L}_{\bar{x}}(d) = \|d\| \odot \mathbf{E}(\mathbf{F}(\bar{x}); d) \ \text{ for all } \ \|d\| < \delta, \qquad \text{(E.1)}$$

where $\mathbf{E}(\mathbf{F}(\bar{x}); d) \to \mathbf{0}$ as $\|d\| \to 0$.

Considering $d = \lambda h$ for $\lambda \neq 0$ and $h \in \mathbb{R}^n$ with $|\lambda|\|h\| < \delta$, from the equation (E.1), we obtain

$$\tfrac{1}{\lambda} \odot [(\mathbf{F}(\bar{x} + \lambda h) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \mathbf{L}_{\bar{x}}(\lambda h)] = \tfrac{|\lambda|\|d\|}{\lambda} \odot \mathbf{E}(\mathbf{F}(\bar{x}); h)$$
$$\text{or, } \tfrac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda h) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \tfrac{1}{\lambda} \odot \mathbf{L}_{\bar{x}}(\lambda h) = \tfrac{|\lambda|\|d\|}{\lambda} \odot \mathbf{E}(\mathbf{F}(\bar{x}); h)$$
$$\text{or, } \lim_{\lambda \to 0} \tfrac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda h) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \mathbf{L}_{\bar{x}}(h) = \mathbf{0}, \ \text{since} \ \mathbf{L}_{\bar{x}}(\lambda h) = \lambda \odot \mathbf{L}_{\bar{x}}(h).$$

Hence,

$$\lim_{\lambda \to 0} \tfrac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda h) \ominus_{gH} \mathbf{F}(\bar{x})) = \mathbf{L}_{\bar{x}}(h). \qquad \text{(E.2)}$$

□

Appendix F. Proof of Theorem 3.1

Proof. If $d = (0, 0, \ldots, 0)^T \in \mathbb{R}^n$, both the sides of the equation (1) become $\mathbf{0}$. Hence, the equation (1) is trivially true for $d = (0, 0, \ldots, 0)^T$.

Let us assume that $d \neq (0, 0, \ldots, 0)^T$. Since $\mathbf{F}$ is gH-differentiable at $\bar{x}$, by Lemma 3.1 there exists a $\delta > 0$ such that

$$\lim_{\lambda \to 0} \tfrac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda h) \ominus_{gH} \mathbf{F}(\bar{x})) = \mathbf{L}_{\bar{x}}(h), \tag{F.1}$$

where $d = \lambda h$ with $\lambda \neq 0$, $h \in \mathbb{R}^n$ and $|\lambda| \|h\| < \delta$.

Taking $h = e_i$, the i-th unit vector in the standard basis of $\mathbb{R}^n$, from the equation (F.1) we obtain

$$\lim_{\lambda \to 0} \tfrac{1}{\lambda} \odot (\mathbf{F}(\bar{x} + \lambda e_i) \ominus_{gH} \mathbf{F}(\bar{x})) = \mathbf{L}_{\bar{x}}(e_i)$$
$$\text{or, } D_i \mathbf{F}(\bar{x}) = \mathbf{L}_{\bar{x}}(e_i).$$

Therefore, all the i-th partial gH-derivative $D_i \mathbf{F}(\bar{x})$ of $\mathbf{F}$ at $\bar{x}$ exist. Hence, the gradient of $\mathbf{F}$ at $\bar{x}$ exists.

According to Definition 3.1 of linear IVF, we get

$$\begin{aligned}
\mathbf{L}_{\bar{x}}(d) &= \mathbf{L}_{\bar{x}}(d_1, d_2, \ldots, d_n) \\
&= d_1 \odot \mathbf{L}_{\bar{x}}(e_1) \oplus d_2 \odot \mathbf{L}_{\bar{x}}(e_2) \oplus \cdots \oplus d_n \odot \mathbf{L}_{\bar{x}}(e_n) \\
&= \bigoplus_{i=1}^{n} d_i \odot D_i \mathbf{F}(\bar{x}) = d^T \odot \nabla \mathbf{F}(\bar{x}).
\end{aligned}$$

