Approximations for Pareto and Proper Pareto solutions and their KKT conditions

P. Kesarwani[*], P.K.Shukla [†], J. Dutta[‡] and K. Deb[§]

February 8, 2019

Abstract

In this article, we focus on approximate solution concepts in multiobjective optimization. We begin with well-known notions of approximate Pareto and weak Pareto solutions and the key result deduced about these notions is that any sequence converging to a weak Pareto minimizer satisfies an approximate Karush-Kuhn-Tucker (KKT) type necessary optimality condition. We then focus on the notion of approximate Geoffrion proper solutions with a preset bound and characterize them entirely through saddle-point type conditions for the problem with convex data. We also describe the approximate Benson proper solutions completely for a multiobjective problem having convex data through KKT type conditions.

1 Introduction

There has been a growing interest among optimizers to analyze the nature of an approximate solution to an optimization problem. This stems from the fact that the most optimization problems cannot be solved exactly. This fact about optimization has been very clearly stated at the beginning of the monograph on convex optimization by Nesterov [29]. In practice, we know that our algorithms only return us some approximate local/global solutions. Thus it is natural to develop a formal theory of approximate solutions and see to what extent their behavior parallels that of an exact solution. Optimizers have attempted to answer these questions over the past several decades, see for example [12], [18], [19] and reference therein.

In vector optimization the key to find a Pareto minimizer or a weak Pareto minimizer largely hinges on scalarization techniques (see Ehrgott [14], Jahn [23], Luc [27], Chankong

[*]Department of Mathematics and Statistics, Indian Institute Of Technology Kanpur,India
[†]Institute AIFB, Karlsruhe Institute of Technology, Karlsruhe, Germany
[‡]Department of Economics Sciences, Indian Institute Of Technology Kanpur, India
[§]College of Engineering, Michigan State University, Michigan, USA

et al. [4]). The solutions that we generate using these techniques, is an approximate solutions. Hence it is natural to study the notion of approximate solutions in vector optimization. Several researchers have already analyzed various notions of the approximate solutions, for instance see Dutta et al. [12], Gutierrez et al. [18, 19], White [37], Rong et al. [32].

It is also essential to a decision maker, who is taking some decisions based on multiobjective optimization models need not necessarily be interested in all the Pareto solutions of the problem at hand. In many cases, the decision maker focuses on the part of the Pareto frontier in the image space which corresponds to a subset of the set of Pareto solutions. These subsets, when chosen in a particular way, gives rise to various classes of proper Pareto solutions (see for example Ehrgott [14]).

In real practice and more specifically in engineering, there is a growing importance of the use of evolutionary algorithms for solving a multiobjective optimization problem (see for example Deb [6]). These evolutionary algorithms which are population-based heuristics, produce an approximation of the Pareto frontier very quickly in the image space. Thus the proper Pareto solutions that one gets in practice are kind of approximation of the proper Pareto solutions. Thus we are again faced with a situation where we need to build up an analysis for approximated proper Pareto solutions. When we consider the natural cone or the non-negative cone for ordering the image space, then the first notion of a proper Pareto optimal solution was put forward by Geoffrion [17]. In [17], Geoffrion defined those Pareto optimal points as proper for which the objective function has a bounded trade-offs. The importance of this notion was clear, when Geoffrion tide this to the weighted-sum scalarization. If we take all the weights positive, then solving the weighted-sum scalar problem will lead to a proper Pareto solution in the sense of Geoffrion. The converse amazingly holds if all objective functions are convex. Such proper solutions are now referred to as the Geoffrion proper Pareto solutions (see [17]).

Around thirteen years ago, Shukla et al. [33] developed a more robust version of the Geoffrion proper Pareto solutions. The notion due to Geoffrion suffers from a drawback that the decision maker does not know beforehand, is how much is the trade-off bound. A decision maker in many situations might like to be concerned with only those proper solutions whose trade-off is bounded by a preset number provided by the decision maker herself/himself. In [34], such solutions are shown to be stable when approximate solution considered. To be more precise a sequence of approximated Geoffrion proper Pareto solutions with a preset trade-off bound converges to a Geoffrion solution with the same trade-off bound. This property is not present if we consider the usual Geoffrion proper solutions. For more details on this new solution, concept see [34].

An example in Eichfelder [15] shows that in many situations the ordering cone that we need to use is not the positive cone. How does one then talk of a proper Pareto solution? Thus the idea of Geoffrion had to be generalized into a new setting. This was achieved for example in Benson [2], Borwein [3] and Henig [21] etc. In this article, we also study proper Pareto solutions in sense of Benson [2].

1.1 Our Aim

The title of the article reflects the fact that the celebrated Karush-Kuhn-Tucker condition (KKT for short) will play an integral role in this paper. There is no doubt that the KKT conditions play a pivotal role in scalar optimization. However, its role in vector optimization is not clearly understood except for the fact that it provides a necessary condition. If we look at the literature of multiobjective optimization, we do not see any direct use of the KKT condition even in algorithms for solving linear multiobjective optimization problems.

On the other hand, as we have mentioned earlier, evolutionary algorithms are gaining prominence. Since these are heuristics, it is essential to check the quality of the points that we finally select as an approximation to Pareto points. To be more robust in our approach, we can devise an approximate version of the KKT conditions which can be used as a stopping criteria. We can accept a point generated by a heuristic if it violates the KKT conditions within a given margin of error. To the best of our knowledge such a use of the KKT condition for multiobjective optimization was first carried out in [8] (see also [7]). This motivated us to rethink, how to develop approximate versions of KKT condition for approximate Geoffrion proper solutions with a preset trade-off bound. We show that such conditions are both necessary and sufficient when the problem data is convex. Thus, at least in the convex case, any point satisfying our approximate KKT conditions will be an approximate Geoffrion proper Pareto optimal point with a preset bound. This is a novel feature, since usually the converse does not hold so easily even in the convex case for approximate weak Pareto solutions see for example Dutta et al. [12]. Further, we develop a saddle point criteria for such type of solution points. We of course not only focus on Geoffrion solutions but also look at the approximate version of Benson proper Pareto solutions for which the necessary approximate KKT conditions are also sufficient when the problem data is cone-convex. Thus, both for the Geoffrion and Benson cases, we have a complete characterization when the problem data is convex with some constraint qualifications.

To make our presentation more complete, we also study some approximate KKT conditions for approximate Pareto and weak Pareto points. The study of approximate KKT conditions for vector optimization problem done in a general setting was carried out for example in Durea et al. [10]. We use their approach but provide results which are very different from that in [10]. An exciting result which we present here is a kind of converse. We show that if the problem data is locally Lipschitz or convex, then if there is a sequence of points which are feasible and converges to a weak Pareto minimum, then there exists a subsequence of that sequence whose elements satisfy an approximate KKT type condition. We are yet to explore the full implications of this result.

1.2 Organization of the Paper

There are five major sections of this article which include the introduction. In section 2, we introduce the standard notations and present the various solution concepts used in this paper. We also introduce the notion of a Geoffrion proper Pareto solution with a fixed upper bound to the trade-offs. This new solution concept was introduced to the best of our knowledge by Shukla et al. [33] and, then was freshly looked into in Shukla et al. [34]. Shukla et al. [34] explored the important properties of the solutions mentioned above.

In section 3, we focus on the usual notion of approximate weak Pareto solutions in the light of approximate KKT conditions. A key notion of approximate KKT condition introduced in this section is modified ϵ-KKT condition for multiobjective problem with locally Lipschitz data. This was motivated by its scalar counterpart introduced in Dutta et al. [11]. The key result of this section is a result of the converse type. It says that if we consider a multiobjective programming problem with locally Lipschitz data and if there is a sequence converging to a local weak Pareto minimizer, then there is a subsequence of the sequence which is feasible and satisfies a modified ϵ-KKT type condition under mild regularity assumption. This result is possible by the application of the Ekeland Variational principle for the vector case which is due to Tammer [35].

The key focus in section 4 is to develop an approximate version of the KKT condition for the approximate version of the Geoffrion proper Pareto solutions with a fixed upper bound on the trade-off which is necessary for general but is also sufficient if in particular, the problem data is convex.

In section 5, we present approximate KKT type conditions associated with the approximate version of some other class of proper Pareto solutions, known as the Benson proper solutions in which the ordering cone is not the non-negative cone itself but a more general cone is considered.

2 Notations and Definitions

Our notations are fairly standard. Let $A \subseteq \mathbb{R}^n$ be a given set, then closure and interior of set A is denoted by cl$A$ and int$A$ respectively. For vectors $x, y \in \mathbb{R}^n$ the inner product given by $\langle x, y \rangle$. A set $A \subset \mathbb{R}^n$ is a cone, if for each $a \in A$ and positive scalar λ, $\lambda a \in A$. A cone A is pointed, if $A \cap (-A) = \{0\}$. We shall write cone(A) to denote the cone generated by the set A which is given as $\text{cone}(A) = \{\lambda a \mid \lambda \geq 0,\ a \in A\}$. A dual cone A^* of set $A \subset \mathbb{R}^n$ is a cone $A^* = \{y \in \mathbb{R}^n \mid \langle x, y \rangle \geq 0,\ \text{for all } x \in A\}$ and $(A^*)^0$ is defined by $(A^*)^0 = \{y \in \mathbb{R}^n \mid \langle x, y \rangle > 0,\ \text{for all } x \in A \setminus \{0\}\}$. A normal cone of a set A at the point x_0, denoted by $N_A(x_0)$, is $N_A(x_0) = \{v \in \mathbb{R}^n : \langle v, x - x_0 \rangle \leq 0,\ \text{for all } x \in U\}$. We consider the following form of multiobjective optimization problem (MOP) in this article:

$$\min f(x) := (f_1(x), \ldots, f_m(x)),$$
$$\text{subject to } \; g_j(x) \leq 0, \; j = 1, 2, .., l.$$

where each $f_i : \mathbb{R}^n \to \mathbb{R}$ and $g_j : \mathbb{R}^n \to \mathbb{R}$. Let us denote the constraint set by $X := \{x \in \mathbb{R}^n : g_j(x) \leq 0, \; j = 1, 2, .., l\} \subseteq \mathbb{R}^n$, $I := \{1, 2, .., m\}$, $L := \{1, 2, .., l\}$ and set of active indices at x as $R(x) := \{r \in L : g_r(x) = 0\}$.
Solving MOP requires a (binary) ordering relation on $\mathbb{R}^m$. Given an ordering relation induced by a cone, one can compare two m-dimensional vectors from the set $f(X) := \{f(x)|x \in X\} \subset \mathbb{R}^m$ and define an optimality notion for MOP. Consider $C \subseteq \mathbb{R}^m$ to be a closed, convex and pointed cone and define a ordering relation " $\geq_C$ " on $\mathbb{R}^m$. For $x, y \in \mathbb{R}^m$, we say $x \geq_C y$ if and only if $x - y \in C$. By choosing different cone C, we can get different notions of optimality for MOP. These optimality notions is used in algorithms to find one or many optimal solutions of MOP.
Whenever, we solve a MOP through an algorithm irrespective of point-by-point or population-based algorithm, we always generate a sequence or set of approximate solutions. These approximate solutions go towards optimal solutions as we increase the number of iterations and we get a good approximation of solution points as we terminate our algorithm after a large number of iterations. Hence the algorithms give rise to approximate solutions or *ε-solutions*, depending on the distance of iterative points to the set of solutions. There are several notions for approximate optimality in the multi-objective theory for example [24], [26], [37], [36], [20] and the references therein. In this article, we consider the notion of approximate solution introduced in Loridan [26].
To formalize our notions, we consider $\epsilon \in \mathbb{R}^m_+$, *i.e.*, $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_m)$, $\epsilon_i \geq 0$ for each $i \in I$. Our focus on this paper is on ϵ- solutions of MOP. In the following definition the order of image space $f(X)$ is induced by natural cone $C = \mathbb{R}^m_+$.

Definition 2.1 *Given $\epsilon \in \mathbb{R}^n_+$, if there is no $x \in X$ such that $f(x) + \epsilon - f(x^*) \in -\mathbb{R}^m_+ \setminus \{0\}$, then the point $x^* \in X$ is said to be an ϵ-Pareto optimal solution of MOP. Further if there is no $x \in X$ such that $f(x) + \epsilon - f(x^*) \in -\mathrm{int}(\mathbb{R}^m_+)$, then the point x^* is said to be a weak ϵ-Pareto optimal solution of MOP.*

Let us make it clear at this stage that all our notations in this paper, specifically those denoting solution sets are borrowed from our paper [34]. The current paper is in many ways a continuation of our study in [34]. Though not always seen in the literature the following notions of a local solutions are also relevant.

