Approximations for Pareto and Proper Pareto solutions and their KKT conditions

P. Kesarwani 1 · P. K. Shukla 2,3 · J. Dutta 4 · K. Deb 5

Received: 6 April 2020 / Revised: 2 May 2022 / Accepted: 3 May 2022 / Published online: 6 June 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
In this article, we view the Pareto and weak Pareto solutions of the multiobjective optimization by using an approximate version of KKT type conditions. In one of our main results Ekeland’s variational principle for vector-valued maps plays a key role. We also focus on an improved version of Geoffrion proper Pareto solutions and it’s approximation and characterize them through saddle point and KKT type conditions.

Keywords
Convex functions · Locally Lipschitz functions · Multi objective optimisation · Pareto minimum · Proper Pareto minimum · Saddle point

1 Introduction
The importance of multiobjective optimization problems in various applications in engineering, economics, business and management can be hardly overstated. For a wide range of applications in engineering design see, for example, (Deb 2001). In this article, we shall study a multiobjective constrained optimization problem with inequality constraints where all the objectives need to be minimized. There are multiple solution concepts in multiobjective optimization. See for example the monograph of Jahn (2004). We shall however focus on three solutions concepts namely, Pareto minimizers, weak Pareto minimizers and a recently introduced modification of Geoffrion proper minimizers in Shukla et al. (2019).

J. Dutta
jdutta@iitk.ac.in

1 Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur, India
2 Alliance Manchester Business School, Manchester, UK
3 Institute AIFB, Karlsruhe Institute of Technology, Karlsruhe, Germany
4 Department of Economic Sciences, Indian Institute of Technology, Kanpur, India
5 College of Engineering, Michigan State University, Michigan, USA
The image of the set of Pareto minimizers and weak Pareto minimizers in the objective space are often referred to as efficient solutions and weak efficient solutions, respectively. The collection of all efficient solutions is often referred to as the Pareto efficient frontier. Two distinct issues have been studied in this article. The first issue is of approximate Karush–Kuhn–Tucker (KKT) conditions while the second involves a class of approximate proper Pareto solutions and their associated saddle point conditions.

The KKT-type conditions associated with multiobjective optimization have a long history and there is a vast literature. It was first introduced in 1951 by Kuhn and Tucker (1951). It is indeed not possible to list down all there references which study KKT type optimality conditions for multiobjective optimization problems. The readers can see the monograph of Ehrgott (2005), Chankong and Haimes (2008), Eichfelder (2008) and Miettinen (1999) and the references there in for optimality conditions for multiobjective optimization in finite dimensions. For more deeper analysis in infinite dimensions kindly see the monographs by Jahn (2004), Luc (1989), Göpfert et al. (2006) and Khan and Tammer (2016) and the references there in.

It is well known that scalarization techniques play a significant role in the solution of multiobjective optimization problems. For more details see Ehrgott (2005) and Miettinen (1999). In engineering sciences, metaheuristics like evolutionary algorithms are very popular. See for example (Deb 2001). For the past fifteen years there has been a lot of interest in the study of descent methods for multiobjective optimization. See for example the excellent survey of Fukuda and Grana Drummond (2014) which appeared in 2014 and the reference therein. For some recent work on descent methods see for example (Tanabe et al. 2019; Goochi et al. 0000; Bennet and Peitz 2021) and the references therein, just to cite a few. There are several approaches to finding a Pareto minimizer where the solving the associated KKT conditions play a crucial role. Consider for example the recent method due to Martin and Schutze (2018) where in they develop a two-step method or a predictor-corrector method for solving a multiobjective problem with inequality constraints. They call their method the Pareto Tracer. Very recently (Beltran et al. 2020) modified the approach of Martin and Schutze (2018) to solve a constrained multiobjective optimization problem with both equality and inequality constraints. Also look the monograph by Hillermeierer (2001) where the focus is to solve the KKT conditions associated with a multiobjective optimization problem.

On the other hand, KKT conditions for multiobjective optimization and their approximate versions play a major role in determining the quality of solution of a multiobjective problem specially when metaheuristics like evolutionary algorithms are applied. See for example (Deb et al. 2007, 2015). While searching the literature during the revision of the first draft of the paper we found two recent papers which are very close in spirit to the issues that we study in the paper though each of these papers also studies several other issues. The first paper is by Giorgi et al. (2016), where they introduce the Approximate KKT condition or AKKT condition for multiobjective optimization problems. This notion of AKKT conditions was first introduced in scalar optimization by Andreani et al. (2011). The AKKT conditions are essentially sequential in nature, in the sense that we say, AKKT conditions hold at a feasible point of a multiobjective optimization problem with smooth data if there is a sequence

Springer
of vectors converging to that feasible point and also a sequence of multipliers with respect to which some approximate version of the KKT conditions hold at each point of such a sequence. As we will see later in our discussion, that AKKT condition says that the gradients of the scalar Lagrangian of the multiobjective optimization problem evaluated the points of the sequence should converge to zero. It also requires an asymptotic version of the complementary slackness condition. The versions of the approximate KKT condition we study here do not require any such asymptotic requirements though we will try to present the relation between our notions and that of the AKKT conditions. The other paper by Chuong and Kim (2016) where they introduces several notions of quasi-Pareto minimizers and studies optimality conditions and duality for these solution concepts. The duality they study here is a generalization of the notion of Lagrangian duality in scalar optimization and thus are very intimately linked to saddle point conditions. In this article we also focus on the saddle point type conditions for very recently introduced notion of approximate Geoffrion proper solutions with pre-set bound in Shukla et al. (2019). Let us also mention that results in our article are different from the two papers that we just mentioned.

At this point it would be nice to motivate the new notion of proper Pareto optimality whose approximate version we intend to study from the KKT viewpoint. It is important to note that 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 Ehrgott 2005). Very recently, the authors discussed an improved version of Geoffrion proper solutions in Shukla et al. (2019). This solution notion is based on the assumption that the decision maker, in practice, usually looks for those proper Pareto solutions whose trade-off is bounded by a value preset by her/him. The detailed analysis of such solutions and their approximate versions has been carried out in Shukla et al. (2019) and shown to be stable than the standard Geoffrion proper solutions. In the present article, one of our goals is to develop saddle point and KKT type conditions for this class of solutions in the particular setting of a convex vector optimization problem with convex inequality constraints.

The whole paper revolves around answering three questions in which first two questions stem from an attempt to generalize two results, which are on approximate solutions for scalar optimization problems which appeared in Dutta et al. (2013). The first result concerns a scalar optimization problem with locally Lipschitz data [(see Theorem 3.2 in Dutta et al. (2013)] which says that if a sequence of points each satisfying an approximate version of KKT conditions converges to point at which a suitable constraint qualification holds, then that point is a KKT point. Thus, we have the following first question:

- **Q1:** Can a similar kind of result be deduced for multiobjective optimization problem?

Our second question stems from Theorem 3.7 in Dutta et al. (2013) in which a kind of converse result of Theorem 3.2 in Dutta et al. (2013). In Theorem 3.7 in Dutta et al. (2013) the following is proved. A local minimizer of a scalar optimization problem
with locally Lipschitz objective and convex inequality constraints is considered. It is assumed that the Slater condition holds. It is then demonstrated that if we consider a sequence of points converging to the local minimizer then there exists a subsequence which satisfies some type of approximate KKT type conditions at each point close to the local minimizer.

- **Q2:** Does the result of Theorem 3.7 in Dutta et al. (2013) have any analog in multiobjective optimization? Further, do the locally Lipschitz data suffice, or we need more assumptions? Can the convexity assumptions give us better results?

Our third question is associated with the KKT-type conditions for the approximate Geoffrion proper solutions with a preset bound.

- **Q3:** Can we develop an approximate KKT type condition which can completely characterize a Geoffrion proper solutions with a preset bound at least in the convex case? Does the saddle point conditions completely characterize such class of solutions?

In Sect. 2, we present the problem, basic definitions and technical tools from convex and non-smooth analysis required in the article. In Sect. 3, we answer the first two questions raised in this section and Sect. 4 assures the last question by trying to develop the saddle point conditions and approximate KKT type conditions for the improved Geoffrion proper solutions. We end our discussion by concluding remarks in Sect. 5.