$\square$

Appendix G. Proof of Theorem 3.2

Proof. Consider an arbitrary $\bar{x} \in \mathcal{X}$. Let for any $d \in \mathcal{N}_\delta(\bar{x}) \cap \mathcal{X}$,

$$\mathbf{F}(\bar{x} + d) = \mathbf{F}(\bar{x}) \oplus \mathbf{F}(d).$$

Since

$$\nabla \mathbf{F}(\bar{x}) = (\mathbf{F}(e_1), \mathbf{F}(e_2), \ldots, \mathbf{F}(e_n))^T,$$

for any $d \in \mathcal{N}_\delta(\bar{x}) \cap \mathcal{X}$, we obtain

$$d^T \odot \nabla \mathbf{F}(\bar{x}) = \bigoplus_{i=1}^{n} d_i \odot \mathbf{F}(e_i) = \mathbf{F}(d).$$

Therefore,

$$\begin{aligned}
& \lim_{\|d\| \to 0} \tfrac{1}{\|d\|} \odot \left((\mathbf{F}(\bar{x} + d) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} d^T \odot \nabla \mathbf{F}(\bar{x}) \right) \\
&= \lim_{\|d\| \to 0} \tfrac{1}{\|d\|} \odot \left((\mathbf{F}(\bar{x}) \oplus \mathbf{F}(d) \ominus_{gH} \mathbf{F}(\bar{x})) \ominus_{gH} \mathbf{F}(d) \right) \\
&= \lim_{\|d\| \to 0} \tfrac{1}{\|d\|} \odot \left(\left[\underline{f}(\bar{x}) + \underline{f}(d) - \underline{f}(\bar{x}), \overline{f}(\bar{x}) + \overline{f}(d) - \overline{f}(\bar{x}) \right] \ominus_{gH} \mathbf{F}(d) \right)
\end{aligned}$$

41

$$= \lim_{\|d\| \to 0} \frac{1}{\|d\|} \odot (\mathbf{F}(d) \ominus_{gH} \mathbf{F}(d))$$
$$= \mathbf{0}.$$

Hence, due to Definition 3.2 and Theorem 3.1, the linear IVF $\mathbf{F}$ is differentiable at $\bar{x} \in \mathcal{X}$. $\quad\square$

Appendix H. Proof of Theorem 3.3

Proof. Let the function $\mathbf{F}$ be convex on $\mathcal{X}$. Then, for any $x,\ y \in \mathcal{X}$ and $\lambda \in (0, 1]$, we get

$$\mathbf{F}(x + \lambda(y - x)) = \mathbf{F}(\lambda y + \lambda' x) \preceq \lambda \odot \mathbf{F}(y) \oplus \lambda' \odot \mathbf{F}(x), \text{ where } \lambda' = 1 - \lambda.$$

Hence,

$$\begin{aligned}
\mathbf{F}(x + \lambda(y - x)) \ominus_{gH} \mathbf{F}(x) &\preceq (\lambda \odot \mathbf{F}(y) \oplus \lambda' \odot \mathbf{F}(x)) \ominus_{gH} \mathbf{F}(x) \\
&= \left[\lambda \underline{f}(y) + \lambda' \underline{f}(x), \lambda \overline{f}(y) + \lambda' \overline{f}(x) \right] \ominus_{gH} \left[\underline{f}(x), \overline{f}(x) \right] \\
&= \left[\min\{ \lambda \underline{f}(y) + \lambda' \underline{f}(x) - \underline{f}(x), \lambda \overline{f}(y) + \lambda' \overline{f}(x) - \overline{f}(x) \}, \right. \\
&\qquad \left. \max\{ \lambda \underline{f}(y) + \lambda' \underline{f}(x) - \underline{f}(x), \lambda \overline{f}(y) + \lambda' \overline{f}(x) - \overline{f}(x) \} \right] \\
&= \left[\min\{ \lambda \underline{f}(y) - \lambda \underline{f}(x), \lambda \overline{f}(y) - \lambda \overline{f}(x) \}, \right. \\
&\qquad \left. \max\{ \lambda \underline{f}(y) - \lambda \underline{f}(x), \lambda \overline{f}(y) - \lambda \overline{f}(x) \} \right] \\
&= \lambda \odot \left[\min\{ \underline{f}(y) - \underline{f}(x), \overline{f}(y) - \overline{f}(x) \}, \right. \\
&\qquad \left. \max\{ \underline{f}(y) - \underline{f}(x), \overline{f}(y) - \overline{f}(x) \} \right], \text{ since } \lambda > 0 \\
&= \lambda \odot (\mathbf{F}(y) \ominus_{gH} \mathbf{F}(x)),
\end{aligned}$$

