Karush-Kuhn-Tucker Proximity Measure for Multi-Objective Optimization Based on Numerical Gradients

Mohamed Abouhawwash
Department of Mathematics, Faculty of Science, Mansoura University
Mansoura 35516, Egypt
saleh1284@mans.edu.eg

Kalyanmoy Deb*
Department of Electrical and Computer Engineering, Michigan State University
East Lansing, MI 48824, USA
kdeb@egr.msu.edu

ABSTRACT

A measure for estimating the convergence characteristics of a set of non-dominated points obtained by a multi-objective optimization algorithm was developed recently. The idea of the measure was developed based on the Karush-Kuhn-Tucker (KKT) optimality conditions which require the gradients of objective and constraint functions. In this paper, we extend the scope of the proposed KKT proximity measure by computing gradients numerically and evaluating the accuracy of the numerically computed KKT proximity measure with the same computed using the exact gradient computation. The results are encouraging and open up the possibility of using the proposed KKTPM to non-differentiable problems as well.

Keywords

Numerical gradients; KKT proximity measure; Evolutionary algorithms

1. INTRODUCTION

Karush-Kuhn-Tucker (KKT) optimality conditions are necessary for a solution to be optimal for single or multi-objective optimization problems [21, 22, 3]. These conditions require first-order derivatives of objective function(s) and constraint functions, although extensions of them for handling non-smooth problems using subdifferentials exist [2, 22]. Although KKT conditions are guaranteed to be satisfied at optimal solutions, the mathematical optimization literature is mostly silent about the regularity in violation of these conditions in the vicinity of optimal solutions. This lack of literature prohibits optimization algorithmists working particularly with approximation based algorithms that attempt to find a near-optimal solution at best to check or evaluate the closeness of a solution to the theoretical optimal solution using KKT optimality conditions.

*http://www.coin-laboratory.com

GECCO '16, July 20-24, 2016, Denver, CO, USA
© 2016 ACM. ISBN 978-1-4503-4206-3/16/07…$15.00
DOI: http://dx.doi.org/10.1145/2908812.2908893

However, recent studies have suggested interesting new definitions of neighboring and approximate KKT points [14, 16, 1], which have been resulted in a KKT proximity measure [23, 14]. The study has clearly shown that a naive measure of the extent of violation of KKT optimality conditions cannot provide a proximity 'measure' to the theoretical optimum, however, the KKT proximity measure (KKTPM) that uses the approximate KKT point definition makes a 'measure' of closeness of a solution from the KKT point. This is a remarkable achievement in its own right, as KKTPM allows a way to know the proximity of a solution from a KKT point without actually knowing the location of the KKT point. A recent study has extended the KKTPM procedure developed for single-objective optimization problems to be applied to multi-objective optimization problems [9, 8]. Since in a multi-objective optimization problem with conflicting objectives, multiple Pareto-optimal solutions exist, KKTPM procedure can be applied to each obtained solution to obtain a proximity measure of the solution from its *nearest* Pareto-optimal solution.

The KKTPM computation opens up several new and promising avenues for further development of evolutionary single [15] and multi-objective optimization algorithms [6, 5]: (i) KKTPM can be used for a reliable convergence measure for terminating a run, (ii) KKTPM value can be used to identify poorly converged non-dominated solutions in a multi-objective optimization problem, (iii) KKTPM-based local search procedure can be invoked in an algorithm to speed up the convergence procedure, and (iv) dynamics of KKTPM variation from start to end of an optimization run can provide useful information about multi-modality and other attractors in the search space.

In this paper, we extend the scope of KKTPM by evaluating its accuracy with gradients computed numerically. If a good enough accuracy is achieved, the computation of exact gradient is not mandatory for KKTPM to be applied to single and multi-objective problems. For this purpose, standard EMO results on a number of multi and many-objective optimization problems are evaluated using numerical and exact gradient based KKTPM values.

In the remainder of the paper, we first discuss the KKTPM computing methods in Section 2. We describe the principle of KKT proximity measure for multi-objective optimization problems based on the *achievement scalarizing function (ASF)* concept proposed in the multiple criterion decision making (MCDM) literature. Also, we discuss three fast yet approximate methods for computing the KKT proximity measure without using any explicit optimization procedure. Section 3 proposed numerical gradient method based on central difference method for calculating the objective and constraint differentiation instead of exact differentiation. Section 4 compares the proposed numerical gradient method with previously-published exact and approximate KKTPM values on standard multi- and many-objective optimization problems including a few engineering design problems. Finally, conclusions are made in Section 5.