## 2 Preliminaries and basic tools

Let \( A \subseteq \mathbb{R}^n \) be a given set, then the closure and interior of a set \( A \) is denoted by \( \text{cl}A \) and \( \text{int}A \) respectively. For vectors \( x, y \in \mathbb{R}^n \) the inner product is given by \( \langle x, y \rangle \) and \( \|x\| \) denotes the Euclidean norm or 2-norm of the vector \( x \). A set \( A \subseteq \mathbb{R}^n \) is a cone, if for each \( a \in A \) and non-negative scalar \( \lambda \), \( \lambda a \in A \). A normal cone of a convex 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:

\[
\begin{align*}
\min f(x) := (f_1(x), \ldots, f_m(x)), \\
\text{subject to } g_j(x) \leq 0, \quad j = 1, 2, \ldots, l.
\end{align*}
\]

where each \( f_i : \mathbb{R}^n \rightarrow \mathbb{R} \) and \( g_j : \mathbb{R}^n \rightarrow \mathbb{R} \). Let us denote the constraints by the set \( X := \{ x \in \mathbb{R}^n : g_j(x) \leq 0, \quad j = 1, 2, \ldots, l \} \subseteq \mathbb{R}^n \), \( I := \{ 1, 2, \ldots, m \} \), \( L := \{ 1, 2, \ldots, l \} \). Our main assumption on the objective and constraint functions would be that they would be locally Lipschitz though need not be differentiable. The special case when the functions are convex will also be discussed. Thus this shows that the feasible set \( X \) is a closed set. As we mentioned earlier, that there are several notions for approximate solutions but in this article, we consider the notion of approximate solution introduced in Loridan (1984). We consider \( \epsilon \in \mathbb{R}_+^m \), i.e., \( \epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_m) \), \( \epsilon_i \geq 0 \) for each \( i \in I \) to formalize our notions. Our focus on this paper is on \( \epsilon \)- solutions of MOP and we provide the relevant definitions below.
Definition 2.1 Given $\epsilon \in \mathbb{R}_+^m$, 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 $\epsilon$-Pareto optimal solution of MOP. Further if there is no $x \in X$ such that $f(x) + \epsilon - f(x^*) \in -\text{int}(\mathbb{R}_+^m)$, then the point $x^*$ is said to be a weak $\epsilon$-Pareto optimal solution of MOP.

An $\epsilon$-Pareto (weak) optimal solution with $\epsilon = 0$ is commonly known as Pareto (weak) optimal solution. There are various notion of approximate solutions other than the ones defined above in the literature (see for example Dutta and Vetrivel 2001; Gutiérrez et al. 2006; Loridan 1984; Vályi 1985 and Gutiérrez et al. 2010).

Definition 2.2 A point $x^*$ is said to be a local 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 $\delta$.

The weak counter part of local solution can be defined in the similar fashion as in Definition 2.2. We want to mention that in several situations we consider the particular form of the vector $\epsilon \in \mathbb{R}_+^m$, given by $\epsilon = \epsilon e$, where $e = (1, 1, \ldots, 1)^T$ and $\epsilon \in \mathbb{R}_+$. In those cases, the solutions referred to as the $\epsilon e$-Pareto and $\epsilon e$-weak Pareto solution respectively. The set of all $\epsilon$-Pareto points is denoted by $S_\epsilon(f, X)$ and the set of all $\epsilon$-weak Pareto points as $S_{\epsilon e}(f, X)$.

Definition 2.3 Given $\epsilon \in \mathbb{R}_+^m$, a point $x_0 \in X$ is called $\epsilon$-Geoffrion proper solution of MOP if $x_0 \in S_\epsilon(f, X)$ 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

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

The upper bound of the trade-off in the above definition is not known beforehand and the definition only assures the existence of such a bound. Further, it is clear form the definition that the trade-off varies as we choose different proper points. The following improved definition introduced in Shukla et al. (2019) eliminates the dependence of the bound on the solution points:

Definition 2.4 Given $\epsilon \in \mathbb{R}_+^m$ and a scalar $\hat{M} > 0$, a point $x_0 \in X$ is called $(\hat{M}, \epsilon)$-Geoffrion proper solution of MOP if $x_0 \in S_\epsilon(f, X)$ 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 as $G_{\hat{M}, \epsilon}(f, X)$. For $\epsilon = 0$, the set of exact $\hat{M}$- Geoffrion proper is denoted by $G_{\hat{M}}(f, X)$. Now we shall present the Ekeland variation principle for vector-valued functions which was introduced in Tammer (1992) when the ordering cone is $\mathbb{R}_+^m$. We first define the
then, there exists (Bazaraa et al. 2013). It is important to note that a point for more details on subdifferentials of finite-valued convex functions see for example 123

\[ f(x) + \rho c_0 - f(x_0) \notin -\text{int}(\mathbb{R}^m_+) \], for all \( x \in U \).

(2.1)

Then, there exists \( \bar{x}_0 = x_0(\rho) \in U \) such that \( \|\bar{x}_0 - x_0\| \leq \sqrt{\rho} \) and for all \( x \in U \setminus \{\bar{x}_0\} \)

1. \( f(x) + \rho c_0 - f(x_0) \notin -\text{int}(\mathbb{R}^m_+) \),
2. \( f(x) + \sqrt{\rho}\|\bar{x}_0 - x\|c_0 - f(\bar{x}_0) \notin -\text{int}(\mathbb{R}^m_+) \).

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. Although 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. 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

\[ \partial (f + g)(x) = \partial f(x) + \partial g(x). \]  

(2.2)

For more details on subdifferentials of finite-valued convex functions see for example (Bazaraa et al. 2013). 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 $\varepsilon$-subdifferential is a relaxed version of the subdifferential which is very useful tool in convex analysis and optimization. We begin with defining the $\varepsilon$-subdifferential of convex function.

**Definition 2.7** Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex function and $\varepsilon \geq 0$. The $\varepsilon$-subdifferential of $f$ at the point $x$ is given as

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

The elements of $\partial_{\varepsilon} f(x)$ are called $\varepsilon$-subgradients 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 $\varepsilon$-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 $\varepsilon$-minimizer of $f$ on $\mathbb{R}^n$ if and only if $0 \in \partial_{\varepsilon} f(x_0)$.

For complete description of properties of $\varepsilon$-subdifferential see (Dhara and Dutta 2011).

Definition 2.7 talks about subdifferential for convex functions. We now discuss subdifferential of a locally Lipschitz function which need not be convex. 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 Lipschitz 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. We now define the Clarke directional derivative of a 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 $0 \in \partial^\circ f(x_0)$ (see Rockafellar 1970 for proof). It also satisfy the sum rule but it gives only one side containment, i.e., for given two locally Lipschitz functions $f$ and $g$, we have

$$\partial^\circ (f + g)(x) \subset \partial^\circ f(x) + \partial^\circ g(x). $$

In fact the inclusion above can be proper. For more details on the Clarke subdifferential see (Clarke 1983).
3 Approximate KKT conditions

In this section, we begin by defining a notion of modified $\varepsilon$-KKT points which is very suitable for the purpose of convex vector optimization problems. This notion is motivated by a similar notion defined in Dutta et al. (2013) for scalar optimization problem and Durea et al. (2011) for convex vector optimization.

Definition 3.1 For an $\varepsilon \in \mathbb{R}_+$, a point $x_0 \in X$ is said to be a modified $\varepsilon$-KKT point of MOP if there exists $x_\varepsilon$ such that $\|x_0 - x_\varepsilon\| \leq \sqrt{\varepsilon}$ and there exists $u_i \in \partial^0 f_i(x_\varepsilon)$ for all $i \in I$, $v_r \in \partial^0 g_r(x_\varepsilon)$ for all $r \in L$, vectors $\lambda \in \mathbb{R}^m_+$ with $\|\lambda\| = 1$ and $\mu \in \mathbb{R}^l_+$ such that

$$\left\| \sum_{i \in L} \lambda_i u_i + \sum_{r \in L} \mu_r v_r \right\| \leq \sqrt{\varepsilon}, \quad \text{and} \quad \sum_{r \in L} \mu_r g_r(x_0) \geq -\varepsilon.$$

Note that a modified $\varepsilon$-KKT point of MOP with $\varepsilon = 0$ is commonly known as a KKT point of MOP. Now we will define two constraint qualifications, Slater constraint qualification (SCQ for short) and Basic constraint qualification (BCQ for short) which will play a significant role in the main result of this article.

Definition 3.2 (Rockafellar and Wets 1998) MOP satisfies Slater constraint qualification if there exists $\hat{x} \in X$ such that $g_r(\hat{x}) < 0$, for all $r \in L$ provided constraint functions $g_r$ for all $r \in L$ are convex.