which implies

$$\frac{1}{\lambda} \odot (\mathbf{F}(x + \lambda(y - x)) \ominus_{gH} \mathbf{F}(x)) \preceq \mathbf{F}(y) \ominus_{gH} \mathbf{F}(x).$$

Since $\mathbf{F}$ is gH-differentiable at $x \in \mathcal{X}$, taking $\lambda \to 0+$, by Theorem 3.1, we have

$$(y - x)^T \odot \nabla \mathbf{F}(x) \preceq \mathbf{F}(y) \ominus_{gH} \mathbf{F}(x) \text{ for all } x,\ y \in \mathcal{X}.$$

$\quad\square$

Appendix I. Proof of Theorem 3.4

Proof. Let the function $\mathbf{F}$ be convex on $\mathcal{X}$. As $\mathbf{F}$ is also gH-differentiable on $\mathcal{X}$, by Theorem 3.3, for all $x, y \in \mathcal{X}$ we obtain

$$(y - x)^T \odot \nabla \mathbf{F}(x) \preceq \mathbf{F}(y) \ominus_{gH} \mathbf{F}(x) \tag{I.1}$$

and

$$(x - y)^T \odot \nabla \mathbf{F}(y) \preceq \mathbf{F}(x) \ominus_{gH} \mathbf{F}(y). \tag{I.2}$$

For a given pair of points $x, y \in \mathcal{X}$, let us suppose

$$\mathbf{F}(x) \ominus_{gH} \mathbf{F}(y) = [\underline{a}, \overline{a}], \ (x - y)^T \odot \nabla \mathbf{F}(x) = [\underline{b}, \overline{b}] \text{ and } (x - y)^T \odot \nabla \mathbf{F}(y) = [\underline{c}, \overline{c}].$$

Thus, from (I.1) and (I.2), respectively, we have

$$[-\overline{b}, -\underline{b}] \preceq [-\overline{a}, -\underline{a}] \text{ and } [\underline{c}, \overline{c}] \preceq [\underline{a}, \overline{a}]$$
$$\implies [\underline{a}, \overline{a}] \preceq [\underline{b}, \overline{b}] \text{ and } [-\overline{a}, -\underline{a}] \preceq [-\overline{c}, -\underline{c}]$$
$$\implies \underline{a} \leq \underline{b}, \ \overline{a} \leq \overline{b}, \ -\underline{a} \leq -\underline{c} \text{ and } -\overline{a} \leq -\overline{c}$$
$$\implies 0 \leq \underline{b} - \underline{c} \text{ and } 0 \leq \overline{b} - \overline{c}$$
$$\implies \mathbf{0} \preceq [\min\{\underline{b} - \underline{c}, \overline{b} - \overline{c}\}, \max\{\underline{b} - \underline{c}, \overline{b} - \overline{c}\}]$$
$$\implies \mathbf{0} \preceq (x - y)^T \odot \nabla \mathbf{F}(x) \ominus_{gH} (x - y)^T \odot \nabla \mathbf{F}(y).$$

Since x and y are arbitrary,

$$\mathbf{0} \preceq (x - y)^T \odot \nabla \mathbf{F}(x) \ominus_{gH} (x - y)^T \odot \nabla \mathbf{F}(y) \text{ for all } x, \ y \in \mathcal{X}.$$

$\square$

Acknowledgement

References

References

[1] Ahmad, I., Jayswal, A., Al-Homidan, S. and Banerjee, J. (2018), Sufficiency and duality in interval-valued variational programming, *Neural Computing and Applications*, 1–11.