2. A BRIEF REVIEW OF KKTPM

In this section, we first summarize the KKTPM computation procedure based on exact gradients of objective and constraint functions, suggested in an earlier study [8]. Thereafter, we discuss a computationally fast yet approximate KKTPM computational procedure which has been recently proposed [7].

2.1 Exact KKT Proximity Measure

Authors [8] used the definition of an approximate KKT solution to suggest a KKT proximity measure (KKTPM) for any iterate (solution), $\mathbf{x}^k \in R^n$ for an M-objective optimization problem of the following type:

$$\begin{array}{rl} \text{Minimize}_{(\mathbf{x})} & \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x})\}, \\ \text{Subject to} & g_j(\mathbf{x}) \leq 0, \quad j = 1, 2, \ldots, J. \end{array} \tag{1}$$

For a given iterate $\mathbf{x}^k$, they formulated an *achievement scalarization function* (ASF) optimization problem [24]:

$$\begin{array}{rl} \text{Minimize}_{(\mathbf{x})} & \text{ASF}(\mathbf{x}, \mathbf{z}, \mathbf{w}) = \max_{m=1}^{M} \left(\frac{f_m(\mathbf{x}) - z_m}{w_m}\right), \\ \text{Subject to} & g_j(\mathbf{x}) \leq 0, \quad j = 1, 2, \ldots, J. \end{array} \tag{2}$$

To have a worse KKTPM value for weak Pareto-optimal solutions, an augmented version of the ASF [19] is employed in the original study. The reference point $\mathbf{z} \in R^M$ in the objective space was then considered as an utopian point and each component of the weight vector $\mathbf{w} \in R^M$ is set as follows:

$$w_i = \frac{f_i(\mathbf{x}^k) - z_i}{\sqrt{\sum_{m=1}^{M} (f_m(\mathbf{x}^k) - z_m)^2}}. \tag{3}$$

Thereafter, the KKT proximity measure was computed by extending the procedure developed for single-objective problems elsewhere [14]. Since the ASF formulation makes the objective function non-differentiable, a smooth transformation was first made by introducing a slack variable x_{n+1} and reformulating the optimization problem (equation 2) as follows:

$$\begin{array}{rl} \text{Minimize}_{(\mathbf{x}, x_{n+1})} & F(\mathbf{x}, x_{n+1}) = x_{n+1}, \\ \text{Subject to} & \left(\frac{f_i(\mathbf{x}) - z_i}{w_i^k}\right) - x_{n+1} \leq 0, \; i = 1, \ldots, M, \\ & g_j(\mathbf{x}) \leq 0, \; j = 1, 2, \ldots, J. \end{array} \tag{4}$$

The KKTPM optimization problem for the above smooth objective function was rewritten for an extended variable set $\mathbf{y} = (\mathbf{x}; x_{n+1})$, as follows:

$$\begin{array}{rl} \text{Minimize}_{(\epsilon_k, x_{n+1}, \mathbf{u})} & \epsilon_k + \sum_{j=1}^{J} \left(u_{M+j} g_j(\mathbf{x}^k)\right)^2, \\ \text{Subject to} & \|\nabla F(\mathbf{y}) + \sum_{j=1}^{M+J} u_j \nabla G_j(\mathbf{y})\|^2 \leq \epsilon_k, \\ & \sum_{j=1}^{M+J} u_j G_j(\mathbf{y}) \geq -\epsilon_k, \\ & \left(\frac{f_j(\mathbf{x}) - z_j}{w_j^k}\right) - x_{n+1} \leq 0, \quad j = 1, \ldots, M, \\ & u_j \geq 0, \quad j = 1, 2, \ldots, (M+J). \end{array} \tag{5}$$

In the above problem, the constraints $G_j(\mathbf{y})$ are given as follows:

$$G_j(\mathbf{y}) = \left(\frac{f_j(\mathbf{x}) - z_j}{w_j^k}\right) - x_{n+1} \leq 0, \quad j = 1, 2, \ldots, M, \tag{6}$$

$$G_{M+j}(\mathbf{y}) = g_j(\mathbf{x}) \leq 0, \quad j = 1, 2, \ldots, J. \tag{7}$$