Definition 2.2 *A point x^* is said to be a loacl Pareto optimal solution of MOP if there exists $\delta > 0$ and no $x \in X \cap B_\delta(x^*)$ such that, $f(x) - f(x^*) \in -\mathbb{R}^m_+ \setminus \{0\}$, where $B_\delta(x_0) \subset \mathbb{R}^n$ is a ball of radius δ.*

The weak counter part of local solution can be defined in the similar fashion as in Definition 2.1. We want to mention that in several situations we would have to consider the particular form of the vector $\epsilon \in \mathbb{R}^m_+$, given by $\epsilon = \varepsilon e$, where $e = (1, 1.., 1)^T$ and $\varepsilon \in \mathbb{R}_+$. In those cases, the solutions referred to as the εe-Pareto and εe-weak Pareto solution respectively.
The set of all ϵ-Pareto points is denoted by $S_\epsilon(f, X)$ and the set of all ϵ-weak Pareto points as $S_{w,\epsilon}(f, X)$. An ϵ-Pareto (weak) optimal solution with $\epsilon = 0$ is commonly known as Pareto (weak) optimal solution. For simplicity, the set of Pareto optimal solutions and the set of weak Pareto optimal solutions will be denoted by $\mathcal{S}(f, X)$ and $\mathcal{S}_w(f, X)$, respectively.
As argued at the beginning of this paper, different Pareto optimal solutions might have different properties that a decision maker may desire. The need to filter out bad Pareto optimal solutions lead to the notion of a properly efficient point. There are different notions of proper optimality (see a nice survey in [28]). For example, if a closed, convex and pointed cone C is used as the ordering cone, Benson proper optimality [2] is a widely-studied notion. In an earlier work [33], we introduced the following approximate version of Benson proper optimality.

Definition 2.3 *Given $\epsilon \in \mathbb{R}^n_+$, a point $x_0 \in X$ is called ϵ-Benson proper solutions of MOP (with respect to the ordering cone C) if*

$$\mathrm{cl}(\mathrm{cone}(f(X) + (C + \epsilon) - f(x_0))) \cap (-C) = \{0\}.$$

Let us denote the problem MOP as (f, X, C) when ordering of image space is induced by a cone C. The set of all ϵ-Benson proper solutions will be denoted by $\mathcal{S}_{\mathcal{B},\varepsilon}(f, X, C)$. If $\epsilon = 0$, this reduces to classical notion of Benson proper efficiency [2], and we will use the notation $\mathcal{S}_{\mathcal{B}}(f, X, C)$ instead of $\mathcal{S}_{\mathcal{B},0}(f, X, C)$.
In the case of $C = \mathbb{R}^m_+$, ϵ-Benson proper optimality is equivalent to the following notion of Geoffrion ϵ-proper optimality (see [25, 14], for example).

Definition 2.4 *Given $\epsilon \in \mathbb{R}^n_+$, a point $x_0 \in X$ is called ϵ-Geoffrion proper solution of MOP if $x_0 \in \mathcal{S}_\epsilon(f, X)$ and if there exists a number $M > 0$ such that for all $i \in I$ and $x \in X$ satisfying $f_i(x) < f_i(x_0) - \epsilon_i$, there exists an index $j \in I$ such that $f_j(x_0) - \epsilon_j < f_j(x)$ and*

$$\frac{f_i(x_0) - f_i(x) - \epsilon_i}{f_j(x) - f_j(x_0) + \epsilon_j} \leq M.$$

Geoffrion ϵ-proper solution says that the trade-offs between objective functions at two points are bounded. It is interesting to ask whether the trade-offs are bounded above for all the points. This would mean that the same M can work for all the Geoffrion ϵ-proper solutions. This may not always be the case unless the set of all trade-offs bounds is bounded above. Now we state a practical version of above definition which has been introduced in Shukla et al [33] .

Definition 2.5 *Given* $\epsilon \in \mathbb{R}^n_+$ *and a scalar* $\hat{M} > 0$*, a point* $x_0 \in X$ *is called* $(\hat{M}, \epsilon)$*-Geoffrion proper solution of MOP if* $x_0 \in \mathcal{S}^\epsilon(f, X)$ *and for all* $i \in I$ *and* $x \in X$ *satisfying* $f_i(x) < f_i(x_0) - \epsilon_i$*, there exists an index* $j \in I$ *such that* $f_j(x_0) - \epsilon_j < f_j(x)$ *and*

$$\frac{f_i(x_0) - f_i(x) - \epsilon_i}{f_j(x) - f_j(x_0) + \epsilon_j} \leq \hat{M}.$$

Given $\hat{M} > 0$, we shall denote the set of all $(\hat{M}, \epsilon)$- Geoffrion proper and ϵ-Geoffrion proper solution as $\mathcal{G}_{\hat{M},\epsilon}(f, X)$ and $\mathcal{G}_\epsilon(f, X)$ respectively. For $\epsilon = 0$, the set of exact $\hat{M}$- Geoffrion proper and Geoffrion proper solution is denoted by $\mathcal{G}_{\hat{M}}(f, X)$ and $\mathcal{G}(f, X)$ respectively.
Now we will state the most celebrated result Ekeland variation principle [16] which will play a key role in coming sections. First, we state the principle for scalar valued function and then for vector-valued function.

Theorem 2.6 *Let* $F : \mathbb{R}^n \to \mathbb{R}$ *be a lower semicontinuous function which is bounded below. Let* $\rho > 0$ *and a point* $u \in \mathbb{R}^n$ *be given such that*

$$F(u) \leq \inf_{x \in \mathbb{R}^n} F + \rho.$$

Then, there exists $u_\rho \in \mathbb{R}^n$ *such that* $\|u - u_\rho\| \leq \sqrt{\rho}$ *and*

1. $F(u_\rho) \leq F(u)$,
2. $F(u_\rho) < F(w) + \sqrt{\rho}\|w - u_\rho\|$*, for all* $w \neq u_\rho$.

To describe the above principle for vector-valued functions which was introduced in [35], we need to understand the notion of lower semicontinuity and boundness of vector-valued functions.

Definition 2.7 *Let* $C \subseteq \mathbb{R}^m$ *be a closed, convex and pointed cone with non empty interior and* $f : U \to \mathbb{R}^m$ *where* U *is a non empty subset of* $\mathbb{R}^n$*. The function* f *is* C*-bounded below if there exists* $y \in \mathbb{R}^m$ *such that* $f(x) \geq_C y$ *for all* $x \in U$*. Let* $c \in int(C)$*, the function* f *is* (c, C)*- lower semi continuous if for all* $t \in \mathbb{R}$*,* $\{x \in U : tc \geq_C f(x)\}$ *is closed.*

Theorem 2.8 *Let* C *and* U *be as defined in Definition 2.7 and let* $f : U \to \mathbb{R}^m$ *be a* (c_0, C)*- lower semi continuous function which is also* C*-bounded below. Further, suppose we are given* $\rho > 0$ *and a point* $x_0 \in U$ *such that,*

$$f(x) + \rho c_0 - f(x_0) \notin -C \setminus \{0\}, \text{ for all } x \in U. \tag{2.1}$$

Then, there exists $\bar{x_0} = \bar{x_0}(\rho) \in U$ *such that* $\|\bar{x_0} - x_0\| \leq \sqrt{\rho}$ *and for all* $x \in U \setminus \{\bar{x_0}\}$

1. $f(x) + \rho c_0 - f(\bar{x_0}) \notin -int(C)$,
2. $f(x) + \sqrt{\rho}\|\bar{x_0} - x\|c_0 - f(\bar{x_0}) \notin -int(C)$.

2.1 Tools from non-smooth analysis

In this article, we rely on two major tools from non-smooth analysis, namely the subdifferential of a convex function and the Clarke subdifferential of a locally Lipschitz function. Though these notions are very well known in the optimization community, we shall provide the definitions for completeness. We shall however restrict ourselves to the class of functions which are finite-valued function on $\mathbb{R}^n$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, then the subdifferential of f at the point x is a set of vectors in $\mathbb{R}^n$, given as

$$\partial f(x) = \{v \in \mathbb{R}^n : f(y) - f(x) \geq \langle v, y - x \rangle, \text{ for all } y \in \mathbb{R}^n\}.$$

The subdifferential set is a non-empty, convex and compact for every $x \in \mathbb{R}^n$. The subdifferential is also deeply linked with the notion of the directional derivative of a convex function. The directional derivative of a convex function at a given x in the direction h is given as

$$f'(x, h) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda h) - f(x)}{\lambda}$$

This directional derivative exists for each x and in each direction h, and, the subdifferential of f can be written as $\partial f(x) = \{v \in \mathbb{R}^n : \ f'(x, h) \geq \langle v, h \rangle, \text{ for all } h \in \mathbb{R}^n\}$.

Thus each of these can be recovered from the other. Now the most common question to ask is whether the generalized notion of derivative has properties like the usual derivative of calculus? We will begin with the most fundamental one, the sum rule. Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are convex functions. Then

$$\partial(f + g)(x) = \partial f(x) + \partial g(x). \tag{2.2}$$

For more details on subdifferentail of convex functions see [1]. It is important to note that a point x_0 is a global minimum of f on $\mathbb{R}^n$ if and only if $0 \in \partial f(x_0)$. Since subdifferential is a generalized version of derivative, it has some limitation. The ε-subdifferential is a relaxed version of the subdifferential which is very useful tool in convex analysis and optimization. We begin with defining the ε-subdifferential of convex function.

Definition 2.9 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex function and $\varepsilon \geq 0$. The ε-subdifferential of f at the point x is given as*

$$\partial_\varepsilon f(x) = \{v \in \mathbb{R}^n : f(y) - f(x) \geq \langle v, y - x \rangle - \varepsilon, \textit{ for all } y \in \mathbb{R}^n\}.$$

The elements of $\partial_\varepsilon f(x)$ are called ε-gradients of f at x and $\partial_\varepsilon f(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$. A point x_0 is called an ε-minimizer of f on $\mathbb{R}^n$ if $f(y) - f(x) \geq -\varepsilon$, for all $y \in \mathbb{R}^n$. Thus x_0 is an ε-minimizer of f on $\mathbb{R}^n$ if and only if $0 \in \partial_\varepsilon f(x_0)$. For complete description of properties of ε-subdifferential see [9].

The subdifferential defined above is only defined for convex functions, so the obvious question is to ask what about subdifferential of non-convex functions? We now discuss subdifferential of a non-convex function which is locally Lipschitz in nature. The relation of subdifferential and directional derivative as above becomes a key to develop the notion of a subdifferential for a locally Lipschitz functions.
A function $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz around $x \in \mathbb{R}^n$, if there exists a neighborhood U_x of x and $L_x \geq 0$ such that $\|f(y) - f(z)\| \leq L_x\|y - z\|$, for all $y, z \in U_x$. The constant L_x is the Lipschtiz constant of the function f at the point x. A function f is said to be locally Lipschitz if f is Lipschitz around x for any $x \in \mathbb{R}^n$.We shall focus in this article on MOP with locally Lipschitz objective and constraint functions. Sometimes, we shall refer to the problem MOP with convex data as CMOP. Both MOP and CMOP represent the problem, when the ordering cone is $\mathbb{R}^m_+$. When the ordering cone is not $\mathbb{R}^l_+$, but some closed convex cone C, then we shall refer to the MOP as (f, X, C).
We now define the Clarke directional derivative of locally Lipschitz function f at x and in the direction $h \in \mathbb{R}^n$ as

$$f^\circ(x, h) = \limsup_{y \to x, t \downarrow 0} \frac{f(y + th) - f(y)}{t}.$$

The Clarke subdifferential of f at $x \in \mathbb{R}^n$ is given as,

$$\partial^\circ f(x) = \{\xi \in \mathbb{R}^n : \ f^\circ(x, h) \geq \langle \xi, h \rangle, \text{ for all } h \in \mathbb{R}^n\}.$$

For each $x \in \mathbb{R}^n$, the set $\partial^\circ f(x)$ is non-empty, convex and compact. It is important to note that when function f is convex, then $\partial^\circ f(x) = \partial f(x)$, for all $x \in \mathbb{R}^n$. Same as subdifferential for convex function, Clarke subdifferential has lots of nice properties. If $x_0 \in \mathbb{R}^n$ is a local minimum of f over $\mathbb{R}^n$, then $0 \in \partial^\circ f(x_0)$ (for proof see [30]). It also satisfy sum rule but it gives only one side containment, *i.e.,* for given two locally Lipschtiz function f and g, we have $\partial^\circ(f + g)(x) \subset \partial^\circ f(x) + \partial^\circ g(x)$.