[2] Antczak T. (2017), Optimality conditions and duality results for nonsmooth vector optimization problems with the multiple intervalvalued objective function, *Acta Mathematica Scientia*, 37B(4), 1133–1150.

[3] Bao, Y., Zao, B. and Bai, E. (2016), Directional differentiability of interval-valued functions, *Journal of Mathematics and Computer Science* 16(4), 507–515.

[4] Bazaraa, M. S., Sherali, H. D. and Shetty C. M. (2006), *Nonlinear Programming*, John Wiley and Sons, Inc. 3rd Edition.

[5] Bede, B. and Gal, S. G. (2005). Generalizations of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations, *Fuzzy Sets and Systems*, 151(3), 581–599.

[6] Bede, B. and Stefanini, L. (2013). Generalized differentiability of fuzzy-valued functions. *Fuzzy Sets and Systems*, 230(1), 119–141.

[7] Bhurjee, A. K. and Panda, G. (2012), Efficient solution of interval optimization problem, *Mathematical Methods of Operations Research* 76(3), 273–288.

[8] Bhurjee, A. K. and Padhan, S. K. (2016), Optimality conditions and duality results for nondifferentiable interval optimization problems, *Journal of Applied Mathematics and Computing*, 50(1-2) 59–71.

[9] Chalco-Cano, Y., Román-Flores, H. and Jiménez-Gamero, M. D. (2011), Generalized derivative and π-derivative for set-valued functions, *Information Sciences*, 181(11), 2177–2188.

[10] Chalco-Cano, Y., Lodwick, W. A. and Rufian-Lizana, A. (2013), Optimality conditions of type KKT for optimization problem with interval-valued objective function via generalized derivative, *Fuzzy Optimization and Decision Making*, 12, 305–322.

[11] Chalco-Cano, Y., Rufian-Lizana, A., Roman-Flores, H. and Jimenez-Gamero, M. D. (2013), Calculus for interval-valued functions using generalized Hukuhara derivative and applications, *Fuzzy Sets and Systems*, 219, 49–67.

[12] Chalco-Cano, Y., Maqui-Huamán, G. G., Silva, G. N. and Jiménez-Gamero, M. D. (2019), Algebra of generalized Hukuhara differentiable interval-valued functions: review and new properties. *Fuzzy Sets and Systems*, 375, 53–69.

[13] Chanas, S. and Kuchta, D. (1996), Multiobjective programming in optimization of interval objective functions–a generalized approach, *European Journal of Operational Research*, 94(3), 594–598.

[14] Chen, S. H., Wu, J. and Chen, Y. D. (2004), Interval optimization for uncertain structures, *Finite Elements in Analysis and Design*, 40, 1379–1398.

[15] Chen, S. H. and Wu, J. (2004). Interval optimization of dynamic response for uncertain structures with natural frequency constraints, *Engineering Structures*, 26(2), 221–232.

[16] Cheng, J., Liu, Z., Wu, Z., Tang, M. and Tan, J. (2016), Direct optimization of uncertain structures based on degree of interval constraint violation, *Computers and Structures* 164, 83–94.

[17] Chinneck, J. W. and Ramadan, K. (2000), Linear programming with interval coefficients, *Journal of the Operational Research Society*, 51(2), 209–220.

[18] Costa, T. M., Chalco-Cano, Y., Lodwick, W. A. and Silva, G. N. (2015), Generalized interval vector spaces and interval optimization, *Information Sciences*, 311, 74–85.

[19] Couso, I. and Dubois, D. (2014). Statistical reasoning with set-valued information: Ontic vs. epistemic views, *International Journal of Approximate Reasoning*, 55(7), 1502–1518.

[20] Csendes, T. (2001), New subinterval selection criteria for interval global optimization, *Journal of Global Optimization*, 19, 307–327.

[21] Effati, S. and Pakdaman, M. (2012), Solving the intervalvalued linear fractional programming problem, *American Journal of Computational Mathematics*, 2, 51–55.

[22] Ghosh, D. (2017), Newton method to obtain efficient solutions of the optimization problems with interval-valued objective functions, *Journal of Applied Mathematics and Computing*, 53, 709–731.