The value of ϵ_k^* at the optimal point of the above problem (Equation 5) corresponds to the exact KKTPM value. It was proven in the follow-up study [7] that for any feasible solution $\mathbf{x}^k$, $\epsilon_k^* \leq 1$. Based on this fact, the exact KKTPM was defined as follows:

$$\text{Exact KKT Proximity Measure}(\mathbf{x}^k) = \begin{cases} \epsilon_k^*, & \text{if } \mathbf{x}^k \text{ is feasible}, \\ 1 + \sum_{j=1}^{J} \left\langle g_j(\mathbf{x}^k) \right\rangle^2, & \text{otherwise}. \end{cases} \tag{8}$$

On a number of multi and many-objective optimization problems, the above exact KKTPM was able to provide useful properties [8]:

1. In general, a solution closer to the true Pareto-optimal front has a better KKTPM value, thereby providing an almost monotonic characteristic of the KKTPM surface on the objective space.
2. Every Pareto-optimal solution has the smallest KKTPM value of zero.
3. For every feasible solution, the KKTPM value has a value smaller than or equal to one.
4. Weak Pareto-optimal solutions have worse KKTPM values than non-weak Pareto-optimal solutions.
5. Non-dominated solutions parallel to the true Pareto-optimal front in the objective space have similar KKTPM values.
6. Above properties have allowed KKTPM to be used as a termination criterion for set-based multi-objective optimization algorithms, such as EMO methods.

2.2 Approximate KKTPM Computation Method

A flip side of the computational procedure of KKTPM is that it involves solution of an optimization problem stated in Equation 5 at every trade-off solution. However, the resulting optimization problem is quadratic and specialized methodologies such as quadratic programming method can be employed to solve it quickly. A later study [7] proposed an optimization-less approximate procedure for estimating KKTPM values, which we briefly discuss here.

There are three constraint sets to the optimization task stated in Equation 5. The first constraint requires the computation of the gradients of F and G functions and can be rewritten as:

$$\epsilon_k \geq \left\| \sum_{j=1}^{M} \frac{u_j}{w_j^k} \nabla f_j(\mathbf{x}^k) + \sum_{j=1}^{J} u_{M+j} \nabla g_j(\mathbf{x}^k) \right\|^2 + \left(1 - \sum_{j=1}^{M} u_j\right)^2. \tag{9}$$

The second constraint can be rewritten as follows:

$$\epsilon_k \geq -\sum_{j=1}^{M} u_j \left(\frac{f_j(\mathbf{x}^k) - z_j}{w_j^k} - x_{n+1}\right) - \sum_{j=1}^{J} u_{M+j} g_j(\mathbf{x}^k). \tag{10}$$

Since x_{n+1} is a slack variable, the third constraint will be satisfied by setting

$$x_{n+1} = \max_{j=1}^{M} \left(\frac{f_j(\mathbf{x}^k) - z_j}{w_j^k} \right). \quad (11)$$

The first approximation proposed in [7] (referred here as the 'Direct' method) ignored the second constraint given in Equation 10 and only the first (quadratic) constraint is used to find the KKTPM value. This is explained in Figure 1. In

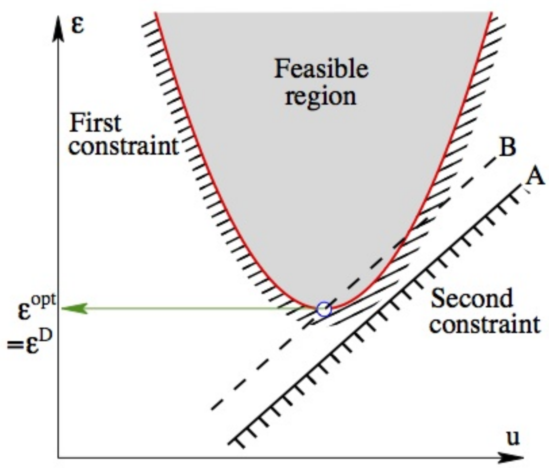


Figure 1: Second constraint governs at iterate $\mathbf{x}_k$ which is not a KKT point.