3 Approximate KKT conditions

We already explained the importance of approximate solutions of a multiobjective optimization problem from a practical point of view in the previous section. Thus it is natural to ask whether these approximate solutions satisfy some kind of approximate KKT condition. In the literature, there are several approaches to approximate KKT conditions [10],[11] etc.
In this section, we begin by defining a notion of approximate KKT points which suits very well for the purpose of convex vector optimization problem. This notion is called modified ϵ-KKT points, which are motivated by a similar notion defined in Dutta et al [11] for scalar optimization problem and Durea et al [10] for vector optimization. The

next question which is a natural one is as follows, *Under what assumptions on the problem data, do an ϵ-weak Pareto point satisfy the approximate KKT conditions and under what assumptions the converse can also be generated*? We shall focus on these questions in the current section.

Definition 3.1 *A feasible point $x_0 \in X$ is said to be a modified ε-KKT point of MOP if for a given $\varepsilon \in \mathbb{R}_+$, there exists x_ε such that $\|x_0 - x_\varepsilon\| \leq \sqrt{\varepsilon}$ and there exists $u_i \in \partial^\circ f_i(x_\varepsilon)$ for all $i \in I$, $v_r \in \partial^\circ g_r(x_\varepsilon)$ for all $r \in L$, vectors $\lambda \in \mathbb{R}^m_+$ with $\|\lambda\| = 1$ and $\mu \in \mathbb{R}^l_+$ such that*

$$\left\| \sum_{i\in 1}^{m} \lambda_i u_i + \sum_{r=1}^{l} \mu_r v_r \right\| \leq \sqrt{\varepsilon}, \ \textit{and}, \ \sum_{r=1}^{r} \mu_r g_r(x_0) \geq -\varepsilon.$$

We now state two constraint qualifications, Slater constraint qualification (SCQ for short) and Basic constraint qualification (BCQ for short) which are used in the main results of this article (see [31]).

Definition 3.2 *The problem MOP satisfies Slater constraint qualification if there exists $\hat{x} \in X$ such that $g_r(\hat{x}) < 0$, for all $r \in L$.*

Definition 3.3 *The problem MOP satisfies Basic constraint qualification at a point $\bar{x}$ if there exists no $p \in \mathbb{R}^l_+ \setminus \{0\}$ such that $0 \in \sum\limits_{r\in L} p_r \partial^\circ g_r(\bar{x})$.*

Theorem 3.4 *Consider the problem MOP with locally Lipschitz data and let $\{\varepsilon_k\}$ to be a decreasing sequence of positive real numbers such that $\varepsilon_k \to 0$ as $k \to \infty$. Consider $\{x^k\}$ to be a sequence of feasible points of MOP with $x^k \to x_0$ as $k \to \infty$. Assume that for each k, x^k is a modified ε_k-KKT point of MOP. Further assume that the BCQ holds at x_0. Then x_0 is a KKT point of MOP.*

Proof: As x^k is a modified ε_k- KKT point, for each k, Definition 3.1 gives the existence of a point $\hat{x}^k$ such that $\|x^k - \hat{x}^k\| \leq \sqrt{\varepsilon_k}$, the existence of $u_i^k \in \partial^\circ f_i(\hat{x}^k)$ and $v_r^k \in \partial^\circ g_r(\hat{x}^k)$ for all $i \in I$ and $r \in L$, and the vectors $\lambda^k \in \mathbb{R}^m_+$ and $\mu^k \in \mathbb{R}^l_+$ with $\|\lambda^k\| = 1$ such that

$$\left\| \sum_{i\in I} \lambda_i^k u_i^k + \sum_{r\in L} \mu_r^k v_r^k \right\| \leq \sqrt{\varepsilon_k}, \text{ and} \tag{3.1}$$

$$\sum_{r\in L} \mu_r^k g_r(x^k) \geq -\varepsilon_k. \tag{3.2}$$

We first claim that $\{\mu^k\}$ is bounded. To prove our claim, on the contrary assume that $\{\mu^k\}$ is unbounded. Thus, $\|\mu^k\| \to \infty$ as $k \to \infty$. Further, Equation (3.1), can be re-written as

$$\left\| \sum_{i\in I} \frac{\lambda_i^k}{\|\mu^k\|} u_i^k + \sum_{r\in L} \frac{\mu_r^k}{\|\mu^k\|} v_r^k \right\| \leq \frac{1}{\|\mu^k\|} \sqrt{\varepsilon_k}. \tag{3.3}$$

Then, in Equation (3.3), we observe the following:

1. As ε_k converges to 0, the same holds for $\frac{1}{\|\mu^k\|}\sqrt{\varepsilon_k}$.

2. Let $p_r^k = \frac{\mu_r^k}{\|\mu^k\|} \in \mathbb{R}_+^l$, for all $r \in L$. As $\|p^k\| = 1$ $\{p^k\}$ is a bounded sequence. So, by the Bolzano-Weierstrass theorem there exists a subsequence of $\{p^k\}$ which converges to $\hat{p}_r \in \mathbb{R}_+^l$ with $\|\hat{p}\| = 1$. In fact, without loss of generality, we can assume that p_r^k converges to $\hat{p}_r$. Hence, for all $r \in L$

$$\frac{\mu_r^k}{\|\mu^k\|} = p_r^k \to \hat{p}_r, \ as \ k \to \infty. \tag{3.4}$$

3. As f_i's are locally Lipschitz functions, their Clarke subdifferential are locally bounded. Hence, the sequence $\{u_i^k\}$, where $u_i^k \in \partial^\circ f_i(\hat{x}^k)$ for each k, is bounded for all $i \in I$. Hence, using the fact that $\|\lambda^k\| = 1$ and $\|\mu^k\| \to \infty$, we deduce that for all $i \in I$,

$$\frac{\lambda_i^k}{\|\mu^k\|} u_i^k \to 0, \ as \ k \to \infty. \tag{3.5}$$

4. An argument similar to the previous part implies that the sequence $\{v_r^k\}$, for each fixed $r \in L$, is bounded. Hence, the sequence $\{v_r^k\}$ has a limit point, for all $r \in L$, say $\hat{v}_r$. Without loss of generality, we can assume that for all $r \in L$,

$$v_r^k \to \hat{v}_r, \ \text{as } k \to \infty. \tag{3.6}$$

Since $\partial^\circ g_r$'s is graph closed and $\hat{x}^k \to x_0$, one has $\hat{v}_r \in \partial^\circ g_r(x_0)$ for all $r \in L$.

Now, take the limit as $k \to \infty$ in Inequality (3.3) and in view of the above observations (3.4),(3.5) and (3.6), we get, $\left\|\sum_{r\in L} \hat{p}_r \hat{v}_r\right\| \leq 0$. Hence, we have $\sum\limits_{r\in L} \hat{p}_r \hat{v}_r = 0$, where $\hat{p} \in \mathbb{R}_+^l$ with $\|\hat{p}\| = 1$ and $\hat{v}_r \in \partial^\circ g_r(x_0)$ for all $r \in L$. This contradicts the assumption that BCQ holds at x_0. Therefore, we have shown the correctness of our claim, *i.e.*, the sequence $\{\mu^k\}$ is a bounded.

As $\{\mu^k\}$ is a bounded sequence, an argument similar to the one above, implies that there exist $\hat{\mu} \in \mathbb{R}_+^l$ such that $\mu_k \to \hat{\mu}$ *as* $k \to \infty$. Similarly, the sequences $\{\lambda^k\}$ and $\{u^k\}$ have limit points, say $\hat{\lambda}$ and $\hat{u}$, respectively, with $\|\hat{\lambda}\| = 1$ and $\lambda_k \to \hat{\lambda}$, $u^k \to \hat{u}$. Now taking $k \to \infty$ in Inequality (3.1), we get $\|\sum\limits_{i\in I} \hat{\lambda}_i \hat{u}_i + \sum\limits_{r\in L} \hat{\mu}_r \hat{v}_r\| \leq 0$. Thus

$$\sum_{i\in I} \hat{\lambda}_i \hat{u}_i + \sum_{r\in L} \hat{\mu}_r \hat{v}_r = 0, \quad \text{where } \hat{\lambda} \in \mathbb{R}_+^m \text{ with } \|\hat{\lambda}\| = 1,$$
$$\hat{\mu} \in \mathbb{R}_+^l, \hat{u}_i \in \partial^\circ f_i(x_0) \text{ and } \hat{v}_r \in \partial^\circ g_r(x_0). \tag{3.7}$$

Now, since x^k's are feasible points, $g_r(x^k) \leq 0$ for all $r \in L$. Using, the continuity property of g_r's, we conclude that $g_r(x_0) \leq 0$ for all $r \in L$. Hence, x_0 is a feasible point for MOP. Since, $\hat{\mu}_r \geq 0$ for all $r \in L$, we have $\sum_{r \in L} \hat{\mu}_r g_r(x_0) \leq 0$. Taking $k \to \infty$ in Inequality (3.2), we get $\sum_{r \in L} \hat{\mu}_r g_r(x_0) \geq 0$ and thus, we conclude that

$$\sum_{r \in L} \hat{\mu}_r g_r(x_0) = 0. \tag{3.8}$$

The Inequalities (3.7) and (3.8) together imply that x_0 is a KKT point of MOP. □

An interesting question one might ask is the following. Do we have some kind of converse for the above theorem? To be more precise, we ask ourselves the following question:
If a sequence of iterates from the feasible set converges to a weak Pareto minimizer, then do these iterates satisfy some approximate KKT conditions? The answer surprisingly turns out to be affirmative, and thus it strengthens the whole premise of studying approximate version of KKT condition for the multiobjective optimization problem. In the following theorem, we consider local weak Pareto solutions in place of weak Pareto solutions. The reason for studying local solution is that it provides an existence of a sequence of feasible points that converges to solution points. We can work with global solutions in place of local, but then we have to assume the existence of a converging sequence. With local solutions, we need to assume less which makes the result more interesting and general than the global case.

Theorem 3.5 *Consider the problem MOP with locally Lipschtiz data and g_r's for all $r \in L$ to be a convex function which satisfies the Slater constraint qualification. Further, assume that x_0 is a local weak Pareto minima and consider $\{\varepsilon_k\}$ to be a decreasing sequence of positive real numbers converging to 0. Then, there exists*

1. *a sequence $\{x^k\}$ of feasible points converging to $x_0 \in X$.*
2. *a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that for each y^k, there exists $\hat{y}^k$ satisfying*
   - *(a) $\|y^k - \hat{y}^k\| \leq \sqrt{\varepsilon_k}$,*
   - *(b) there exists $u_i^k \in \partial^\circ f_i(\hat{y}^k)$ and $v_r^k \in \partial^\circ g_r(\hat{y}^k)$, for all $i \in I$ and $r \in L$, such that*

$$\left\| \sum_{i \in I} \lambda_i^k u_i^k + \sum_{r \in L} \mu_r^k v_r^k \right\| \leq \sqrt{\varepsilon_k}, \tag{3.9}$$

$$\sum_{r \in L} \mu_r^k g_r(\hat{y}^k) = 0, \tag{3.10}$$

where $\lambda^k \in \mathbb{R}^m_+$ with $\|\lambda^k\| = 1$ and $\mu^k \in \mathbb{R}^l_+$.

Proof: **Part 1:** By assumption, x_0 is a locally weak Pareto minimizer of MOP, *i.e.*, there exists $\delta > 0$ such that

$$f(x) - f(x_0) \notin -int(\mathbb{R}^m_+), \text{ for all } x \in V, \tag{3.11}$$

where $V = X \cap \overline{B_\delta(x_0)}$. The convexity of the constraint functions g_r's together with closed convex feasible set X implies that V is a closed, convex and bounded set. As $x_0 \in V$, there exists a sequence x^k in X with x^k converging to $x_0 \in V$ and $x^k \in V$, for all k sufficiently large. This, completes the proof of Part 1.