[23] Ghosh, D. (2017), A quasi-newton method with rank-two update to solve interval optimization problems, *International Journal of Applied and Computational Mathematics* 3(3), 1719–1738.

[24] Ghosh, D. and Chakraborty, D. (2019), *An Introduction to Analytical Fuzzy Plane Geometry*, Springer International Publishing.

[25] Ghosh, D., Ghosh, D., Bhuiya, S. K. and Patra, L. K. (2018), A saddle point characterization of efficient solutions for interval optimization problems, *Journal of Applied Mathematics and Computing*, 58(1–2), 193–217.

[26] Ghosh, D., Singh, A., Shukla, K. K. and Manchanda, K. (2019), Extended Karush-Kuhn-Tucker condition for constrained interval optimization problems and its application in support vector machines, *Information Sciences*, 504, 276–292.

[27] Ghosh, D., Chauhan, R. S., Mesiar, R. and Debnath, A. K. (2020), Generalized Hukuhara Gâteaux and Fréchet derivatives of interval-valued functions and their application in optimization with interval-valued functions, *Information Sciences*, 510, 317–340.

[28] Ghosh, D., Debnath, A. K. and Pedrycz, W. (2020), A variable and a fixed ordering of intervals and their application in optimization with interval-valued functions, *International Journal of Approximate Reasoning*, 121, 187–205.

[29] Gong, D., Sun, J. and Miao, Z. (2018), A set-based genetic algorithm for interval many-objective optimization problems, *IEEE Transactions on Evolutionary Computation*, 22(1), 47–60.

[30] Hong, F. X. and Li, D. F. (2017), Nonlinear programming method for interval-valued n-person cooperative games, *Operational Research* 17(2), 479–497.

[31] Hukuhara, M. (1967), Intégration des applications measurables dont la valeur est un compact convexe, *Funkcialaj Ekvacioj*, 10, 205–223.

[32] Ishibuchi, H. and Tanaka, H. (1990), Multiobjective programming in optimization of the interval objective function, *European Journal of Operational Research*, 48(2), 219–225.

[33] Jayswal, A., Stancu-Minasian, I. and Ahmed, I. (2011), On sufficient and duality for a class of interval-valued programming problems, *European Journal of Operational Research*, 188(1), 1–13.

[34] Jiang, C., Han, X. and Liu, G. P. (2008), A sequential nonlinear interval number programming method for uncertain structures, *Computer Methods in Applied Mechanics and Engineering* 197, 4250–4265.

[35] Jianga, C., Xiea, H. C., Zhanga, Z. G. and Hana, X. (2015), A new interval optimization method considering tolerance design, *Engineering Optimization*, 47(12), 1637–1650.

[36] Karmakar, S. and Bhunia, A. K. (2013), A new multi-section based technique for constrained optimization problems with interval-valued objective function, *Applied Mathematics and Computation*, 225, 487–502.

[37] Kumar, P. (2015), Inventory model with price-dependent demand rate and no shortages: an interval-valued linear fractional programming approach, *Operations Research and Applications: An International Journal* 2(4), 1–14.

[38] Limbourg, P. and Aponte, D. E. S. (2005), An optimization algorithm for imprecise multiobjective problem functions, *IEEE Congress on Evolutionary Computation, Edinburgh, Scotland, UK, Date 2nd–5th Sept, 2005*, 1, 459–466.

[39] Liu, S. T. and Wang, R. T. (2007), A numerical solution method to interval quadratic programming, *Applied Mathematics and Computation*, 189(2), 1274–1281.

[40] Lodwick, W. A. and Jamison, K. D. (2018), A constraint fuzzy interval analysis approach to fuzzy optimization, *Information Sciences*, 426, 38–49.

[41] Lupulescu, V. (2013), Hukuhara differentiability of interval-valued functions and interval differential equations on time scales, *Information Sciences*, 248, 50–67.

[42] Lupulescu, V. (2015), Fractional calculus for interval-valued functions, *Fuzzy Sets and Systems*, 265, 63–85.

[43] Markov, S. (1979), Calculus for interval functions of a real variable, *Computing*, 22(4), 325–337.