this case, the second constraint gets automatically satisfied at a point where the first constraint is feasible. The respective optimization problem is equivalent to solving a set of linear system of equations and the process gives rise to the following approximate KKTPM value:

$$\epsilon_k^D = 1 - 1_{M\times 1}^T \mathbf{u}_M^D - \left(\mathcal{G}^T \mathbf{u}_J^D\right)^2. \quad (12)$$

Since $u_m^D \geq 0$, $u_j^D \geq 0$, and $\mathcal{G}\mathcal{G}^T$ is a matrix with positive elements, it can be concluded that $\epsilon_k^D \leq 1$ for any feasible iterate $\mathbf{x}^k$. The previous study [7] also observed that this scenario happens only when the following condition is true at $(\mathbf{u}_M^D, \mathbf{u}_J^D)$-vector:

$$1 - 1_{M\times 1}^T \mathbf{u}_M^D - \left(\mathcal{G}^T \mathbf{u}_J^D\right)^2 \geq -\mathcal{G}^T \mathbf{u}_J, \quad (13)$$

$$\text{or,} \quad 1_{M\times 1}^T \mathbf{u}_M^D - \left(\mathcal{G}^T \mathbf{u}_J^D\right)\left(1 - \mathcal{G}^T \mathbf{u}_J^D\right) \leq 1. \quad (14)$$

For a more generic scenario, illustrated in Figure 2, in which the condition 14 is not satisfied for an iterate $\mathbf{x}^k$, $\epsilon_k^* \neq \epsilon_k^D$ and in fact $\epsilon_k^* > \epsilon_k^D$. Authors proposed approximate values of ϵ_k^* by using three computationally fast approaches, which we discuss in the following subsections.

2.2.1 Adjusted KKTPM Computation Method

From the direct solution $\mathbf{u}^D$ and corresponding ϵ_k^D (point 'D' in the Figure 2), we compute an *adjusted* point 'A' (marked in the figure) by simply computing the ϵ_k value from the second constraint boundary at $\mathbf{u} = \mathbf{u}^D$, as follows:

$$\epsilon_k^{\text{Adj}} = -\mathcal{G}^T \mathbf{u}_J^D. \quad (15)$$

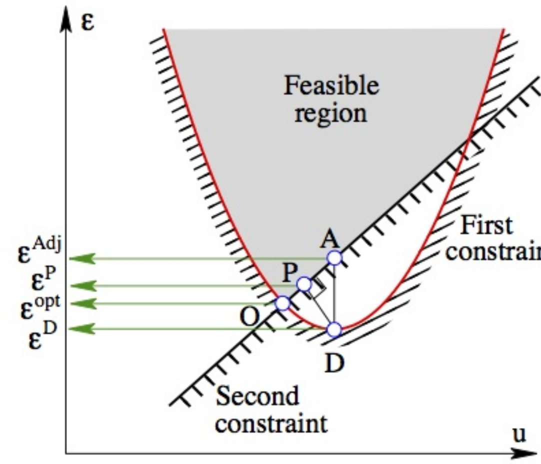


Figure 2: Second constraint governs at iterate $\mathbf{x}_k$ which is not a KKT point.

2.2.2 Projected KKTPM Computation Method

Next, we consider another approximation method using $\mathbf{u}_j^D$. This time, we make a projection from the direct solution (point 'D') $(\mathbf{u}^D, \epsilon_k^D)$ on the second (linear) constraint boundary and obtain the *projected* KKTPM value (for point P), as follows:

$$\epsilon_k^P = \frac{\mathcal{G}^T \left(\epsilon_k^D \mathcal{G} - \mathbf{u}_j^D\right)}{1 + \mathcal{G}^T \mathcal{G}}. \quad (16)$$

2.2.3 Estimated KKTPM Computation Method

After calculating above approximate KKTPM values on many test problems and on a number of engineering design problems, the authors have suggested an aggregate KKTPM value by averaging them and referring to it as the *estimated* KKTPM, as follows:

$$\epsilon^{\text{est}} = \frac{1}{3}\left(\epsilon_k^D + \epsilon_k^P + \epsilon_k^{\text{Adj}}\right). \quad (17)$$

On a number of multi and many-objective optimization problems, the above approximate KKTPM values were found to be extremely well correlated to the exact KKTPM values and in most cases the difference was small [7]. This study allowed a computationally fast way to compute KKTPM values for multiple trade-off points in a multi-objective optimization algorithm.