Definition 3.3 (Rockafellar and Wets 1998) MOP satisfies Basic Constraint Qualification (BCQ) at a point $\bar{x}$ if there exists no $p \in \mathbb{R}_+^l \setminus \{0\}$ such that $0 \in \sum_{r \in L} p_r \partial^0 g_r(\bar{x})$ provided constraint functions $g_r$ for all $r \in L$ are locally Lipschitz.

The next theorem answers the first question (Q1) raised in this article which says that if a sequence $\{x^k\}$ of modified $\varepsilon^k$-KKT points of MOP converges to a point $x_0$ where the Basic Constraint Qualification (BCQ) holds, then $x_0$ is a KKT point. Observe that we do not need convexity of the objective functions to prove the following result whereas we only require local Lipschitz continuity of the objectives and constraints. It is important to note that the similar kind of result has been discussed under convexity assumptions in Durea et al. (2011).

Theorem 3.4 Consider the problem MOP with locally Lipschitz data and let $\{\varepsilon_k\}$ 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 $\varepsilon^k$-KKT point of MOP. Further, assume that the BCQ holds at $x_0$. Then, $x_0$ is a KKT point of MOP.

Proof Note that $x^k$’s are feasible points, i.e., $g_r(x^k) \leq 0$ for all $r \in L$ and $x^k \to x_0$. Hence, using the convexity 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 of MOP. Now, as $x^k$ is a modified $\varepsilon^k$-KKT point, for each $k$, Definition 3.1 gives the existence of a point $x^k$ such that $\|x^k - x_0\| \leq \sqrt{\varepsilon_k}$, the
existence of \( u^k_i \in \partial^o f_i(\hat{x}^k) \) and \( v^k_r \in \partial^o g_r(\hat{x}^k) \) for all \( i \in I \) and \( r \in L \), and the vectors \( \lambda^k \in \mathbb{R}^n \) and \( \mu^k \in \mathbb{R}^l \) with \( \| \lambda^k \| = 1 \) such that

\[
\left\| \sum_{i \in I} \lambda^k_i u^k_i + \sum_{r \in L} \mu^k_r v^k_r \right\| \leq \sqrt{\varepsilon_k}, \quad \text{and} \quad \sum_{r \in L} \mu^k_r g_r(x^k) \geq -\varepsilon_k. \tag{3.1}
\]

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, Eq. (3.1), can be re-written as

\[
\left\| \sum_{i \in I} \frac{\lambda^k_i}{\| \mu^k \|} u^k_i + \sum_{r \in L} \frac{\mu^k_r}{\| \mu^k \|} v^k_r \right\| \leq \frac{1}{\| \mu^k \|} \sqrt{\varepsilon_k}. \tag{3.3}
\]

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

1. As \( \varepsilon_k \) converges to 0, the same holds for \( \frac{1}{\| \mu^k \|} \sqrt{\varepsilon_k} \).
2. Let \( p^k_r = \frac{\mu^k_r}{\| \mu^k \|} \in \mathbb{R}_+ \), 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 \( \hat{p} \in \mathbb{R}_+^l \) with \( \| \hat{p} \| = 1 \). In fact, without loss of generality, we can assume that \( p^k_r \) converges to \( \hat{p}_r \). Hence, for all \( r \in L \)

\[
\frac{\mu^k_r}{\| \mu^k \|} = p^k_r \to \hat{p}_r, \quad \text{as} \quad k \to \infty. \tag{3.4}
\]

3. As \( f_i \)'s are locally Lipschitz functions, their Clarke subdifferential are locally bounded, i.e., for \( x_0 \in X \), there exists \( \delta > 0 \) such that for all \( z \in B_\delta(x_0) \), \( \partial^o f_i(z) \subset K_i \), where, for all \( i \in I \), \( K_i \)'s are bounded sets on \( \mathbb{R}^n \). Since \( x^k \to x_0 \), there exists \( k_0 \in \mathbb{N} \) such that, for all \( k \geq k_0 \), \( x^k \in B_\delta(x_0) \). Therefore, by choosing \( K = \bigcup_{i \in I} \hat{K}_i \) where \( \hat{K}_i = K_i \cup \partial^o f_i(x^1) \cup \ldots \partial^o f_i(x^n) \), we get \( \partial^o f_i(x^k) \subset K \), for all \( i \in I \) and \( k \geq 0 \). Hence, the sequence \( \{ u^k_i \} \), where \( u^k_i \in \partial^o f_i(x^k) \), 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 \( i \in I \),

\[
\frac{-\lambda^k_i}{\| \mu^k \|} u^k_i \to 0, \quad \text{as} \quad k \to \infty. \tag{3.5}
\]

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

\[
v^k_r \to \hat{v}_r, \quad \text{as} \quad k \to \infty. \tag{3.6}
\]

Since each \( \partial^o g_r \) is graph closed set-valued map (see Rockafellar 1970) and \( \hat{x}_k \to x_0 \), one has \( \hat{v}_r \in \partial^o g_r(x_0) \) 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 $\sum_{r \in L} \hat{p}_r \hat{v}_r = 0$, where $\hat{p} \in \mathbb{R}_+^l$ with $\| \hat{p} \| = 1$ and $\hat{v}_r \in \partial^o g_r(x_0)$ 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 $\hat{\lambda} \in \mathbb{R}_+^l$ with $\| \hat{\lambda} \| = 1$, $\hat{\mu} \in \mathbb{R}_+^l$, $\hat{u}_i \in \partial^o f_i(x_0)$ and $\hat{v}_r \in \partial^o g_r(x_0)$.

Now taking $k \to \infty$ in Inequality (3.1), we get $\sum_{i \in I} \hat{\lambda}_i \hat{u}_i + \sum_{r \in L} \hat{\mu}_r \hat{v}_r = 0$. Thus

$$\sum_{i \in I} \hat{\lambda}_i \hat{u}_i + \sum_{r \in L} \hat{\mu}_r \hat{v}_r = 0, \quad \text{where} \quad \hat{\lambda} \in \mathbb{R}_+^m \text{ with } \| \hat{\lambda} \| = 1,$$

$\hat{\mu} \in \mathbb{R}_+^l$, $\hat{u}_i \in \partial^o f_i(x_0)$ and $\hat{v}_r \in \partial^o g_r(x_0).$  \hspace{1cm} (3.7)

Since, $x_0$ is a feasible point of MOP and $\hat{\mu}_r \geq 0$ 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

$$\sum_{r \in L} \hat{\mu}_r g_r(x_0) = 0. \hspace{1cm} (3.8)$$

The Inequalities (3.7) and (3.8) together imply that $x_0$ is a KKT point of MOP. \hfill \Box

The next theorem deals with the second question asked in the article. Basically, $Q_2$ for multiobjective problem can be framed as follows: for every local Pareto points of MOP, does there exists sequence which converges to the point and the sequence has a subsequence which satisfies some type of approximate KKT conditions? We answer this question in Theorem 3.6 for MOP with locally Lipschitz objective functions and with convex inequality constraints and satisfying the Slater constraint qualification. This result shows that we always have a sequence converging to a local Pareto point of MOP with approximate KKT type of conditions which implies that the idea of constructing approximate KKT type conditions is essential in multiobjective optimization. Note that $Q_2$ has not been addressed in Durea et al. (2011). Before we state the theorem, we present the following lemma which will be needed in the proof of the result. This lemma is a special case of Theorem 2.44 in Mordukhovich and Nam (2013).

**Lemma 3.5** Let $A$ and $B$ be two non-empty subsets of $\mathbb{R}^m$. Let $A \cap B \neq \emptyset$ and $\bar{x} \in A \cap B$. Assume that the following qualification condition holds:

$$N_A(\bar{x}) \cap (-N_B(\bar{x})) = \{0\}.$$
Then, \( N_{A \cap B}(\bar{x}) = N_A(\bar{x}) + N_B(\bar{x}) \).