[44] Miettinen, K. (1999), *Nonlinear Multiobjective Optimization*, Second Edition, Kluwer Academic Publishers, Boston.

[45] Moore, R. E. (1966), *Interval Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey.

[46] Moore, R. E. (1987), Method and applications of interval analysis, *Society for Industrial and Applied Mathematics*.

[47] Osuna-Gómez, R., Chalco-Cano, Y., Hernández-Jiménez, B. and Ruiz-Garzón, G. (2015), Optimality conditions for generalized differentiable interval-valued functions, *Information Sciences*, 321, 136–146.

[48] Osuna-Gómez, R., Hernández-Jiménez, B., Chalco-Cano, Y. and Ruiz-Garzón, G. (2017), New efficiency conditions for multiobjective interval-valued programming problems, *Information Sciences*, 420, 235–248.

[49] Singh, D., Dar, B. A. and Kim, D. S. (2016), KKT optimality conditions in interval valued multiobjective programming with generalized differentiable functions, *European Journal of Operational Research*, 254(1), 29–39.

[50] Stefanini, L. (2009), A generalization of Hukuhara difference - soft methods for handling variability and imprecision, *Advances in Soft Computing*, 48, 203–210.

[51] Stefanini, L. and Bede, B. (2009), Generalized Hukuhara differentiability of interval-valued functions and interval differential equations, *Nonlinear Analysis: Theory, Methods and Applications*, 71(3–4), 1311–1328.

[52] Stefanini, L. (2010), A generalization of Hukuhara difference and division for interval and fuzzy arithmetic, *Fuzzy Sets and Systems*, 161, 1564–1584.

[53] Stefanini, L. and Arana-Jiménez, M. (2019) Karush–Kuhn–Tucker conditions for interval and fuzzy optimization in several variables under total and directional generalized differentiability, *Fuzzy Sets and Systems*, 362, 1–34.

[54] Tao, J. and Zhang, Z. (2016), Properties of interval-valued function space under the gH-difference and their application to semi-linear interval differential equations, *Advances in Difference Equations*, 45, 1–28.

[55] Wang, H. and Zhang, R. (2015), Optimality conditions and duality for arcwise connected interval optimization problems, *Opsearch*, 52, 870–883.

[56] Wolfe, M. A. (2000), Interval mathematics, algebraic equations and optimization, *Journal of Computational and Applied Mathematics*, 124, 263–280.

[57] Wu, X. Y., Huang, G. H., Liu, L. and Li, J. B. (2006), An interval nonlinear program for the planning of waste management systems with economies-of-scale effects—a case study for the region of Hamilton, Ontario, Canada, *European Journal of Operational Research*, 171(2), 349–372.

[58] Wu, H. C. (2007), The Karush-Kuhn-Tucker optimality conditions in an optimization problem with interval-valued objective function, *European Journal of Operational Research*, 176, 46–59.

[59] Wu, H. C. (2008), On interval-valued nonlinear programming problems, *Journal of Mathematical Analysis and Applications*, 338(1), 299–316.

[60] Wu, H. C. (2009), Dulity theory for optimization problems with interval-valued objetive function, *Journal of Optimization Theory and Applications*, 144(3), 615–628.

[61] Zhang, J., Liu, S., Li, L. and Feng, Q. (2014), The KKT optimality conditions in a class of generalized convex optimization problems with an interval-valued objective function, *Optimization Letters*, 8(2), 607–631.

[62] Zhang, Z., Wang, X. and Lu, J. (2018), Multi-objective immune genetic algorithm solving nonlinear intervalvalued programming, *Engineering Applications of Artificial Intelligence*, 67, 235–245.

[63] Zhao, Z., Han, X., Jiang, C. and Zhou, X. (2010), A nonlinear interval-based optimization method with local-densifying approximation technique, *Structural and Multidisciplinary Optimization*, 42(4), 559–573.

[64] Zhou, H. C. and Wang, Y. J. (2009), Optimality condition and mixed duality for interval-valued optimization, *Fuzzy Information and Engineering*, 2, 1315–1323.