3. PROPOSED NUMERICAL GRADIENT KKT PROXIMITY MEASURE

Although the approximate KKTPM computational procedure enabled a fast approach to computing KKTPM, it still required the knowledge of exact gradients of objective and constraint functions at every solution. In this section, we investigate the accuracy of both exact and approximate KKTPM computation procedures when gradients are computed numerically. The purpose of this study is that if a good enough accuracy can be obtained, the KKTPM procedure becomes more flexible and more widely applicable.

KKTPM requires gradient information of the original objective functions and constraints at the point at which the

KKTPM value needs to be computed. The original study used exact gradients for most examples but demonstrated by working with numerical gradient computation (forward difference method) on a single problem that the induced error is small and is proportional to the step size used for numerical gradient computation. In this paper we make a systematic and more detailed study of computing KKTPM values with numerical gradients. Gradients are calculated using two-point central difference method [20].

For a specified point vector $\mathbf{x}$ and a step size vector $\Delta\mathbf{x}$, the partial derivative of a function for i-th variable is given as follows:

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{f(\mathbf{x} + e_i^T \cdot \Delta\mathbf{x}) - f(\mathbf{x} - e_i^T \cdot \Delta\mathbf{x})}{2(e_i^T \cdot \Delta\mathbf{x})}. \quad (18)$$

Here e_i is the unit vector describing the i-th variable axis. A little thought will reveal that every numerical gradient computation requires $2n$ solution evaluations (n is the number of variables). Since the same $2n$ solutions will be used to compute all objective and constraint functions for their derivative values, a total of $2n$ solution evaluations are needed to compute all gradients numerically for a trade-off solution. Thus, a total of $2nNT$ solution evaluations are required to complete an entire optimization run involving N non-dominated solutions computed for a total of T generations.

4. RESULTS

We are now ready to evaluate the KKTPM values computed using numerical gradients compared to the exact KKTPM values.

We present simulation results on a number of test problems and a few engineering design problems. In all simulations, we use NSGA-II [12] for bi-objective problems and NSGA-III [11, 17] for three and higher objective problems, although other recent EMO methods can also be used. For problems solved using NSGA-II and NSGA-III, we use the simulated binary crossover (SBX) operator [10] with $p_c = 0.9$ and $\eta_c = 30$ and polynomial mutation operator [6] with $p_m = 1/n$ (where n is the number of variables) and $\eta_m = 20$. All other parameters are shown in Table 1.

For each problem, the non-dominated points obtained at each generation are recorded and both optimal and approximate KKTPM values are computed using exact and numerical gradients. Three different step size values ($\Delta x_i = 10^{-2}$, 10^{-3} and 10^{-4}) are chosen to show the level of accuracy obtained for each case.

4.1 Bi-Objective Test Problems

First, we present results of our proposed numerical gradient based KKTPM procedure applied to a number of bi-objective constrained and unconstrained problems. For all problems, we compare our results with the optimal and estimated KKTPM values obtained with gradient functions for an identical number of solution evaluations.

First, we consider results on two specific test problems to illustrate in detail that our exact or approximate KKTPM using mathematical or numerical gradients for calculating KKTPM have very similar values. We consider the constrained problem SRN [6] for this purpose. Figure 3 presents the relationship between the exact KKTPM with mathematical and numerical gradients for each non-dominated solution in each generation of an NSGA-II run. The x-axis marks the exact KKTPM for each individual using exact mathematical gradient functions. As mentioned earlier, for feasible solutions the proposed KKTPM values lie within zero and one. The y-axis marks the exact KKTPM values for the same individuals using numerical gradients computed using the central difference method with a step size of 10^{-4} for each variable. It is clear from the plot that there is a perfect correlation between the two quantities for each non-dominated solution. This is a remarkable result and suggests that the numerical gradients can be used to compute the KKTPM values.

Figure 4 shows the relationship between mathematical and numerical gradient based approximate KKTPM values computed for non-dominated solutions using different step sizes. Once again, a perfect correlation even with a step size of 0.01 is obtained. This indicates that numerical gradients can replace the exact mathematical gradient function computation even for computing the approximate KKTPM values for speeding up the overall computational process.

Next, we consider the BNH problem which is solved using

Table 1: Parameters used in the study are shown here. Columns represent from left to right: problem name, population size, number of generations used by the EMO algorithm.