**Theorem 3.6** Consider the problem MOP with locally Lipschitz objectives \( f_i \)'s for all \( i \in I \) and \( g_r \)'s for all \( r \in L \) to be a convex functions which satisfies the Slater constraint qualification. Further, assume that \( x_0 \) is a local Pareto minima and consider \( \{\varepsilon_k\} \) to be a decreasing sequence of positive real numbers converging to \( 0 \). Then, for any sequence \( \{x^k\} \subset X \) of feasible points converging to \( x_0 \) there exists a subsequence \( \{y^k\} \) of \( \{x^k\} \) such that for each \( y^k \), there exists \( \hat{y}^k \) satisfying

1. \( \|y^k - \hat{y}^k\| \leq \sqrt{\varepsilon_k} \),
2. there exists \( u^k_i \in \partial^o f_i(\hat{y}^k) \) and \( v^k_r \in \partial^o g_r(\hat{y}^k) \), for all \( i \in I \) and \( r \in L \), such that

\[
\begin{align*}
\left\| \sum_{i \in I} \lambda^k_i u^k_i + \sum_{r \in L} \mu^k_r v^k_r \right\| & \leq \sqrt{\varepsilon_k}, \\
\sum_{r \in L} \mu^k_r g_r(\hat{y}^k) & = 0,
\end{align*}
\]

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

**Proof** By assumption, \( x_0 \) is a local Pareto minimizer of MOP, i.e., there exists \( \delta > 0 \) such that

\[
f(x) - f(x_0) \notin -\mathbb{R}^m_+ \setminus \{0\}, \text{ for all } x \in V,
\]
equivalently,

\[
f(x) - f(x_0) \in \bar{W}, \text{ for all } x \in V,
\]

where \( \bar{W} := \mathbb{R}^m \setminus (-\mathbb{R}^m_+ \setminus \{0\}) \) and \( V = X \cap B_\delta(x_0) \). 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. Consider any sequence \( \{x^k\} \) in \( X \) and let \( x^k \to x_0 \) as \( k \to \infty \). Thus \( x^k \in B_\delta(x_0) \) for all \( k \) sufficiently large. This shows that \( x^k \in V \) for \( k \) sufficiently large and hence \( \{x_k\} \) is bounded.

We have broken the rest of the 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-Pareto minima of MOP with feasible set as \( V \) where \( e = (1, \ldots, 1)^T \) and \( \varepsilon_k > 0 \). 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 \( N^i_1 \), such that

\[
|f_i(x^k) - f_i(x_0)| < \varepsilon_1, \text{ for all } k \geq N^i_1.
\]

Now choose \( N_1 = \max\{N^1_1, N^2_1, \ldots, N^m_1\} \). Thus, for all \( i \in I \)

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

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

\[
f(x_0) + e\varepsilon_1 - f(y^1) \in \text{int}(\mathbb{R}^m_+).
\]
Note that $\hat{W} + \text{int}(\mathbb{R}^m_+) \subseteq \hat{W}$, hence, (3.11) and (3.13) together gives
\[ f(x) + e\varepsilon_1 - f(y^1) \notin -\mathbb{R}^m_+ \setminus \{0\}, \text{ for all } x \in V. \quad (3.14) \]

Take $\varepsilon_2 < \varepsilon_1$ and 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 $f(x) + \varepsilon_2 e - f(y^2) \notin -\mathbb{R}^m_+ \setminus \{0\}$ for all $x \in V$. Proceeding as above, gives a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that $y^k \in V$ and
\[ f(x) + \varepsilon_k e - f(y^k) \notin -\mathbb{R}^m_+ \setminus \{0\}, \text{ for all } x \in V. \quad (3.15) \]

Hence, $y^k \in V$ is an $\varepsilon_k e$-Pareto minima of MOP with feasible set as $V$. This completes the proof of the first step.

We now come to the second step to complete the proof. Since each $f_i$ is locally Lipschitz, $f$ is $(e, \mathbb{R}^m_+)$-lower semi continuous and $\mathbb{R}^m_+$-bounded below on the convex compact set $V$. Thus, the vector Ekeland Variational Principle (Theorem 2.6) 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 $x \in V \setminus \{y^k\}$,
\begin{enumerate}
  \item $f(x) + \varepsilon_k e - f(\hat{y}^k) \notin -\text{int}(\mathbb{R}^m_+)$, and
  \item $f(x) + \sqrt{\varepsilon_k}\|\hat{y}^k - y^k\|e - f(\hat{y}^k) \notin -\text{int}(\mathbb{R}^m_+)$.\end{enumerate}
Thus from 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
\[ 0 \in \sum_{i \in I} \lambda^k_i \partial^o f_i(\hat{y}^k) + N_V(\hat{y}^k), \]

where $N_V(\hat{y}^k)$ is the normal cone to the set $V$ at $\hat{y}^k$. For proof of above result see for example, page 137 of Chapter 5 in Dutta (2012). Now applying sum rule for the Clarke subdifferential (see Clarke 1983) and using the fact that subdifferential of the norm function at origin is the unit ball, we get
\[ 0 \in \sum_{i \in I} \lambda^k_i \partial^o f_i(\hat{y}^k) + \sqrt{\varepsilon_k}B_1(0) + N_V(\hat{y}^k). \quad (3.16) \]

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 $\hat{y}^k \in B_{\sqrt{\varepsilon_k}}(y^k)$ and $\varepsilon_k \to 0$, for sufficiently large $k$, $B_{\sqrt{\varepsilon_k}}(y^k) \subset B_\delta(x_0)$. Hence, $\hat{y}^k \in B_\delta(x_0)$, for $k$ sufficiently large.

Clearly, $X \cap B_\delta(x_0) \neq \emptyset$. We will now see that the qualification condition for Lemma 3.5 holds in this case. Since $\hat{y}^k \in B_\delta(x_0)$, we see that $\hat{y}^k \in \text{int}B_\delta(x_0)$, thus $N_{B_\delta(x_0)}(\hat{y}^k) = \{0\}$. Hence, $N_V(\hat{y}^k) \cap (-N_{B_\delta(x_0)}(\hat{y}^k)) = \{0\}$. Therefore, using Lemma 3.5, we conclude that
\[ N_V(\hat{y}^k) = N_{X \cap B_2(x_0)}(\hat{y}^k) = N_X(\hat{y}^k) + N_{B_2(x_0)}(\hat{y}^k). \]

Thus \( N_V(\hat{y}^k) = N_X(\hat{y}^k) \). Hence, we can rewrite (3.16) as

\[ 0 \in \sum_{i \in I} \lambda_i \partial f_i(\hat{y}^k) + \sqrt{\epsilon_k} B_1(0) + N_X(\hat{y}^k). \quad (3.17) \]

Further as the Slater constraint qualification holds, using Corollary 23.7.1 of Rockafellar (1970),

\[ N_X(\hat{y}^k) = \left\{ \sum_{r \in L} \lambda_r v_r : v_r \in \partial g_r(\hat{y}^k), \mu_r \geq 0, \mu_r g_r(\hat{y}^k) = 0, r \in L \right\}. \]

Now using the above form of \( N_X(\hat{y}^k) \) and (3.17), it is evident that there exists \( u_i^k \in \partial f_i(\hat{y}^k) \) for all \( i \in I \), \( v_r^k \in \partial g_r(\hat{y}^k) \) for all \( r \in L \) and scalars \( \lambda^k \in \mathbb{R}_+^n \) 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. \( \Box \)

**Remark 3.7** The key fact to note about Theorem 3.6 is that the results hold for any sequence converging to a local Pareto minimizer. So the multipliers in the above result depends on the sequence that we have chosen. Of course in a very naive sense one simple choice of a sequence is \( x_0 = x_k \) for all \( k \in \mathbb{N} \), then naturally the non-smooth KKT conditions hold at that point and hence the approximate KKT conditions also hold. However there can naturally be sequences \( \{x_k\} \) in \( X \) which converges to \( x_0 \) such that \( x_k \neq x_0 \) for all \( k \in \mathbb{N} \). This is of course the case when we run an algorithm for solving (MOP) and generate iterates which converge to a local Pareto minimizer. We would like to mention that to the best of our knowledge that result in Theorem 3.6 is the first of its kind in multiobjective optimization. Further the above result also demonstrates that the use of modified \( \varepsilon \)-KKT conditions can be indeed meaningfully used as a proximity measure or stopping criteria.

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 3.8 and a result from Durea et al. (2011) which will play a key role in proving the Theorem 3.10.

**Lemma 3.8** Consider the problem MOP with each objective functions \( f_i \)'s and constraint function \( g_r \)'s to be convex. Then every local Pareto minima is a global Pareto minima.

**Theorem 3.9** (Theorem 3.6 of Durea et al. (2011)) Let \( x_0 \) be a \( \varepsilon \)-\( \varepsilon \)-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 \( \sigma \)-KKT point where \( \sigma \in (0, \|\varepsilon\| \varepsilon) \).
Theorem 3.10 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 Pareto minima and let the Slater constraint qualification hold. Then, for decreasing sequence of positive real numbers $\{\varepsilon_k\}$ converging to 0, then for any feasible sequence $\{x^k\}$ converging to $x_0$ there exists a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that each $y^k$ is a modified $\sigma_k$-KKT point with $\sigma_k \in (0, \|\varepsilon\|\varepsilon_k]$.