Part 2: We have broken this proof in two steps. For the first step, we prove that there exists a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that $y^k \in V$ and is an $\varepsilon_k e$-weak Pareto minima of MOP with feasible set as V.

As f_i's, for $i \in I$, are locally Lipschitz, $f_i(x^k) \to f_i(x_0)$ as $k \to \infty$, for all $i \in I$. So, for a given $\varepsilon_1 > 0$, for each $i \in I$ there exist natural numbers N_1^i, such that

$$|f_i(x^k) - f_i(x_0)| < \varepsilon_1, \text{ for all } k \geq N_1^i.$$

Now choose $N_1 = \max\{N_1^1, N_1^2, \ldots, N_1^m\}$. Thus, for all $i \in I$

$$|f_i(x^k) - f_i(x_0)| < \varepsilon_1, \text{ for all } k \geq N_1. \tag{3.12}$$

Choose $y^1 = x^{N_1}$, then $|f_i(y^1) - f_i(x_0)| < \varepsilon_1$, or equivalently,

$$f(x_0) + e\epsilon_1 - f(y^1) \in int(\mathbb{R}^m_+). \tag{3.13}$$

Note that the sum of the spaces $\mathbb{R}^m \setminus (-int(\mathbb{R}^m_+))$ and $\mathbb{R}^m_+$ equals $\mathbb{R}^m \setminus (-int(\mathbb{R}^m_+))$ and hence adding Inequalities (3.11) and (3.13) gives

$$f(x) + e\epsilon_1 - f(y^1) \notin -int(\mathbb{R}^m_+), \text{ for all } x \in V. \tag{3.14}$$

Since $\epsilon_2 < \epsilon_1$, a similar argument applied to the sequence $\{x_{N_1}, x_{N_1+1}, x_{N_1+2}, \ldots\}$ gives an element $y^2 = x^{N_2}$, with $N_2 > N_1$, such that $f(x) + \varepsilon_2 e - f(y^2) \notin -int(\mathbb{R}^m_+)$ for all $x \in V$. Proceeding as above, gives a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that $y^k \in V$ and

$$f(x) + \varepsilon_k e - f(y^k) \notin -int(\mathbb{R}^m_+) \text{ for all } x \in V. \tag{3.15}$$

This completes the proof of the first step. We now come to the second step to complete the proof of the second part.

Since each f_i is locally Lipschitz, f is $(e, \mathbb{R}^m_+)$-lower semi continuous and $\mathbb{R}^m_+$-bounded below. Thus, the vector Ekeland Variational Principle (Theorem 2.8) gives the existence of $\hat{y}^k \in V$, for each $y^k \in V$, such that $\|\hat{y}^k - y^k\| \leq \sqrt{\varepsilon_k}$, and for all $x \in V \setminus \{\hat{y}^k\}$,

1. $f(x) + \varepsilon_k e - f(\hat{y}^k) \notin -int(\mathbb{R}^m_+)$, and

2. $f(x) + \sqrt{\varepsilon_k}\|\hat{y}^k - y^k\|e - f(\hat{y}^k) \notin -int(\mathbb{R}^m_+)$.

Thus from (c) above, we conclude that $\hat{y}^k$ is a weak Pareto minimizer of the problem

$$\min_{x \in V} g(x), \text{ where } g(x) = f(x) + \sqrt{\varepsilon_k}\|x - \hat{y}^k\|e.$$

Now, using the necessary optimality condition for the above multiobjective problem, there exists $\lambda^k \in \mathbb{R}^m_+$ with $\|\lambda^k\| = 1$ such that

$$0 \in \sum_{i \in I} \lambda_i^k \partial^\circ g_i(\hat{y}^k) + N_V(\hat{y}^k),$$

where $N_V(\hat{y}^k)$ is the normal cone to the set V at $\hat{y}^k$. Now applying sum rule for the Clarke subdifferential (see Clarke [5]) and using the fact that subdifferential of the norm function at origin is the unit ball, we get

$$0 \in \sum_{i \in I} \lambda_i^k \partial^\circ f_i(\hat{y}^k) + \sqrt{\varepsilon_k} B_1(0) + N_V(\hat{y}^k). \tag{3.16}$$

Since $x^k \to x_0$ and y^k is a sub-sequence of $\{x^k\}$, $y^k \in X \cap B_\delta(x_0)$, for sufficiently large k. As $\hat{y}^k \in B_{\sqrt{\epsilon_k}}(y^k)$ and $\epsilon_k \to 0$, for sufficiently large k, $B_{\sqrt{\epsilon_k}}(y^k) \subset B_\delta(x_0)$. Hence, $\hat{y}^k \in B_\delta(x_0)$, for k sufficiently large.

Clearly, $X \cap B_\delta(x_0) \neq \emptyset$ and hence $ri(X \cap B_\delta(x_0)) \neq \emptyset$. Thus, using Theorem 6.5 in Rockafellar [30], one obtains $ri(X \cap B_\delta(x_0)) = ri(X) \cap ri(B_\delta(x_0)) \neq \emptyset$. Now, using Theorem 23.8 in Rockafellar [30], we have

$$N_V(\hat{y}^k) = N_{X \cap \overline{B_\delta(x_0)}}(\hat{y}^k) = N_X(\hat{y}^k) + N_{\overline{B_\delta(x_0)}}(\hat{y}^k).$$

As $\hat{y}^k \in B_\delta(x_0) = int\overline{B_\delta(x_0)}$, we see that $N_{\overline{B_\delta(x_0)}}(\hat{y}^k) = \{0\}$. Thus $N_V(\hat{y}^k) = N_X(\hat{y}^k)$. Hence, we can rewrite (3.16) as

$$0 \in \sum_{i \in I} \lambda_i^k \partial^\circ f_i(\hat{y}^k) + \sqrt{\varepsilon_k} B_1(0) + N_X(\hat{y}^k). \tag{3.17}$$

Further as the Slater constraint qualification holds, using Corollary 23.7.1 of Rockafellar [30],

$$N_X(\hat{y}^k) = \{\sum_{r \in L} \mu_r^k v_r^k : v_r \in \partial^\circ g_r(\hat{y}^k),\ \mu_r^k \geq 0,\ \mu_r^k g_r^k(\hat{y}^k) = 0,\ r \in L\}.$$

Now using the above form of $N_X(\hat{y}^k)$ and (3.17), it is evident that there exists $u_i^k \in \partial^\circ f_i(\hat{y}^k)$ for all $i \in I$, $v_r^k \in \partial^\circ g_r(\hat{y}^k)$ for all $r \in L$ and scalars $\lambda^k \in \mathbb{R}^m_+$ with $\|\lambda^k\| = 1$, $\mu^k \in \mathbb{R}^l_+$ such that (3.9) and (3.10) holds. This completes the proof of the second part and hence the proof of the theorem is complete. □

Remark 3.6 *In the above theorem, the objective functions are taken to be locally Lipschitz only. If the objective function f_i's are convex as well, then we have a more concrete result. To proof the next result we need the following lemma.*

Lemma 3.7 *Consider the problem MOP with each objective functions f_i's to be convex. Then every local weak Pareto minima is a global weak Pareto minima.*

Proof: Let x_0 be a local weak Pareto minima of MOP, *i.e.,* there exists $\delta > 0$ such that for all $x \in X \cap B_\delta(x_0)$, $f(x) - f(x_0) \notin -int(\mathbb{R}^m_+)$. On the contrary assume that x_0 is not a global weak minima, *i.e.,* there exists $x^* \in X$ with

$$f(x^*) - f(x_0) \in -int(\mathbb{R}^m_+). \tag{3.18}$$

Thus, $f_i(x^*) < f_i(x_0)$, for all $i \in I$. As each f_i is a convex function, for all $t \in (0,1)$

$$f_i((1-t)x^* + tx_0) \leq (1-t)f_i(x^*) + tf_i(x_0) < f_i(x_0), \text{ for all } i \in I. \tag{3.19}$$

Now we can always choose $t_0 \in (0,1)$ such that $(1-t_0)x^* + t_0x_0 \in B_\delta(x_0)$. This contradicts the fact that x_0 is local weak Pareto minima. □

Before going to the theorem, we first discuss a result from Durea [10] which will play a key role in proving the next theorem.

Theorem 3.8 (Theorem 3.6 of [10]) *Let x_0 be a εe-weak Pareto minima of the problem MOP with each f_i's and g_r's to be convex functions and assume that Slater constraint qualification holds. Then x_0 is a modified σ-KKT point where $\sigma \in (0, \|e\|\varepsilon]$.*

Theorem 3.9 *Consider the problem MOP with each f_i and g_r being convex functions, for all $i \in I$ and $r \in L$. Let x_0 be a weak Pareto minima and let the Slater constraint qualification hold. Then, for decreasing sequence of positive real numbers $\{\varepsilon_k\}$ converging to 0, there exists a feasible sequence $\{x^k\}$ converging to x_0 and a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that each y^k is a modified σ_k-KKT point with $\sigma_k \in (0,\ \|e\|\varepsilon_k]$.*

Proof: Since the problem data is convex, local weak Pareto point is global. Now proceed as in the proof of Theorem 3.5 to get a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that y^k is a $\varepsilon_k e$-weak Pareto minima of MOP with feasible set as V), where $V = X \cap \overline{B_\delta(x_0)}$ with $\delta > 0$, *i.e.,* y^k is a local $\varepsilon_k e$-weak Pareto minima of MOP. So, by using the assumption of convexity and Lemma 3.7, we conclude that y^k is a $\varepsilon_k e$-weak Pareto minima of MOP. Now using Theorem 3.8, we conclude that y^k is a modified σ_k-KKT point with $\sigma_k \in (0,\ \|e\|\varepsilon_k]$. □

4 Approximate $\hat{M}$-Geoffrion solutions, Saddle points, and KKT conditions

In this section, we analyze saddle point conditions and KKT type conditions for the $(\hat{M}, \epsilon)$-Geoffrion solutions which give a complete characterization of the considered proper points. We also discuss a scalarization rule for the $(\hat{M}, \epsilon)$-Geoffrion solutions which is a connecting bridge for deducing saddle point and KKT type conditions. Before discussing the mentioned results, we shall observe that there is a characterization of $(\hat{M}, \epsilon)$-Geoffrion proper points by the system of inequalities. For a given $\epsilon \in \mathbb{R}^m_+$ and $\hat{M} > 0$, consider $x_0 \in X$, $i \in I$ and define the following system of inequalities $(\mathcal{Q}_i(x_0))$ as

$$\begin{cases} -f_i(x_0) + f_i(x) + \epsilon_i < 0, \\ -f_i(x_0) + f_i(x) + \epsilon_i < \hat{M}(f_j(x_0) - f_j(x) - \epsilon_j), \text{ for all } j \in I \setminus \{i\} \\ x \in X. \end{cases}$$

Proposition 4.1 *For given $\epsilon \in \mathbb{R}^m_+$ and $\hat{M} > 0$, consider the problem MOP. Then a point $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$ if and only if for each $i \in I$, the system $\mathcal{Q}_i(x_0)$ is inconsistent.*

The above proposition follows from the definition of Proof of the $(\hat{M}, \epsilon)$-Geoffrion proper solutions, for complete proof, see [34]. Before discussing the saddle point conditions for the $(\hat{M}, \epsilon)$-Geoffrion proper solutions, let us discuss the correspondence between (M, ϵ)-Geoffrion proper solutions and solution of the weighted sum scalar problem. This correspondence plays a pivotal role to prove saddle point conditions for (M, ϵ)-Geoffrion proper solutions. To this end, let for $s^* \in \mathbb{R}^m_+$, the weighted sum scalar problem $P(s^*)$ be defined as $\min\limits_{x \in X} \langle s^*, f(x) \rangle$.