Proof Since the problem data is convex, local Pareto point is global. Now proceed as in the proof of Theorem 3.6 to get a sub-sequence $\{y^k\}$ of $\{x^k\}$ such that $y^k$ is a $\varepsilon_k$-Pareto minima of MOP with feasible set as $V$, where $V = X \cap B_\delta(x_0)$ with $\delta > 0$, i.e., $y^k$ is a local $\varepsilon_k$-Pareto minima of MOP. So, by using the assumption of convexity and Lemma 3.8, we conclude that $y^k$ is a $\varepsilon_k$-Pareto minima of MOP. Now using Theorem 3.9, we conclude that $y^k$ is a modified $\sigma_k$-KKT point with $\sigma_k \in (0, \|\varepsilon\|\varepsilon_k]$.

It will be interesting to seek a relationship between points satisfying the AKKT conditions (Giorgi et al. 2016) and modified $\varepsilon$ KKT points. The AKKT condition introduced in Giorgi et al. (2016) is for a smooth multiobjective optimization problem with both equality and inequality constraints. However, we shall describe the AKKT conditions for our problem MOP with only inequality constraints. For this purpose, we will assume that all functions in MOP are smooth.

Definition 3.11 Let $x_0$ be a feasible point of (MOP), where we we assume that each function $f_i, i = 1, \ldots, m$ and each $g_r, r = 1, \ldots, l$ are smooth functions i.e. they are continuously differentiable. We say that the Approximate KKT or AKKT condition holds at $x_0$ if there exists a sequence $\{x^k\}$ in $\mathbb{R}^n$ and $(\lambda^k, \mu^k) \in \mathbb{R}_+^m \times \mathbb{R}_+^l$ such that

(a) $x^k \to x_0$ as $k \to \infty$.
(b) $\sum_{i \in I} \lambda^k_i \nabla f_i(x^k) + \sum_{r \in L} \mu^k_r \nabla g_r(x^k) \to 0$ as $k \to \infty$.
(c) $\sum_{i \in I} \lambda^k_i = 1$ for all $k$.
(d) $g_r(x_0) < 0$ implies that $\mu^k_r = 0$ for sufficiently large $k$ and $r \in L$.

If we denote the set of active indices by $J(x_0) = \{r : g_r(x_0) = 0\}$, then it is clearly mentioned that the condition (d) is equivalent to the following condition:

$$
\mu^k_r g_r(x^k) \geq 0 \text{ for sufficiently large } k \text{ for all } r \notin J(x_0). \quad (3.18)
$$

Proposition 3.12 Let $x_0$ be a feasible point of MOP with smooth data for which the AKKT condition holds. Then given $\varepsilon > 0$, there exists $x'$ such that $\|x_0 - x'\| \leq \sqrt{\varepsilon}$ and there exists $(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}_+^l$ such that $\lambda \neq 0$, $\|\lambda\| \leq 1$ such that

(i) $\left\| \sum_{i \in I} \lambda_i \nabla f_i(x') + \sum_{r \in L} \mu_r \nabla g_r(x') \right\| \leq \sqrt{\varepsilon}$.
(ii) $\sum_{r \in L} \mu_r g_r(x_0) \geq -\varepsilon$

Proof Form the conditions (a) and (b) of AKKT (Definition 3.11), it is clear that we can find $x^k$ for sufficiently large $k$ such that $\|x_0 - x^k\| \leq \sqrt{\varepsilon}$ and there exists...
\((\lambda^k, \mu^k) \in \mathbb{R}^m_+ \times \mathbb{R}^l_+\) such that
\[
\left\| \sum_{i \in I} \lambda^k_i \nabla f_i(x') + \sum_{r \in L} \mu^k_r \nabla g_r(x') \right\| \leq \sqrt{\epsilon}.
\]

Now form Definition 3.11 (c), we get \(\|\lambda^k\|_1 = 1\), where \(\|\cdot\|_1\) denotes the 1-norm. Hence, \(\lambda^k \neq 0\) and since \(\|\lambda^k\| \leq \|\lambda^k\|_1 = 1\), we have \(\|\lambda^k\| \leq 1\). Further, as we have chosen \(k\) to be sufficiently large from equation (3.18), we get
\[
\sum_{r \notin J(x_0)} \mu^k_r g_r(x^k) \geq 0.
\]

Thus by continuity of \(\sum_{r \notin J(x_0)} \mu^k_r g_r(x)\) at \(x = x_0\), we have, for \(k\) sufficiently large,
\[
\left| \sum_{r \notin J(x_0)} \mu^k_r g_r(x^k) - \sum_{r \notin J(x_0)} \mu^k_r g_r(x_0) \right| \leq \epsilon.
\]

This shows that
\[
\sum_{r \notin J(x_0)} \mu^k_r g_r(x^k) - \epsilon \leq \sum_{r \notin J(x_0)} \mu^k_r g_r(x_0).
\]

By using (3.19) we deduce from the inequality (3.20), that
\[
-\epsilon \leq \sum_{r \notin J(x_0)} \mu^k_r g_r(x_0).
\]

Since \(g_r(x_0) = 0\) for \(r \in J(x_0)\), from above we get \(\sum_{r \in L} \mu^k_r g_r(x_0) \geq -\epsilon\). Hence setting \(x^k = x', \lambda^k = \lambda\) and \(\mu^k = \mu\) the result is established. \(\square\)

From the above proposition it is clear that for a multiobjective optimization problem (MOP) with smooth data a point satisfying the AKKT conditions also satisfies all the requirements to be a modified \(\varepsilon\)-KKT point but does not satisfy the requirement that \(\|\lambda\| = 1\) and hence we cannot say that an AKKT point is a modified \(\varepsilon\)-KKT point. As we were finishing the revision of this paper a very interesting paper due to Eichfelder and Warnow (2021) was published online. In that paper they have also defined the notion of modified \(\varepsilon\)-KKT condition for the smooth version of the MOP problem but have assumed that the multiplier vector associated with the objective function is a unit vector with respect to 1-norm. Thus from the above proposition we can immediately say that if \(x_0\) is an AKKT point for the smooth version of MOP is also a modified \(\varepsilon\)-KKT point in the sense of Eichfelder and Warnow (2021). It is important to note that an earlier version of this paper where the idea of the modified \(\varepsilon\)-KKT point was also defined and uploaded in optimization-online (Kesarwani et al. 2019) was also cited in Eichfelder and Warnow (2021). The interesting part of the work of Eichfelder...
and Warnow (2021) is that they provide numerical examples to show that a slightly tweaked version the modified $\varepsilon$-KKT condition which they call simplified $\varepsilon$-KKT condition is a useful measure of proximity for (MOP) in actual computations and also provides us with an open question. For more details see (Eichfelder and Warnow 2021). Thus the analysis that is carried out here is a first step towards building good proximity measures or stopping criteria for multiobjective optimization problems even with non-smooth data.

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 $(\hat{M}, \varepsilon)$-Geoffrion solutions. Our aim is to present a complete characterization of this class of approximate proper Pareto points in the convex case. We also discuss a scalarization rule for the $(\hat{M}, \varepsilon)$-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 $(\hat{M}, \varepsilon)$-Geoffrion proper points by the system of inequalities which appeared in Shukla et al. (2019). For a given $\varepsilon \in \mathbb{R}^m_+$ and $\hat{M} > 0$, consider $x_0 \in X$, $i \in I$ and define the following system of inequalities ($Q_i(x_0)$) as

\[
\begin{align*}
-f_i(x_0) + f_i(x) + \varepsilon_i &< 0, \\
-f_i(x_0) + f_i(x) + \varepsilon_i &< \hat{M}(f_j(x_0) - f_j(x) - \varepsilon_j), \text{ for all } j \in I \setminus \{i\} \\
x \in X.
\end{align*}
\]

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

The above proposition follows from the definition of the $(\hat{M}, \varepsilon)$-Geoffrion proper solutions, for complete proof, see (Shukla et al. 2019). Before discussing the saddle point conditions for the $(\hat{M}, \varepsilon)$-Geoffrion proper solutions, let us discuss the correspondence between $(M, \varepsilon)$-Geoffrion proper solutions and solution of the weighted sum scalar problem. As mentioned earlier, this correspondence plays a pivotal role to prove main results of this section. To this end, let for $s^* \in \mathbb{R}^m_+$, the weighted sum scalar problem $P(s^*)$ be defined as

\[
\min_{x \in X} \langle s^*, f(x) \rangle.
\]

Theorem 4.2 For a given $\varepsilon \in \mathbb{R}^m_+$ and $\hat{M} > 0$, let $x_0$ be a $(s^*, \varepsilon)$-minimum of $P(s^*)$, where $s^* \in \text{int}(\mathbb{R}^m_+)$. If $\hat{M} \geq (m - 1) \max_{i, j \in I} \frac{\varepsilon_i}{s^*_j}$, then $x_0$ is a $(\hat{M}, \varepsilon)$-Geoffrion proper solution of MOP, i.e., $x_0 \in G_{\hat{M}, \varepsilon}(f, X)$.

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

\[
\begin{align*}
-s_1(x_0) + f_1(x) + \epsilon_1 &< 0, \\
-s_1(x_0) + f_1(x) + \epsilon_1 &< \hat{M}(f_j(x_0) - f_j(x) - \epsilon_j), \quad j \in I \setminus \{1\}
\end{align*}
\]

for all \( x \in X \).

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

\[
s_1^*(-f_1(x_0) + f_1(x) + \epsilon_1) < s_j^*(m - 1)(f_j(x_0) - f_j(x) - \epsilon_j), \quad \text{for all } j \in I \setminus \{1\}.
\]

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

\[
\langle s^*, f(x_0) \rangle - \langle s^*, f(x) \rangle - \langle s^*, \epsilon \rangle > 0.
\]

(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 \( \hat{G}_{M, \epsilon} (f, X) \) 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. Therefore, it is of interest to find the minimum \( M \) such that \( \hat{G}_{M, \epsilon} (f, X) \) 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 \( \hat{G}_2(f, X) \) and \( \hat{G}_1(f, X) \) can be easily computed as follows:

\[
\hat{G}_2(f, X) = \{(0, 0, 1)^\top, (0, 1, 0)^\top, (1, 0, 0)^\top, (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^\top\},
\]

\[
\hat{G}_1(f, X) = \{(0, 0, 1)^\top, (0, 1, 0)^\top, (1, 0, 0)^\top\}.
\]

Moreover, \( \hat{G}_M(f, X) = \emptyset \) 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 \( s^* \in \text{int}(\mathbb{R}^m_+) \) such that \( x_0 \) is a \((s^*, \epsilon)\)-minimum of \( P(s^*) \).

The proof of above theorem can be deduced form Theorem 3 of Liu (1999) which states the above result for \( \epsilon \)-Geoffrion proper solutions of MOP. As every \((\hat{M}, \epsilon)\)-Geoffrion proper solution is a \( \epsilon \)-Geoffrion proper solution of MOP, Theorem 4.4 follows from the aforementioned result.

Remark 4.5 Theorem 4.4 can also be proved by noting the fact that each \((\hat{M}, \epsilon)\)-Geoffrion proper point is \( \epsilon \)-Geoffrion proper point with constant \( \hat{M} \geq 0 \). Hence using Theorem 3.15 form Ehrgott (2005), we can deduce the above result. Now if we denote the set of \((s^*, \epsilon)\)-minimum of \( P(s^*) \) by \( \text{Sol}_\epsilon(P(s^*)) \), 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

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

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

\[
L^\hat{M}_i(x, \tau^i, \mu^i) = f_i(x) + \sum_{j \in I, j \neq i} \tau^i_j \hat{M}_j f_j(x) + \sum_{r \in L} \mu^i_r g_r(x),
\]

where \( \mu^i = (\mu^i_1, \mu^i_2, \ldots, \mu^i_l) \in \mathbb{R}^l_+ \) and \( \tau^i = (\tau^i_1, \tau^i_2, \ldots, \tau^i_m) \in S^m \) 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 \( i \)th-objective Lagrangian problem defined in Chapter 4 of Chankong and Haimes (2008). In Chankong and Haimes (2008), they used the above Lagrangian form as a scalarization scheme of multiobjective problems. In the same spirit as Chankong and Haimes (2008), 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 \( \mu \in \mathbb{R}^m_+ \),

\[
\begin{align*}
(i) & \quad L^\hat{M}_i(x_0, \bar{\tau}^i, \bar{\mu}^i) - \bar{\epsilon}_i \leq L^\hat{M}_i(x_0, \bar{\tau}^i, \mu^i) \leq L^\hat{M}_i(x, \bar{\tau}^i, \bar{\mu}^i) + \bar{\epsilon}_i \\
(ii) & \quad \sum_{r \in L} \bar{\mu}^i_r g_r(x_0) \geq -\bar{\epsilon}_i,
\end{align*}
\]
where \( \bar{\epsilon}_i = \epsilon_i + \sum_{j=1, j \neq i}^m \tau_j^i \hat{\epsilon}_j \). 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 \( x_0 \in G_{M, 2\epsilon} (f, X) \), where \( \hat{M} \geq (1 + \hat{M})(m - 1) \).

**Proof** It is evident from Proposition 4.1 that if \( x_0 \in G_{M, \epsilon} (f, X) \), then for each \( i \in I \), the system \( Q_i (x_0) \), re-written as

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

has no solution, for all \( x \in \mathbb{R}^n \). It is easy to observe that the system \( Q_i (x_0) \) 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 Rockafellar 1970), there exists \( \tau^i = (\tau^i_1, \ldots, \tau^i_m) \in \mathbb{R}^m_+ \) and \( \mu^i = (\mu^i_1, \ldots, \mu^i_l) \in \mathbb{R}^+_l \) with \( (\tau^i, \mu^i) \neq 0 \) such that for all \( x \in \mathbb{R}^n \),

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

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

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

Now, we first claim that \( \tau^i = (\tau^i_1, \ldots, \tau^i_m) \neq 0 \). For if, \( \tau^i = 0 \) then \( \mu^i \neq 0 \) and Inequality \( (4.3) \) reduces to \( \sum_{r \in L} \mu^i_r g_r (x) \geq 0 \), 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 \( \sum_{r \in L} \mu^i_r g_r (x) < 0 \), a contradiction to \( \sum_{r \in L} \mu^i_r g_r (x) \geq 0 \).

Hence, \( \tau^i \neq 0 \) and thus \( \tau^i_j > 0 \). Thus, dividing Inequality \( (4.3) \) by \( \sum_{j \in I} \tau^i_j \), we get

\[
f_i (x) - f_i (x_0) + \epsilon_i + \sum_{j \in I, j \neq i} \left[ \frac{\bar{\tau}^i_j \hat{M} f_j (x) - \bar{\tau}^i_j \hat{M} f_j (x_0) + \bar{\tau}^i_j \hat{M} \epsilon_j}{\tau^i_j} \right] + \sum_{r \in L} \mu^i_r g_r (x) \geq 0, \tag{4.4}
\]

for all \( x \in \mathbb{R}^n \), where \( \bar{\tau}^i_j = \frac{\tau^i_j}{\sum_{j \in I} \tau^i_j} \) and \( \bar{\mu}^i_r = \frac{\mu^i_r}{\sum_{j \in I} \tau^i_j} \). In particular, for \( x = x_0 \), Inequality \( (4.4) \) gives \( \epsilon_i + \sum_{j \in I, j \neq i} \bar{\tau}^i_j \hat{M} \epsilon_j + \sum_{r \in L} \bar{\mu}^i_r g_r (x_0) \geq 0 \). By setting \( \bar{\epsilon}_i = \epsilon_i + \sum_{j \in I, j \neq i} \bar{\tau}^i_j \hat{M} \epsilon_j + \sum_{r \in L} \bar{\mu}^i_r g_r (x_0) \geq 0 \).
\[\epsilon_i + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} \epsilon_j,\]
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 \(x \in \mathbb{R}^n\),
\[f_i(x) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x) + \sum_{r \in L} \bar{\mu}_r^i g_r(x) + \bar{\epsilon}_i \geq f_i(x_0) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x_0).\] (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
\[f_i(x) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x) + \sum_{r \in L} \bar{\mu}_r^i g_r(x) + \bar{\epsilon}_i \geq f_i(x_0) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x_0),\]
which implies that for each \(i \in I\) and for all \(x \in \mathbb{R}^n\),
\[L_i^i(x, \tilde{t}_i^i, \bar{\mu}_i^i) + \bar{\epsilon}_i \geq L_i^i(x_0, \tilde{t}_i^i, \bar{\mu}_i^i).\] (4.6)