Theorem 4.2 *For a given $\epsilon \in \mathbb{R}^m_+$, $\hat{M} > 0$, let x_0 is a $\langle s^*, \epsilon \rangle$-minimum of $P(s^*)$, where $s^* \in int(\mathbb{R}^m_+)$. If $\hat{M} \geq (m-1) \max\limits_{i,j} \{\frac{s^*_i}{s^*_j}\}$, then x_0 is a $(\hat{M}, \epsilon)$-Geoffrion proper solution of MOP, i.e., $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$.*

Proof: Let us assume on the contrary that $x_0 \notin \mathcal{G}_{\hat{M},\epsilon}(f, X)$. Therefore, from Proposition 4.1 we obtain an $i \in I$ such that $\mathcal{Q}_i(x_0)$ is consistent. Without loss of generality, we assume that $i = 1$. Thus, the system $\mathcal{Q}_i(x_0)$, written as

$$\begin{cases} -f_1(x_0) + f_1(x) + \epsilon_1 < 0, \\ -f_1(x_0) + f_1(x) + \epsilon_1 < \hat{M}(f_j(x_0) - f_j(x) - \epsilon_j), \quad j \in I \setminus \{1\} \\ x \in X. \end{cases}$$

has a solution. As $\hat{M} \geq (m-1)\{\frac{s^*_j}{s^*_i}\}$ for all $s^* \in \text{int}(\mathbb{R}^m_+)$, the consistency of system $\mathcal{Q}_i(x_0)$ implies that

$$s^*_1(-f_1(x_0) + f_1(x) + \epsilon_1) < s^*_j(m-1)(f_j(x_0) - f_j(x) - \epsilon_j), \text{ for all } j \in I \setminus \{1\}.$$

Summing the above equation for all $j \in I \setminus \{1\}$, we obtain that

$$s_1^*(-f_1(x_0) + f_1(x) + \epsilon_1) < \sum_{j=2}^{m} s_j^*(f_j(x_0) - f_j(x) - \epsilon_j),$$

which further implies

$$\langle s^*, f(x_0)\rangle - \langle s^*, f(x)\rangle - \langle s^*, \epsilon\rangle > 0. \tag{4.1}$$

Since (4.1) is a contradiction to the $\langle s^*, \epsilon\rangle$-minimality of $P(s^*)$. Therefore, the theorem follows. □

All the solutions from $\mathcal{G}_{\hat{M},\epsilon}(f, X)$ satisfy an upper trade-off bound of $\hat{M}$ (in the sense of Geoffrion-proper efficiency). Smaller bounds are more relevant to the decision maker as they provide tighter trade-offs among the criteria values. It would, therefore, be of interest to find the minimum M such that $\mathcal{G}_{M,\epsilon}(f, X)$ is non-empty. Under the conditions of Theorem 4.2, we need minimum value of $\hat{M}$ equals $m - 1$, and this occurs when all components of s^* are identical. The next example shows that if conditions in Theorem 4.2 are not satisfied, then even smaller values of $\hat{M}$ are possible. This is the case with non-convex or discrete multicriteria optimization problems. In the following example, we consider $\epsilon = 0$ and find $\hat{M}$-Geoffrion proper points.

Example 4.3 Let $X := \{(0, 0, 1)^\top, (0, 1, 0)^\top, (1, 0, 0)^\top, (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^\top\}$, $m = 3$, and f be the identity mapping. The sets $\mathcal{G}_2(f, X)$ and $\mathcal{G}_1(f, X)$ can be easily computed as follows:

$$\begin{aligned} \mathcal{G}_2(f, X) &= \{(0, 0, 1)^\top, (0, 1, 0)^\top, (1, 0, 0)^\top, (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^\top\}, \\ \mathcal{G}_1(f, X) &= \{(0, 0, 1)^\top, (0, 1, 0)^\top, (1, 0, 0)^\top\}. \end{aligned}$$

Moreover, $\mathcal{G}_M(f, X) = \emptyset$ for $M < 1$. Therefore, the minimum value of M is 1.

The converse of Theorem 4.2 also holds with convexity assumption on the objective functions and the feasible set. Since, if for each $r \in L$, g_r is convex, then the feasible set X is a convex set. We have the following result.

Theorem 4.4 *Let us consider the problem MOP where for each $i \in I$ and $r \in L$, f_i and g_r are convex functions. If $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$, then there exists an $s^* \in \operatorname{int}(\mathbb{R}^m_+)$ such that x_0 is a $\langle s^*, \epsilon\rangle$-minimum of $P(s^*)$.*

Proof: Let $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$. Then using Proposition 4.1, we obtain that the system $\mathcal{Q}_i(x_0)$ is inconsistent, for each $i \in I$. Applying the Gordan's Theorem of the alternative (see [30]), we conclude, after some rearrangements, that for each $i \in I$, there exists scalars $\lambda^i_j \geq 0$ with $\sum\limits_{j\in I} \lambda^i_j = 1$ such that, for all $x \in X$

$$f_i(x) + \hat{M} \sum_{j\in I, j\neq i} \lambda^i_j f_j(x) \geq f_i(x_0) + \hat{M} \sum_{j\in I, j\neq i} \lambda^i_j f_j(x_0) - \left[\epsilon_i + \hat{M} \sum_{j\in I, j\neq i} \lambda^i_j \epsilon_j\right].$$

Therefore, by summing over all i, we get

$$\begin{aligned} \sum_{i\in I} f_i(x) + \hat{M} \sum_{i\in I} \sum_{j\in I, j\neq i} \lambda^i_j f_j(x) \quad &\geq \quad \sum_{i\in I} f_i(x_0) + \hat{M} \sum_{i\in I} \sum_{j\in I, j\neq i} \lambda^i_j f_j(x_0) \\ &\qquad - \sum_{i\in I} \left[\epsilon_i + \hat{M} \sum_{j\in I, j\neq i} \lambda^i_j \epsilon_j\right]. \end{aligned}$$

Hence, for all $x \in X$,

$$\sum_{j\in I} \left[1 + \hat{M} \sum_{i\in I, i\neq j} \lambda^i_j\right] f_j(x) \geq \sum_{j\in I} \left[1 + \hat{M} \sum_{i\in I, i\neq j} \lambda^i_j\right] f_j(x_0) - \sum_{j\in I} \left[1 + \hat{M} \sum_{i\in I, i\neq j} \lambda^i_j\right] \epsilon_j.$$

Setting $s_j = 1 + \hat{M} \sum\limits_{i\in I, i\neq j} \lambda^i_j$, gives $s \in int(\mathbb{R}^m_+)$ and x_0 is a $\langle s, \epsilon\rangle$-minimum of $P(s)$. □

Remark 4.5 Theorem 4.4 can also be proved by noting the fact that each $(\hat{M}, \epsilon)$-Geoffrion proper point is ϵ-Geoffrion proper point with constant $\hat{M} > 0$. Hence using Theorem 3.15 form [14], we can deduce the above result. Now if we denote the set of $\langle s^*, \epsilon\rangle$-minimum of $P(s^*)$ by $Sol_\epsilon(P(s^*))$, then Theorem 4.2 and 4.4 implies that under convexity assumption on data and for a given $\hat{M}$, there exists $s^* \in \text{int}(\mathbb{R}^m_+)$ such that

$$Sol_\epsilon(P(s^*)) \subseteq \mathcal{G}_{\hat{M},\epsilon}(f, X) \subseteq \bigcup_{s\in \text{int}(\mathbb{R}^m_+)} Sol_\epsilon(P(s)).$$

Now we come to the main attraction of this section, the saddle point conditions for $(\hat{M}, \epsilon)$-Geoffrion proper solutions. For this study, we consider the problem MOP where each f_i, $i \in I$ and g_j, $j \in L$ are a convex function. Whenever the data of problem is convex , we shall denote the problem MOP as CMOP. Given $\hat{M} > 0$, and any index $i \in I$, we define the $(\hat{M}, i)$-Lagrangian associated with CMOP as follows

$$L^{\hat{M}}_i(x, \tau^i, \mu^i) = f_i(x) + \sum_{j\neq i} \tau^i_j \hat{M} f_j(x) + \sum_{r=1}^{l} \mu^i_r g_r(x), \tag{4.2}$$

where $\mu^i = (\mu^i_1, \mu^i_2, ..., \mu^i_l) \in \mathbb{R}^l_+$ and $\tau^i = (\tau^i_1, \tau^i_2, ..., \tau^i_m) \in S^m$ with $S^m = \{x \in \mathbb{R}^m : 0 \leq x_i \leq 1, i \in I, \sum_{i=1}^m x_i = 1\}$, the unit simplex in $\mathbb{R}^m$. The motivation behind considering the above Lagrangian comes from the ith-objective Lagrangian problem defined in

Chapter 4 of [4]. In [4], they used the above Lagrangian form as a scalarization scheme of multiobjective problems. In the same spirit as [4], we get a scalar structure of Lagrangian functions which is comparatively easy than vector-valued Lagrangian to work with. Our aim here is to show the key role played by the $(\hat{M}, i)$-Lagrangian in analyzing and characterizing the Geoffrion $(\hat{M}, \epsilon)$-Proper solutions.

Theorem 4.6 *For a given $\epsilon \in \mathbb{R}^m_+$ and $\hat{M} > 0$, let us consider the problem CMOP which satisfy the Slater constraint qualification. If $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$ then for each i, there exists $\bar{\tau}^i \in S^m$, $\bar{\mu}^i \in \mathbb{R}^l_+$ such that for all $x \in \mathbb{R}^n$ and $\mu \in \mathbb{R}^m_+$,*

(i) $L_i^{\hat{M}}(x_0, \bar{\tau}^i, \mu) - \bar{\epsilon}_i \leq L_i^{\hat{M}}(x_0, \bar{\tau}^i, \bar{\mu}^i) \leq L_i^{\hat{M}}(x, \bar{\tau}^i, \bar{\mu}^i) + \bar{\epsilon}_i$

(ii) $\sum_{r \in L} \bar{\mu}^i_r g_r(x_0) \geq -\bar{\epsilon}_i$,

where $\bar{\epsilon}_i = \epsilon_i + \sum_{j=1, j \neq i}^{m} \tau^i_j \hat{M} \epsilon_j$. Conversely if $x_0 \in \mathbb{R}^n$ be such that for each $i \in I$, there exists $(\bar{\tau}^i, \bar{\mu}^i) \in S^m \times \mathbb{R}^l_+$ such that (i) and (ii) holds then $x_0 \in \mathcal{G}_{\tilde{M},2\epsilon}(f, X)$, where $\tilde{M} \geq (1 + \hat{M})(m - 1)$.

Proof: It is evident from Proposition 4.1 that if $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$, then for each $i \in I$, the system $\mathcal{Q}_i(x_0)$, re-written as

$$\begin{aligned} &- f_i(x_0) + f_i(x) + \epsilon_i < 0, \\ &- f_i(x_0) + f_i(x) + \epsilon_i < M(f_j(x_0) - f_j(x) - \epsilon_j), \text{ for all } j \in I \setminus \{i\} \\ &g_r(x) \leq 0, \; r \in L \end{aligned}$$

has no solution, for all $x \in \mathbb{R}^n$. It is easy to observe that the system $\mathcal{Q}_i(x_0)$ has no solution, if we replace $g_r \leq 0$ by $g_r < 0$ for all $r \in L$. Now by applying the Gordan's theorem of the alternative (see [30]), there exists $\tau^i = (\tau^i_1, \dots, \tau^i_m) \in \mathbb{R}^m_+$ and $\mu^i = (\mu^i_1, \dots, \mu^i_l) \in \mathbb{R}^l_+$ with $(\tau^i, \mu^i) \neq 0$ such that for all $x \in \mathbb{R}^n$,

$$\tau^i_i(f_i(x) - f_i(x_0) + \epsilon_i) + \sum_{j \in I, j \neq i} \tau^i_j(f_i(x) + \hat{M} f_j(x) - f_i(x_0) - \hat{M} f_j(x_0) + \epsilon_i + \hat{M}\epsilon_j) + \sum_{r \in L} \mu^i_r g_r(x) \geq 0.$$

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

$$\Big(\sum_{j \in I} \tau^i_j\Big)(f_i(x) - f_i(x_0) + \epsilon_i) + \sum_{j \in I, j \neq i} \Big[\tau^i_j \hat{M} f_j(x) - \tau^i_j \hat{M} f_j(x_0) + \tau^i_j \hat{M} \epsilon_j\Big] + \sum_{r \in L} \mu^i_r g_r(x) \geq 0. \qquad (4.3)$$