Further, from Equation (4.2), we observe that for all \(i \in I\) and any \(\mu \in \mathbb{R}_+^I\)
\[L_i^i(x_0, \tilde{t}_i^i, \mu) \leq f(x_0) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x_0),\]
which can be written as \[L_i^i(x_0, \tilde{t}_i^i, \mu) \leq f(x_0) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x_0) + \sum_{r \in L} \bar{\mu}_r^i g_r(x) + \bar{\epsilon}_i.\] Thus, for all \(x \in \mathbb{R}^n\) and \(\mu \in \mathbb{R}_+^I\),
\[L_i^i(x_0, \tilde{t}_i^i, \mu) \leq L_i^i(x_0, \tilde{t}_i^i, \bar{\mu}_i^i) + \bar{\epsilon}_i.\] (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 \(\tilde{t}_i^i \in S^n\) and \(\bar{\mu}_i^i \in \mathbb{R}_+^I\) such that Conditions (i) and (ii) hold. Our first step is to show that \(x_0\) is feasible to CMOP. As we know from (i), for all \(\mu \in \mathbb{R}_+^I\)
\[L_i^i(x_0, \tilde{t}_i^i, \mu) - \bar{\epsilon}_i \leq L_i^i(x_0, \tilde{t}_i^i, \bar{\mu}_i^i).\]
Thus, \[f_i(x_0) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x_0) + \sum_{r \in L} \bar{\mu}_r^i g_r(x_0) - \bar{\epsilon}_i \leq f_i(x_0) + \sum_{j \in I, j \neq i} \tilde{t}_j^i \hat{M} f_j(x_0).\] This shows that for all \(\mu \in \mathbb{R}_+^I\),
\[\sum_{r \in L} \bar{\mu}_r^i g_r(x_0) \leq \bar{\epsilon}_i.\] (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 \( x \in \mathbb{R}^n \)

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

(4.9)

which implies

\[
\begin{align*}
&f_i(x) + \sum_{j \in I, j \neq i} \bar{\tau}_j \hat{M}_j f_j(x) + \sum_{r \in L} \bar{\mu}_r g_r(x) + \epsilon_i + \sum_{j \in I, j \neq i} \bar{\epsilon}_j \hat{M}_j \epsilon_j \geq f_i(x_0) \\
&+ \sum_{j \in I, j \neq i} \bar{\tau}_j \hat{M}_j f_j(x_0) + \sum_{r \in L} \bar{\mu}_r g_r(x_0).
\end{align*}
\]  

(4.10)

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

\[
\begin{align*}
&f_i(x) + \sum_{j \in I, j \neq i} \bar{\tau}_j \hat{M}_j f_j(x) + \epsilon_i + \sum_{j \in I, j \neq i} \bar{\epsilon}_j \hat{M}_j \epsilon_j \geq f_i(x_0) \\
&+ \sum_{j \in I, j \neq i} \bar{\tau}_j \hat{M}_j f_j(x_0) + \sum_{r \in L} \bar{\mu}_r g_r(x_0).
\end{align*}
\]  

(4.10)

Using Condition (ii), we have

\[
\begin{align*}
&f_i(x) + \sum_{j \in I, j \neq i} \bar{\tau}_j \hat{M}_j f_j(x) + \epsilon_i + \sum_{j \in I, j \neq i} \bar{\epsilon}_j \hat{M}_j \epsilon_j \geq f_i(x_0) \\
&+ \sum_{j \in I, j \neq i} \bar{\tau}_j \hat{M}_j f_j(x_0) - (\epsilon_i + \sum_{j \in I, j \neq i} \bar{\epsilon}_j \hat{M}_j \epsilon_j).
\end{align*}
\]  

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

\[
\begin{align*}
\sum_{i \in I} (1 + \hat{M} \sum_{j \in I, j \neq i} \bar{\tau}_j) f_i(x) + \sum_{i \in I} (1 + \hat{M} \sum_{j \in I, j \neq i} \bar{\epsilon}_j)(2\epsilon_j) \\
\geq \sum_{i \in I} (1 + \hat{M} \sum_{j \in I, j \neq i} \bar{\tau}_j) f_i(x_0).
\end{align*}
\]
Hence, \( x_0 \) is \( (s, 2\epsilon) \)-minimizer of \( P(s) \), where \( s = (s_1, \ldots, s_m) \) with \( s_i = 1 + \hat{M} \sum_{k \in I, k \neq i} \bar{\tau}_k^i \), for \( i \in I \). Now since \( \bar{\tau}^i \in S^m \) for all \( i \), we have for all \( i, j \in I \)

\[
\frac{s_i}{s_j} = \frac{1 + \hat{M} \sum_{k \in I, k \neq i} \bar{\tau}_k^i}{1 + \hat{M} \sum_{k \in I, k \neq j} \bar{\tau}_k^j} = \frac{1 + \hat{M} (1 - \bar{\tau}_i^i)}{1 + \hat{M} (1 - \bar{\tau}_j^j)} \leq 1 + \hat{M}.
\]

Since the above inequality is true for every \( i \) and \( j \), we have \( \max_{i,j} \{ \frac{s_i}{s_j} \} \leq 1 + \hat{M} \).

Now consider \( \hat{M} \geq (1 + \hat{M})(m - 1) \) and using Theorem 4.2, we conclude that \( x_0 \in G_{\hat{M}, 2\epsilon}(f, X) \). This completes the proof. \( \square \)

**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

\[
\text{Remark 4.7} \quad \text{for each } \theta_i = \arg \min_{\mu_i} \bar{\tau}_i \text{ implies that } \epsilon_i \in \theta_i \text{ and using Theorem 4.2, we conclude that } x_0 \in G_{\hat{M}, 2\epsilon}(f, X). \text{ This completes the proof.} \]

In order to prove the above statement, note that \( \bar{\epsilon}_i = \epsilon_i + \sum_{j=1, j \neq i} \bar{\tau}_j^i \hat{M} \epsilon_j \). 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 \( x_0 \in G_{\hat{M}, 2\epsilon}(f, X) \), where \( \hat{M} \geq (1 + \hat{M})(m - 1) \). 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 \( \bar{\tau}_j^i \). 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 \( \epsilon \)-subdifferentials, however this rule will be quite different. Observe that if \( x_0 \in G_{\hat{M}, \epsilon}(f, X) \), 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 \( x \in \mathbb{R}^n \)

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

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

\[
0 \in \sum_{i \in I} \partial \bar{\epsilon}_i L_i^\hat{M}(x_0, \bar{\tau}^i, \bar{\mu}^i) \text{ with } \sum_{r \in L} \bar{\mu}^i g_r(x_0) \geq -\bar{\epsilon}_i. \tag{4.11}
\]
Theorem 4.8 For a given $\epsilon \in \mathbb{R}_+^m$ and $\hat{M} > 0$, let us consider the problem CMOP. If $x_0 \in G_{\hat{M}, \epsilon}(f, X)$, then there exist vectors $\tau^i \in S^m$ and $\mu^i \in \mathbb{R}_+^l$, $i \in I$ such that

(A) $0 \in \sum_{i \in I} \partial \hat{v}^i L_i^\hat{M}(x_0, \tau^i, \mu^i),$

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

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

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 $\tilde{v}^i \in \partial \hat{v}^i L_i^\hat{M}(x_0, \tau^i, \mu^i)$ for all $i \in I$ such that

$$0 = \tilde{v}^1 + \tilde{v}^2 + \ldots + \tilde{v}^m. \quad (4.12)$$

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

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

Hence,

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

Now using Equation (4.12), we get

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

$$+ \sum_{j \in I, j \neq i} \tau^i_j \hat{M} f_j(x_0) + \sum_{r \in L} \mu^i_r g_r(x_0)) \geq - \sum_{i \in I} \tilde{\epsilon}_i.$$

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

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

which can be rewritten as

$$\sum_{i \in I} (1 + \sum_{i \in I, i \neq j} \tau^i_j \hat{M}) f_i(x) \geq \sum_{i \in I} (1 + \sum_{i \in I, i \neq j} \tau^i_j \hat{M}) f_i(x_0)$$

$$- \sum_{i \in I} (1 + \sum_{i \in I, i \neq j} \tau^i_j \hat{M}) 2\tilde{\epsilon}_i.$$

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

5 Concluding remarks

To analyze the behaviour of an optimization problem from the viewpoint of KKT conditions is deep-rooted in psyche of researchers in optimization theory. Further, when it comes developing stopping criteria of algorithms or checking the quality of solutions KKT conditions and their approximations can play a pivotal role. In this article, we in fact study a type of approximate KKT conditions called modified \( \epsilon \)-KKT type conditions for Pareto and Weak Pareto minimizers. In fact in the convex case, we achieve a complete characterization. For example, Theorem 3.6 demonstrates that a sequence of points which converge to a Pareto minimizer, for the convex case, has a subsequence where each point satisfies a modified \( \epsilon \)-KKT type conditions. This result thus demonstrates the reason why the modified \( \epsilon \)-KKT type conditions could be first step towards building robust proximity measures for multiobjective optimization problems. The following questions can be taken up for future research regarding approximate KKT conditions in multiobjective optimization.