Now, we first claim that $\tau^i = (\tau_1^i, \ldots, \tau_m^i) \neq 0$. For if, $\tau^i = 0$ then $\mu^i \neq 0$ and Inequality (4.3) reduces to $\sum_{r \in L} \mu_r^i g_r(x) \geq 0$, for all $x \in \mathbb{R}^n$. But, the Slater constraint qualification implies that there exists a point, say $\hat{x} \in \mathbb{R}^n$, such that $g_r(\hat{x}) < 0$. As $\mu^i \neq 0$ and $\mu^i \in \mathbb{R}_+^l$, we obtain $\sum_{r \in L} \mu_r^i g_r(x) < 0$, a contradiction to $\sum_{r \in L} \mu_r^i g_r(x) \geq 0$. Hence, $\tau^i \neq 0$ and thus $\sum_{j \in I} \tau_j^i > 0$. Thus, dividing Inequality (4.3) by $\sum_{j \in I} \tau_j^i$, we get

$$f_i(x) - f_i(x_0) + \epsilon_i + \sum_{j \in I, j \neq i} [\bar{\tau}_j^i \hat{M} f_j(x) - \bar{\tau}_j^i \hat{M} f_j(x_0) + \bar{\tau}_j^i \hat{M} \epsilon_j] + \sum_{r \in L} \bar{\mu}_r^i g_r(x) \geq 0, \quad (4.4)$$

for all $x \in \mathbb{R}^n$, where $\bar{\tau}_j^i = \frac{\tau_j^i}{\sum_{j \in I} \tau_j^i}$ and $\bar{\mu}_r^i = \frac{\mu_r^i}{\sum_{j \in I} \tau_j^i}$. In particular, for $x = x_0$, Inequality (4.4) gives $\epsilon_i + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} \epsilon_j + \sum_{r \in L} \bar{\mu}_r^i g_r(x_0) \geq 0$. By setting $\bar{\epsilon_i} = \epsilon_i + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} \epsilon_j$, we get Part (ii) as $\sum_{r \in L} \bar{\mu}_r^i g_r(x_0) \geq -\bar{\epsilon_i}$. Further, Inequality (4.4) reduces to, for all $x \in \mathbb{R}^n$,

$$f_i(x) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x) + \sum_{r \in L} \bar{\mu}_r^i g_r(x) + \bar{\epsilon_i} \geq f_i(x_0) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x_0). \quad (4.5)$$

As x_0 is feasible to CMOP, $\sum_{r \in L} \bar{\mu}_r^i g_r(x_0) \leq 0$. Thus, Inequality (4.5) becomes

$$f_i(x) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x) + \sum_{r \in L} \bar{\mu}_r^i g_r(x) + \bar{\epsilon_i} \geq f_i(x_0) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x_0) + \sum_{r \in L} \bar{\mu}_r^i g_r(x_0),$$

which implies that for each $i \in I$ and for all $x \in \mathbb{R}^n$,

$$L_i^{\hat{M}}(x, \bar{\tau}^i, \bar{\mu}^i) + \bar{\epsilon_i} \geq L_i^{\hat{M}}(x_0, \bar{\tau}^i, \bar{\mu}^i). \quad (4.6)$$

Further, from Equation (4.2), we observe that for all $i \in I$ and any $\mu \in \mathbb{R}_+^l$

$$L_i^{\hat{M}}(x_0, \bar{\tau}^i, \mu) \leq f(x_0) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x_0),$$

which can be written as $L_i^{\hat{M}}(x_0, \bar{\tau}^i, \mu) \leq f(x_0) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x_0) + \sum_{r \in L} \bar{\mu}_r^i g_r(x) + \bar{\epsilon_i}$. Thus, for all $x \in \mathbb{R}^n$ and $\mu \in \mathbb{R}_+^l$,

$$L_i^{\hat{M}}(x_0, \bar{\tau}^i, \mu) \leq L_i^{\hat{M}}(x_0, \bar{\tau}^i, \bar{\mu}^i) + \bar{\epsilon_i}. \quad (4.7)$$

The Inequalities (4.6) and (4.7) together prove Part (i). Now, for the sufficient part, let us assume that for a given $x_0 \in \mathbb{R}^n$ and each $i \in I$ there exists $\bar{\tau}^i \in S^m$ and $\bar{\mu}^i \in \mathbb{R}_+^l$ such that Conditions (i) and (ii) hold. Our first step would be to show that x_0 is feasible to CMOP. As we know from (i), for all $\mu \in \mathbb{R}^l_+$

$$L_i^{\hat{M}}(x_0, \bar{\tau}^i, \mu) - \bar{\epsilon}_i \leq L_i^{\hat{M}}(x_0, \bar{\tau}^i, \bar{\mu}^i).$$

Thus, $f_i(x_0) + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x_0) + \sum_{r\in L} \mu_r g_r(x_0) - \bar{\epsilon}_i \leq f_i(x_0) + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x_0)$. This shows that for all $\mu \in \mathbb{R}^l_+$,

$$\sum_{r\in L} \mu_r g_r(x_0) \leq \bar{\epsilon}_i. \tag{4.8}$$

On the contrary, suppose x_0 is not feasible. Then, there exists $r_0 \in L$ such that $g_{r_0}(x_0) > 0$. Then, choose $\mu = (0, \ldots, 0, \mu_{r_0}, 0, \ldots, 0)$, with $\mu_{r_0} > 0$ and sufficiently large such that $\mu_{r_0} g_{r_0}(x_0) > \bar{\epsilon}_i$. Note that this contradicts Inequality (4.8). Hence, we conclude that x_0 is a feasible solution of CMOP.

Now from right hand side of (i) we also have, for all $x \in \mathbb{R}^n$

$$L_i^{\hat{M}}(x, \bar{\tau}^i, \bar{\mu}^i) + \bar{\epsilon}_i \geq L_i^{\hat{M}}(x_0, \bar{\tau}^i, \bar{\mu}^i). \tag{4.9}$$

which implies

$$\begin{aligned} f_i(x) + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x) + \sum_{r\in L} \bar{\mu}_r^i g_r(x) + \epsilon_i + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} \epsilon_j \geq f_i(x_0) \\ + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x_0) + \sum_{r\in L} \bar{\mu}_r^i g_r(x_0). \end{aligned}$$

Now, for any feasible x, $\sum_{r\in L} \bar{\mu}_r^i g_r(x) \leq 0$. Thus, from the above inequality we have,

$$\begin{aligned} f_i(x) + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x) + \epsilon_i + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} \epsilon_j \geq f_i(x_0) \\ + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x_0) + \sum_{r\in L} \bar{\mu}_r^i g_r(x_0). \end{aligned} \tag{4.10}$$

Using Condition (ii), we have

$$\begin{aligned} f_i(x) + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x) + \epsilon_i + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} \epsilon_j \geq f_i(x_0) \\ + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} f_j(x_0) - (\epsilon_i + \sum_{j\in I, j\neq i} \bar{\tau}_j^i \hat{M} \epsilon_j). \end{aligned}$$

Since, it holds for each i, by summing over all the i's we get,

$$\sum_{i\in I}(1 + \hat{M} \sum_{j\in I, j\neq i} \bar{\tau}_j^i) f_i(x) + \sum_{i\in I}(1 + \hat{M} \sum_{j\in I, j\neq i} \bar{\tau}_j^i)(2\epsilon_j) \geq \sum_{i\in I}(1 + \hat{M} \sum_{j\in I, j\neq i} \bar{\tau}_j^i) f_i(x_0).$$

Hence, x_0 is $\langle s, 2\epsilon\rangle$-minimizer of $P(s)$, where $s = (s_1, \ldots, s_m)$ with $s_i = 1 + \hat{M} \sum_{k\in I, k\neq i} \bar{\tau}^i_k$, for $i \in I$. Now since $\bar{\tau}^i \in S^m$ for all i, we have for all $i, j \in I$

$$\frac{s_i}{s_j} = \frac{1 + \hat{M} \sum\limits_{k\in I, k\neq i} \bar{\tau}^i_k}{1 + \hat{M} \sum\limits_{k\in I, k\neq j} \bar{\tau}^j_k} = \frac{1 + \hat{M}(1 - \bar{\tau}^i_i)}{1 + \hat{M}(1 - \bar{\tau}^j_j)} \leq 1 + \hat{M}.$$

Since the above inequality is true for every i and j, we have $\max\limits_{i,j}\{\frac{s_i}{s_j}\} \leq 1 + \hat{M}$. Now consider $\tilde{M} \geq (1+\hat{M})(m-1)$ and using Theorem 4.2, we conclude that $x_0 \in \mathcal{G}_{\tilde{M},2\epsilon}(f, X)$. This completes the proof. □

Remark 4.7 The saddle point type conditions are useful as a sufficient condition if the number of objectives are only few in number. In fact, for sufficiency we can have a much simpler condition which we now state. Let $x_0 \in \mathbb{R}^n$ be a point that satisfies:
for each $i \in I$, there exists $\bar{\tau}^i \in S^m$ and $\bar{\mu}^i \in \mathbb{R}^l_+$ such that for all $\mu \in \mathbb{R}^l_+$ and $x \in \mathbb{R}^n$,

(a) $L^{\hat{M}}_i(x_0, \bar{\tau}^i, \mu) - \epsilon_i \leq L^{\hat{M}}_i(x_0, \bar{\tau}^i, \bar{\mu}^i) \leq L^{\hat{M}}_i(x, \bar{\tau}^i, \bar{\mu}^i) + \epsilon_i$,

(b) $\sum\limits_{r\in L} \bar{\mu}^i_r g_r(x_0) \geq -\epsilon_i$.

Then, $x_0 \in \mathcal{G}_{\tilde{M},2\epsilon}(f, X)$.

In order to prove the above statement, note that $\bar{\epsilon}_i = \epsilon_i + \sum\limits_{j=1, j\neq i} \bar{\tau}^i_j \hat{M}\epsilon_j$. So, $\bar{\epsilon}_i \geq \epsilon_i$. Hence, Conditions (a) and (b) above implies that Conditions (i) and (ii) of Theorem 4.6 are satisfied. Therefore, we can simply apply the converse part of Theorem 4.6 to get $x_0 \in \mathcal{G}_{\tilde{M},2\bar{\epsilon}}(f, X)$, where $\tilde{M} \geq (1 + \hat{M})(m - 1)$. Note that Condition (a) and (b) above are much simpler as compared to checking Conditions (i) and (ii) as $\bar{\epsilon}_i$ involves the multipliers $\bar{\tau}^i_j$. Hence, for the sufficiency part of Theorem 4.6 which requires the verification of Conditions (i) and (ii), we will be using Conditions (a) and (b).

Of course from the necessary part of Theorem 4.6, we can also derive a multiplier rule involving ϵ-subdifferentials, however this rule will be quite different. Observe that if $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$, then Condition (i) of Theorem 4.6 implies that for any $i \in I$ there exists $\bar{\tau}^i \in S^m$ and $\bar{\mu}^i \in \mathbb{R}^l_+$ such that for all $x \in \mathbb{R}^n$,

$$L^{\hat{M}}_i(x_0, \bar{\tau}^i, \bar{\mu}^i) \leq L^{\hat{M}}_i(x, \bar{\tau}^i, \bar{\mu}^i) + \bar{\epsilon}_i,$$

which implies that $x_0 \in \bar{\epsilon}_i - \arg\min\limits_{x\in\mathbb{R}^n} L^{\hat{M}}_i(\ \cdot, \bar{\tau}^i, \bar{\mu}^i)$, where $\bar{\epsilon}_i - \arg\min$ is the set of $\bar{\epsilon}_i$-minima of the function $L^{\hat{M}}_i(x, \bar{\tau}^i, \bar{\mu}^i)$. Thus, for each $i \in I$, $0 \in \partial_{\bar{\epsilon}_i} L^{\hat{M}}_i(x_0, \bar{\tau}^i, \bar{\mu}^i)$. In fact a more compact necessary condition of the KKT type is given as follows,

$$0 \in \sum_{i=\in I} \partial_{\bar{\epsilon}_i} L^{\hat{M}}_i(x_0, \bar{\tau}^i, \bar{\mu}^i) \quad \text{with} \quad \sum_{r\in L} \bar{\mu}^i_r g_r(x_0) \geq -\bar{\epsilon}_i. \tag{4.11}$$

Theorem 4.8 *For a given* $\epsilon \in \mathbb{R}^m_+$ *and* $\hat{M} > 0$*, let us consider the problem CMOP. If* $x_0 \in \mathcal{G}_{\hat{M},\epsilon}(f, X)$*, then there exist vectors* $\bar{\tau}^i \in S^m$ *and* $\bar{\mu}^i \in \mathbb{R}^l_+$*,* $i \in I$ *such that*

(A) $0 \in \sum_{i \in I} \partial_{\bar{\epsilon}_i} L_i^{\hat{M}}(x_0, \bar{\tau}^i, \bar{\mu}^i)$,

(B) $\sum_{r=1}^{l} \bar{\mu}_r^i g_r(x_0) \geq -\bar{\epsilon}_i$,

where $\bar{\epsilon}_i = \epsilon_i + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} \epsilon_j$*,* $i \in I$*. Conversely, if* $x_0 \in X$ *be a point for which there exist vectors* $(\bar{\tau}^i, \bar{\mu}^i) \in S^m \times \mathbb{R}^l_+$*,* $i \in I$ *such that (A) and (B) hold then* $x_0 \in \mathcal{G}_{\tilde{M},2\epsilon}(f, X)$*, where* $\tilde{M} = (1 + \hat{M})(m - 1)$*.*

Proof: The necessary part has already been done in above remark. For sufficient part, let conditions (A) and (B) hold for $x_0 \in X$. This means that there exists $\bar{v}^i \in \partial_{\bar{\epsilon}_i} L_i^{\hat{M}}(x_0, \bar{\tau}^i, \bar{\mu}^i)$ for all $i \in I$ such that

$$0 = \bar{v}^1 + \bar{v}^2 + \ldots + \bar{v}^m. \tag{4.12}$$

Thus, from definition of ϵ-subdifferential, for each $i \in I$,

$$L_i^{\hat{M}}(x, \bar{\tau}, \bar{\mu}^i) - L_i^{\hat{M}}(x, \bar{\tau}^i, \bar{\mu}^i) \geq \langle \bar{v}^i, x - x_0 \rangle - \bar{\epsilon}^i.$$

Hence,

$$\sum_{i \in I} L_i^{\hat{M}}(x, \bar{\tau}, \bar{\mu}^i) - \sum_{i \in I} L_i^{\hat{M}}(x, \bar{\tau}^i, \bar{\mu}^i) \geq \langle \sum_{i \in I} \bar{v}^i, x - x_0 \rangle - \sum_{i \in I} \bar{\epsilon}^i.$$

Now using Equation (4.12), we get

$$\sum_{i \in I} (f_i(x) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x) + \sum_{r \in L} \bar{\mu}_r^i g_r(x)) - \sum_{i \in I} (f_i(x_0) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x_0) + \sum_{r \in L} \bar{\mu}_r^i g_r(x_0)) \geq -\sum_{i \in I} \bar{\epsilon}_i.$$

So, if x is a feasible point then using Condition (B), the above inequality reduces to

$$\sum_{i \in I} (f_i(x) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x)) \geq \sum_{i \in I} (f_i(x_0) + \sum_{j \in I, j \neq i} \bar{\tau}_j^i \hat{M} f_j(x_0)) - \sum_{i \in I} 2\bar{\epsilon}_i,$$

which can be rewritten as

$$\sum_{i \in I} (1 + \sum_{i \in I, i \neq j} \bar{\tau}_j^i \hat{M}) f_i(x) \geq \sum_{i \in I} (1 + \sum_{i \in I, i \neq j} \bar{\tau}_j^i \hat{M}) f_i(x_0) - \sum_{i \in I} (1 + \sum_{i \in I, i \neq j} \bar{\tau}_j^i \hat{M}) 2\epsilon_i.$$

Hence, x_0 is $\langle s, 2\epsilon \rangle$-minimizer of $P(s)$ where $s_i = 1 + \hat{M} \sum_{k \in I, k \neq i} \bar{\tau}_k^i$. Now using the same argument as in Theorem 4.2, we conclude that $x_0 \in \mathcal{G}_{\tilde{M},2\epsilon}(f, X)$, where $\tilde{M} = (1 + \hat{M})(m - 1)$. This completes the proof. □

5 Approximate KKT conditions for Approximate Benson solutions

In this section, we are concerned with a multiobjective problem in which the ordering of image space is induced with a closed, convex cone $C \subset \mathbb{R}^m$. Recall that these problems are denoted by (f, X, C) in the first section. The center of the present section is ϵ-Benson proper solutions and our aim is to develop a KKT type necessary and sufficient optimality conditions for ϵ-Benson proper solutions with the help of scalarization techniques. First, we will develop Scalarization rules for ϵ-Benson proper solutions which are keys tools needed to derive optimality conditions. It is important to note that when $C = \mathbb{R}^n_+$, ϵ-Benson proper Pareto solutions reduce to Geoffrion ϵ-proper solutions. We would like to note that our result is different from the optimality condition for approximate solutions studied in Dutta and Vetrivel [13].
To develop scalarization techniques for ϵ-Benson proper, we will make use of the following cone separation result.

Proposition 5.1 (Borwein, [3]) *Let K, N be closed, convex cones in $\mathbb{R}^m$ and let $N \cap K = \{0\}$. If the dual cone K^* has nonempty interior, then*

$$(K^*)^0 \cap (-N^*) \neq \emptyset. \tag{5.1}$$

The next two theorems present and analyze a scalarization technique for ϵ-Benson proper solutions.

Theorem 5.2 *Let us consider the problem (f, X, C), where C be a closed, convex cone such that $(C^*)^0 \neq \emptyset$. For $s^* \in (C^*)^0$ and $\epsilon \in \mathbb{R}^m_+$, if x_0 is an $\langle s^*, \epsilon\rangle$-minimum of $P(s^*)$, then $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$.*

Proof: Let $h \neq 0$ and $h \in \text{cl}(\text{cone}(f(X) + (C + \epsilon) - f(x_0)))$. Our aim is to establish that $h \notin -C$. From the definition of h, there exists a sequence $\{h_n\}$ such that $h_n \in \text{cone}(f(X) + (C + \epsilon) - f(x_0)))$ and $h_n = t_n(f(x_n) + s_n + \epsilon - f(x_0)) \to h$, with $t_n \geq 0$, $x_n \in X$, and $s_n \in C$. Now since x_0 is an $\langle s^*, \epsilon\rangle$-minimum of $P(s^*)$, for each n it holds that $\langle s^*, f(x_0)\rangle \leq \langle s^*, f(x_n)\rangle + \langle s^*, \epsilon\rangle$. This shows that

$$\langle s^*, f(x_n) + \epsilon - f(x_0)\rangle \geq 0. \tag{5.2}$$

Since $s_n \in C$ and $s^* \in (C^*)^0$, we have that

$$\langle s^*, s_n\rangle \geq 0. \tag{5.3}$$

Adding (5.2) and (5.3) gives $\langle s^*, h_n\rangle \geq 0$. Now as $h_n \to h$, we obtain

$$\langle s^*, h\rangle \geq 0. \tag{5.4}$$

As $s^* \in (C^*)^0$ and $h \neq 0$, assuming $h \in -C$ would result in $\langle s^*, -h\rangle > 0$, which is a contradiction to (5.4). Therefore, $h \notin -C$ and $x_0 \in \mathcal{S}^{\epsilon}_{\mathcal{B}}(f, X, C)$. □

We shall now show that the converse of the above result by assuming that the vector function f is a C-convex function. The function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called C-convex, where C is a closed convex cone if for any $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$

$$\lambda f(y) + (1 - \lambda)f(x) - f(\lambda y + (1 - \lambda)x) \in C.$$

The first and foremost consequence of C-convexity of f is that $\langle s^*, f(x)\rangle$ is a convex function in x, for any $s \in C^*$, which is a trivial implication of the definition of a C-convex function. The convex structure of the scalar-valued function $\langle s^*, f(x)\rangle$ is very useful in devising KKT type conditions for ϵ-Benson proper solutions and it also allow us to use convex subdifferential for this function. Now under appropriate convexity assumptions, we can show the converse of Theorem 5.2.

Theorem 5.3 *Consider the problem (f, X, C), where f is a C-convex function, and X is a closed convex set and let $C \subset \mathbb{R}^m$ be a closed, convex cone such that $(C^*)^0 \neq \emptyset$. If for a given $\epsilon \in \mathbb{R}^m_+$, $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$, then there exists an $s^* \in (C^*)^0$ such that x_0 is an $\langle s^*, \epsilon\rangle$-minimum of $P(s^*)$.*

Proof: Since f is C-convex and X is a convex set, $f(X) + C$ is convex. Therefore, $f(X) + C + \epsilon - f(x_0)$ is also a convex set. For simplicity let us set $\bar{K} := \mathrm{cl}(\mathrm{cone}(f(X) + C + \epsilon - f(x_0)))$. Hence $\bar{K}$ is a closed, convex cone. Furthermore, as $x_0 \in \mathcal{S}^{\epsilon}_{\mathcal{B}}(f, X, C)$, it follows that $\bar{K} \cap (-C) = \{0\}$. Equivalently, $(-\bar{K}) \cap C = \{0\}$. Using Proposition 5.1, we conclude that

$$(C^*)^0 \cap \bar{K}^* \neq \emptyset.$$

Let $s^* \in (C^*)^0 \cap \bar{K}^*$. This means that $\langle s^*, k\rangle \geq 0$ for all $k \in \bar{K}$. In particular, for every $x \in X$ and $c \in C$,

$$\langle s^*, f(x) + c + \epsilon - f(x_0)\rangle \geq 0.$$

This further implies that

$$\langle s^*, f(x)\rangle - \langle s^*, f(x_0)\rangle \geq -\langle s^*, c\rangle - \langle s^*, \epsilon\rangle. \tag{5.5}$$

Setting $c = 0$ in (5.5) we obtain

$$\langle s^*, f(x)\rangle - \langle s^*, f(x_0)\rangle \geq -\langle s^*, \epsilon\rangle,$$

which implies that x_0 is an $\langle s^*, \epsilon\rangle$-minimizer of $P(s^*)$. □

As a direct consequence of Theorems 5.2 and 5.3 we obtain the following result.

Corollary 5.4 *Consider the problem (f, X, C), where f is a C-convex function, and X is a closed convex set and let $C \subset \mathbb{R}^m$ be a closed, convex cone such that $(C^*)^0 \neq \emptyset$. Then, $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$ if and only if there exists an $s^* \in (C^*)^0$ such that x_0 is an $\langle s^*, \epsilon \rangle$-minimizer of $P(s^*)$.*

The next theorem characterizes ϵ-Benson proper solutions in the convex case through a KKT type approximate optimality condition. Now we shall state and prove the main result of the section.