1. Can the notion of AKKT conditions be generalized for the nonsmooth locally Lipschitz case. Can the results known for the smooth case for AKKT conditions be generalized to the non-smooth settings?.

2. It has been mentioned as an open question in Eichfelder and Warnow (2021), which results for the modified \( \epsilon \)-KKT conditions can be actually carried over to the simplified \( \epsilon \)-KKT condition. This would be interesting since one of the fundamental ways that these two conditions differ is that in the simplified \( \epsilon \)-KKT condition the reference point is used rather that a point in the neighborhood in the case of the modified \( \epsilon \)-KKT conditions. It would be interesting to see if the idea of simplified \( \epsilon \)-KKT conditions can be generalized to the case where the problem MOP has nonsmooth convex functions as objective and constraints. It will be interesting to see whether in such a case the simplified \( \epsilon \)-KKT conditions can be used in actual computations as a proximity measure.

We would also like to mention in the definition of the modified \( \epsilon \)-KKT point we can replace the Clarke subdifferential by the limiting subdifferential. The limiting subdifferential is also mentioned in the current literature as Morudukhovich subdifferential. For more details on the limiting subdifferential see for example (Mordukhovich 2006). Consequently the results in Theorem 3.4 and Theorem 3.6 remains valid if the Clarke subdifferential is replaced with the limiting subdifferential.

The analysis of the approximate versions of the \( \hat{M} \)-Geoffrion 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 using these optimality conditions in algorithms that are designed to approximate the set of efficient solutions and have been successfully applied to real-world multiobjective problems (Braun et al.
Another interesting study would be to extend these results to vector optimization problem with general ordering structures (Eichfelder 2014; Shukla and Braun 2013).

Acknowledgements The authors are very grateful to the anonymous referees for their constructive comments which has greatly improved the presentation of the paper. We are also very thankful to the Associate Editor for bringing to our notice the reference Beltran et al. (2020).

References

Andreani R, Haeser G, Martínez JM (2011) On sequential optimality conditions for smooth constrained optimization. Optimization 60(5):627–641
Bazaraa MS, Sherali HD, Shetty CM (2013) Nonlinear programming: theory and algorithms. Wiley, New York
Beltran F, Cuate O, Schutze O (2020) The Pareto tracer for general inequality constrained multi-objective optimization problems. Math Comput Appl 25:80. https://doi.org/10.3390/mca25040080
Bennet G, Peitz S (2021) An efficient descent method for locally Lipschitz multiobjective problem. J Optim Th Appl 188:696–723
Braun M, Seijo S, Echanobe J, Shukla PK, Campo I, Garcia-Sedano J, Schmeck H (2016) A neuro-genetic approach for modeling and optimizing a complex cogeneration process. Appl Soft Comput 48:347–358
Braun M, Shukla PK, Schmeck H (2015) Obtaining optimal Pareto front approximations using scalarized preference information. Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, pp 631–638
Chankong V, Haines YY (2008) Multiobjective decision making: theory and methodology. Courier Dover Publications, New York
Chuong TD, Kim DS (2016) Approximate solutions of multiobjective optimization problems. Positivity 20(1):187–207
Clarke FH (1983) Optimization and nonsmooth analysis, vol 5. Wiley, New York (Republished by SIAM 1990)
Deb K (2001) Multi-objective optimization using evolutionary algorithms. Wiley, New York
Deb K, Tewari R, Dixit M, Dutta J (2007) Finding trade-off solutions close to KKT points using evolutionary multiobjective optimization. Proceedings of the 2007 IEEE Congress on Evolutionary Computation, pp 2109–2116
Deb K, Abouhawwash M, Dutta J (2015) An optimality theory based proximity measure for evolutionary multiobjective and many objective optimization. Lecture Notes Comput Sci 9019:18–33
Dhara A, Dutta J (2011) Optimality conditions in convex optimization: a finite-dimensional view. CRC Press, Cambridge
Durea M, Dutta J, Tammer C (2011) Stability properties of KKT points in vector optimization. Optimization 60(7):823–838
Dutta J (2012) Strong KKT, second order conditions and non-solid cones in vector optimization. In: Recent Developments in Vector Optimization, pp 127–167. Springer
Dutta J, Deb K, Tulsyan R, Arora R (2013) Approximate KKT points and a proximity measure for termination. J Global Optim 56(4):1463–1499
Dutta J, Vetrivel V (2001) On approximate minima in vector optimization. Numer Funct Anal Optim 22(7–8):845–859
Ehrgott M (2005) Multicriteria optimization, 2nd edn. Springer, Berlin
Eichfelder G (2008) Adaptive scalarization methods in multiobjective optimization, vol 436. Springer, Berlin
Eichfelder G (2014) Variable ordering structures in vector optimization. Springer, Berlin
Eichfelder G, Warnow L (2021) Proximity measures based on KKT points for constrained multi-objective optimization. J Global Optim 80:63–86
Fukuda E, H, Grana Drummond LM (2014) A survey on multiobjective descent methods. Pesquisa Operacional 34(3):585–620
Giorgi G, Jiménez B, Novo V (2016) Approximate Karush–Kuhn–Tucker condition in multiobjective optimization. J Optim Theory Appl 171(1):70–89
Göpfert Alfred, Riahi Hassan, Tammer Christiane, Zalinescu Constantin (2006) Variational methods in partially ordered spaces. Springer, Berlin

Goochi G, Liuzzi G, Luidi S, Sciandrone M (2020) On the convergence of steepest descent methods for multiobjective optimization. Comput Optim Appl. https://doi.org/10.1007/s10589-020-00192-0

Gutiérrez C, Jiménez B, Novo V (2006) On approximate efficiency in multiobjective programming. Math Methods Oper Res 64(1):165–185

Gutiérrez C, Jiménez B, Novo V (2010) Optimality conditions via scalarization for a new $\varepsilon$-efficiency concept in vector optimization problems. Eur J Oper Res 201(1):11–22

Hillermeirer C (2001) Nonlinear multiobjective optimization-a generalized homotopy approach. Birkhauser

Jahn J (2004) Vector optimization?: theory, applications, and extensions. Springer, Berlin

Khan AA, Tammer C, Zalinescu C (2016) Set-valued optimization. Springer, Berlin

Kesarwani P, Shukla PK, Dutta J, Deb K (2019) Approximations for Pareto and Proper Pareto Solutions and their KKT conditions. Optimization Online, http://www.optimization-online.org/DB_HTML/2018/10/6845.html

Kuhn HW, Tucker AW (1951) Nonlinear Programming. Proceedings of the 2nd Berkeley symposium on Mathematical Statistics

Liu JC (1999) $\varepsilon$-properly efficient solutions to nondifferentiable multiobjective programming problems. Appl Math Lett 12(6):109–113

Loridan P (1984) $\varepsilon$-solutions in vector minimization problems. J Optim Theory Appl 43(2):265–276

Luc DT (1989) Theory of vector optimization. Springer, Berlin

Martin A, Schutze O (2018) Pareto Tracer: a predictor-corrector method for multiobjective optimization problem. Eng Optim 50(3):516–536

Miettinen K (1999) Nonlinear Programming. Kluwer Academic Publishers, Norwell

Mordukhovich BS, Nam NM (2013) An easy path to convex analysis and applications, vol 6. Morgan & Claypool Publishers, San Rafael

Mordukhovich Boris S (2006) Variational analysis and generalized differentiation I: Basic theory, vol 330. Springer, Berlin

Rockafellar RT (1970) Convex analysis. Princeton University Press, Oxford

Rockafellar RT, Wets RJB (1998) Variational analysis, vol 317. Springer, Berlin

Shukla PK, Dutta J, Deb K, Kesarwani P (2019) On a practical notion of Geoffrion proper optimality in multicriteria optimization. Optimization, pp 1–27

Shukla PK, Braun M (2013) Indicator based search in variable orderings: theory and algorithms. Lect Notes Comput Sci 7811:66–80

Tammer C (1992) A generalization of Ekeland’s variational principle. Optimization 25(2–3):129–141

Tanabe H, Fukuda EH, Yamashita N (2019) Proximal gradient methods for multiobjective optimization and application. Comput Optim Appl 72(2):339–361

Valyi I (1985) Approximate solutions of vector optimization problems. Annu Rev Autom Program 12:246–250

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