Theorem 5.5 *Let us consider the problem (f, X, C) with f to be a C-convex function and each g_r, $r \in L$ to be a convex function for all $r \in L$ and assume that the Slater's constraint qualification holds. Then, for a given $\epsilon \in \mathbb{R}^m_+$, $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$ if and only if there exist $s^* \in (C^*)^0$, ε_0, λ_r, $\varepsilon_r \geq 0$ for all $r \in L$ such that*

$$0 \in \partial_{\varepsilon_0} \langle s^*, f(x_0) \rangle + \sum_{r \in L} \partial_{\varepsilon_r} (\lambda_r g_r)(x_0) \text{ and} \tag{5.6}$$

$$\sum_{r=0}^{l} \varepsilon_r - \langle s^*, \epsilon \rangle \leq \sum_{r \in L} \lambda_r g_r(x_0) \leq 0. \tag{5.7}$$

Proof: Let us first demonstrate that the condition are necessary. In this direction let us consider $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$. From Theorem 5.3, there exists $s^* \in (C^*)^0$ such that x_0 is an $\langle s^*, \epsilon \rangle$- minimum of the scalar problem $P(s^*)$. Since the Slater Constraints Qualification holds, consider the problem $P(s^*)$,

$$\begin{cases} \ \min \ \langle s^*, f(x) \rangle \\ \text{s.t. } g_r(x) \leq 0, \ r \in L, \end{cases}$$

Now from Theorem 10.9 of [9], we get that there exists scalars $\varepsilon_0 \geq 0,$, $\varepsilon_r \geq 0$ for all $r \in L$, and $\lambda_r \geq 0$ for all $r \in L$ such that the following relations holds:

$$0 \in \partial_{\varepsilon_0} \langle s^*, f(x_0) \rangle + \sum_{r \in L} \partial_{\varepsilon_r} (\lambda_r g_r)(x_0), \text{ and}$$

$$\sum_{r=0}^{l} \varepsilon_r - \langle s^*, \epsilon \rangle \leq \sum_{r \in L} \lambda_r g_r(x_0) \leq 0.$$

This yields the desired necessary part of the theorem. In order to establish the sufficient part, let us assume that conditions (5.6) and (5.7) of the theorem hold. From (5.6) we obtain $\phi \in \partial_{\varepsilon_0} \langle s^*, f(x_0) \rangle$ and $\eta_r \in \partial_{\varepsilon_r} (\lambda_r g_r)(x_0)$, for all $r \in L$ such that

$$\phi + \sum_{r \in L} \eta_r = 0. \tag{5.8}$$

Now using definition of ε_0-subdifferential of $\langle s^*, f(x)\rangle$ and $\varepsilon_r(x)$-subdifferential of $\lambda_r g_r$ for $r \in L$, we have that for all $x \in X$ the following two inequalities hold,

$$\langle s^*, f(x)\rangle - \langle s^*, f(x_0)\rangle \geq \langle \phi, x - x_0\rangle - \varepsilon_0, \tag{5.9}$$

$$\sum_{r\in L} \lambda_r g_r(x) - \sum_{r\in L} \lambda_r g_r(x_0) \geq \left\langle \sum_{r\in L} \eta_r, x - x_0 \right\rangle - \sum_{r\in L} \varepsilon_r. \tag{5.10}$$

By adding (5.9) and (5.10) and using Equation (5.8), we get

$$\langle s^*, f(x)\rangle - \langle s^*, f(x_0)\rangle + \sum_{r\in L} \lambda_r g_r(x) - \sum_{r\in L} \lambda_r g_r(x_0) \geq \sum_{r=0}^{l} \varepsilon_r$$

Now, noting that for any feasible point $x \in X$, $\sum_{r\in L} \lambda_r g_r(x) \leq 0$, and using Condition (5.7), we conclude that $\langle s^*, f(x)\rangle - \langle s^*, f(x_0)\rangle \geq -\langle s^*, \epsilon\rangle$, for all $x \in X$. Therefore, x_0 is a $\langle s^*, \epsilon\rangle$-minimum of $P(s^*)$. Using C-convexity of function f and Corollary 5.4, $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$. □

In Theorem 5.5 we obtained a characterization for ϵ-Benson proper solution using ϵ-subdifferentials. Computing ϵ-subdifferentials is not straight forward unless the convex functions has special structures. Meanwhile in several cases there are formulas for computing the exact subdiffererntial of convex function (see for example[22]). Thus in simple cases we can compute the exact subsdifferential more easily than ϵ-subdifferentials. In the following theorem we show how subdifferentials of a convex function can be used to obtain a necessary and sufficient condition for ϵ-Benson proper solutions.

Theorem 5.6 *Let us consider the problem (f, X, C) with f to be a C-convex function and X to be a closed, convex set. If for $\epsilon \in \mathbb{R}^m_+$, $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$ with $\epsilon > 0$, then there exists $s^* \in (C^*)^0$, $\hat{x} \in B\left(x_0, \sqrt{\langle s^*, \epsilon\rangle}\right)$ such that $\hat{x} \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$. Furthermore,*

$$0 \in \partial\langle s^*, f(\hat{x})\rangle + N_X(\hat{x}) + \sqrt{\langle s^*, \epsilon\rangle} B_1(0), \tag{5.11}$$

where $N_X(\hat{x})$ is the normal cone to the convex set X at $\hat{x} \in X$ and $B_1(0)$ is the unit ball in $\mathbb{R}^n$.

Proof: Let us first consider the case $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$. As f is C-convex, X closed and convex, and $x_0 \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$, applying Theorem 5.3, we obtain $s^* \in (C^*)^0$ such that x_0 is an $\langle s^*, \epsilon\rangle$-minimum of $P(s^*)$. Therefore, x_0 is an $\langle s^*, \epsilon\rangle$-minimum of the unconstrained problem

$$\inf_{x\in\mathbb{R}^n} \langle s^*, f(x)\rangle + \delta_X(x), \tag{5.12}$$

where δ_X is the indicator function of the convex set X. This means that for all $x \in \mathbb{R}^n$, $\langle s^*, f(x_0)\rangle + \delta_X(x_0) \leq \langle s^*, f(x)\rangle + \delta_X(x) + \langle s^*, \epsilon\rangle$.

Note that the indicator function is a proper lower-semicontinuous function as X is closed and convex [13], therefore using the Ekeland Variational Principle (Theorem 2.6), there exists a point $\hat{x} \in B\left(x_0, \sqrt{\langle s^*, \epsilon \rangle}\right)$ such that

$$\langle s^*, f(\hat{x}) \rangle + \delta_X(\hat{x}) \leq \langle s^*, f(x_0) \rangle + \delta_X(x_0) + \langle s^*, \epsilon \rangle, \tag{5.13}$$

and, for all $x \neq \hat{x}$,

$$\langle s^*, f(\hat{x}) \rangle + \delta_X(\hat{x}) \leq \langle s^*, f(x) \rangle + \delta_X(x) + \sqrt{\langle s^*, \epsilon \rangle} \|x - \hat{x}\|. \tag{5.14}$$

From (5.13), we obtain that $\hat{x} \in X$ (otherwise $\delta_X(\hat{x}) = +\infty$, whereas the right side of (5.13) is finite) which imply that $\hat{x}$ is an $\langle s^*, \epsilon \rangle$-minimum of $P(s^*)$. Thus using Theorem 5.2, $\hat{x} \in \mathcal{S}_{\mathcal{B},\epsilon}(f, X, C)$. Furthermore, using (5.14), we obtain

$$0 \in \partial \left(\langle s^*, f \rangle + \delta_X + \sqrt{\langle s^*, \epsilon \rangle} \|x - .\| \right)(\hat{x}).$$

Further noting that $\partial \delta_X(\hat{x}) = N_X(\hat{x})$, $\partial(\|x - \hat{x}\|)|_{x=\hat{x}} = B_1(0)$. Thus by the sum rule for subdifferentials of convex functions (see [30]), we obtain (5.11) which completes the proof. □

6 Concluding remarks

To analyze the behavior of an optimization problem from the view point of KKT conditions is deep rooted in psyche of researchers in optimization theory. Though KKT conditions may not have been used very heavily in multiobjective optimization but they can however act very well as a tool to develop stopping criteria. In this article, we characterize approximate versions of Pareto and proper Pareto solution using KKT type conditions. In fact, in the convex case, we achieve a complete characterization, for example, Theorem 3.5 demonstrates that a sequence of points which converge to weak Pareto minimizer has a subsequence where each point satisfies as approximate version of the KKT conditions. This result thus demonstrates the reason why approximate KKT type conditions can be used as a stopping criteria.
The analysis of the approximate versions of the M-Geoffrion proper solutions and Benson proper solutions in terms of approximate KKT conditions is a starting point for building stopping criteria to identify such points. Our future research would involve more computational studies by using these optimality conditions as a stopping criteria.

References

[1] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming: theory and algorithms*. John Wiley & Sons, 2013.

[2] H. P. Benson. An improved definition of proper efficiency for vector maximization with respect to cones. *J. Math. Anal. Appl.*, 71(1):232–241, 1979.

[3] J. Borwein. Proper efficient points for maximizations with respect to cones. *SIAM Journal on Control and Optimization*, 15(1):57–63, 1977.

[4] V. Chankong and Y. Y. Haimes. *Multiobjective decision making: theory and methodology.* Courier Dover Publications, 2008.

[5] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5. Siam, 1990.

[6] K. Deb. *Multi-objective optimization using evolutionary algorithms*, volume 16. John Wiley & Sons, 2001.

[7] K. Deb, M. Abouhawwash, and J. Dutta. An optimality theory based proximity measure for evolutionary multi-objective and many-objective optimization. In *International Conference on EMO*, pages 18–33. Springer, 2015.

[8] K. Deb, R. Tewari, M. Dixit, and J. Dutta. Finding trade-off solutions close to kkt points using evolutionary multi-objective optimization. In *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*, pages 2109–2116. IEEE, 2007.

[9] A. Dhara and J. Dutta. *Optimality conditions in convex optimization: a finite-dimensional view.* CRC Press, 2011.

[10] M. Durea, J. Dutta, and C. Tammer. Stability properties of KKT points in vector optimization. *Optimization*, 60(7):823–838, 2011.

[11] J. Dutta, K. Deb, R. Tulshyan, and R. Arora. Approximate KKT points and a proximity measure for termination. *Journal of Global Optimization*, 56(4):1463–1499, 2013.

[12] J. Dutta and V. Vetrivel. On approximate minima in vector optimization. *Numerical Functional Analysis and Optimization*, 22(7-8):845–859, 2001.

[13] J. Dutta and V. Vetrivel. On approximate minima in vector optimization. *Numer. Funct. Anal. Optim.*, 22(7-8):845–859, 2001.

[14] M. Ehrgott. *Multicriteria optimization.* Springer-Verlag, Berlin, second edition, 2005.

[15] G. Eichfelder. *Adaptive scalarization methods in multiobjective optimization*, volume 436. Springer, 2008.

[16] I. Ekeland. On the variational principle. *Journal of Mathematical Analysis and Applications*, 47(2):324–353, 1974.

[17] A. M. Geoffrion. Proper efficiency and the theory of vector maximization. *J. Math. Anal. Appl.*, 22:618–630, 1968.

[18] C. Gutiérrez, B. Jiménez, and V. Novo. On approximate efficiency in multiobjective programming. *Mathematical Methods of Operations Research*, 64(1):165–185, 2006.

[19] C. Gutiérrez, B. Jiménez, and V. Novo. Optimality conditions via scalarization for a new ε-efficiency concept in vector optimization problems. *European Journal of Operational Research*, 201(1):11–22, 2010.

[20] S. Helbig and D. Pateva. On several concepts for ϵ-efficiency. *Operations-Research-Spektrum*, 16(3):179–186, 1994.

[21] M. I. Henig. Proper efficiency with respect to cones. *J. Optim. Theory Appl.*, 36(3):387–407, 1982.

[22] J. B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms. II*, volume 306. Springer-Verlag, Berlin, 1993.

[23] J. Jahn. *Vector optimization.* Springer-Verlag, Berlin, 2004. Theory, applications, and extensions.

[24] S. S. Kutateladze. Convex ε-programming. In *Soviet Math. Dokl*, volume 20, pages 391–393, 1979.

[25] J. C. Liu. ϵ-properly efficient solutions to nondifferentiable multiobjective programming problems. *Appl. Math. Lett.*, 12(6):109–113, 1999.

[26] P. Loridan. ε-solutions in vector minimization problems. *J. Optim. Theory Appl.*, 43(2):265–276, 1984.

[27] D. T. Luc. Scalarization of vector optimization problems. *J. Optim. Theory Appl.*, 55(1):85–102, 1987.

[28] E. K. Makarov and N. N. Rachkovski. Unified representation of proper efficiency by means of dilating cones. *J. Optim. Theory Appl.*, 101(1):141–165, 1999.

[29] Y. Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2013.

[30] R. T. Rockafellar. *Convex analysis.* Princeton university press, 2015.

[31] R. T. Rockafellar and R. J.B. Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.

[32] W. D. Rong and Y. N. Wu. e-weak minimal solutions of vector optimization problems with set-valued maps. *J. Optim. Theory Appl.*, 106(3):569–579, 2000.

[33] P. K. Shukla, J. Dutta, and K. Deb. On properly Pareto optimal solutions. *Practical Approaches to Multi-Objective Optimization, Dagstuhl Seminar Proceedings 04461*, 2005.

[34] P. K. Shukla, J. Dutta, K. Deb, and P. Kesarwani. On a practical notion of Geoffrion proper optimality in multicriteria optimization. *optimization-online.org/DB_FILE/2016/08/5584.pdf*, 2016.

[35] C. Tammer. A generalization of Ekeland's variational principle. *Optimization*, 25(2-3):129–141, 1992.

[36] I. Valyi. Approximate solutions of vector optimization problems. *Annual Review in Automatic Programming*, 12:246–250, 1985.

[37] D. J. White. Epsilon efficiency. *J. Optim. Theory Appl.*, 49(2):319–337, 1986.