Problem	Popsize	# Gen. for EMO
ZDT1	40	200
ZDT2	40	200
ZDT3	40	200
ZDT4	48	300
ZDT6	40	200
TNK	12	250
BNH	200	500
SRN	200	500
OSY	200	1000
DTLZ1(3)	92	1000
DTLZ1(5)	212	1000
DTLZ1(10)	276	2000
DTLZ2(3)	92	400
DTLZ2(5)	212	400
DTLZ2(10)	276	1000
DTLZ5(3)	92	500
WELD	60	500
CAR	92	1000

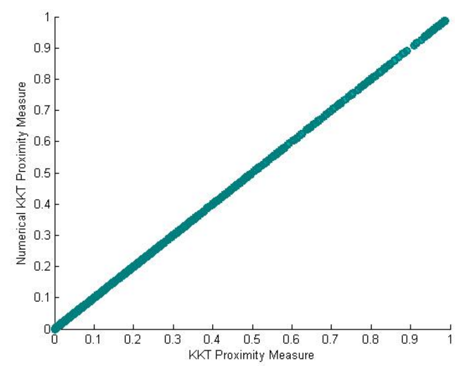
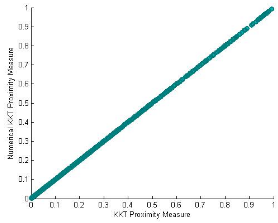
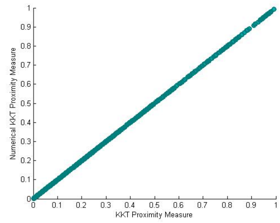


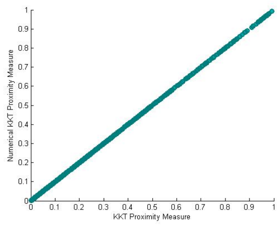
Figure 3: Exact KKTPM values with mathematical and numerical gradients on problem SRN.



(a) Approximate KKTPM using step size 10^{-2} for SRN.



(b) Approximate KKTPM using step size 10^{-3} for SRN.



(c) Approximate KKTPM using step size 10^{-4} for SRN.

Figure 4: Approximate KKTPM values with mathematical and numerical gradients for different step sizes on SRN.

NSGA-II with a population of size 40. Figure 5 shows the variation of exact KKTPM values of non-dominated solutions from the start to finish of multiple runs of NSGA-II with mathematical and numerical gradients. The figure depicts that both methods (with and without numerical gradient) have the same values for KKT proximity measure. To illustrate the effect of step size used in numerical gradient, we plot the variation between estimated KKTPM with and without numerical differentiation in Figure 6. It can be observed that there is no need to use a very small step size in the gradient computation for KKTPM computation and get into the difficulties of numerical instabilities – a value of 0.01 as a step size (1% of the variable range) is adequate for KKTPM computation purpose.

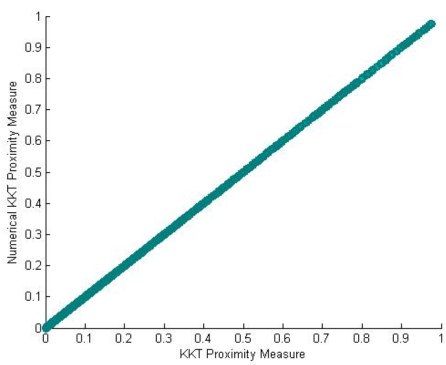
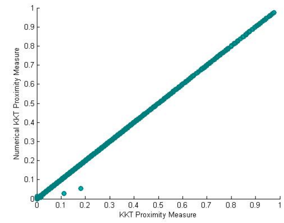
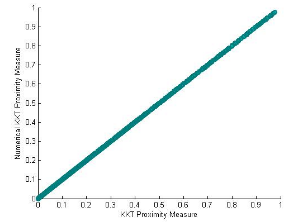


Figure 5: Exact KKTPM values with and without numerical gradient on BNH problem.

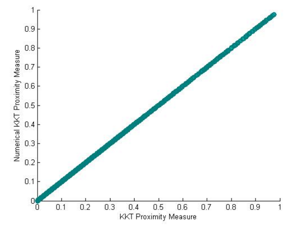
To evaluate the issue further, next, we conduct extensive simulation studies on a wide range of bi-objective problems.



(a) Approximate KKTPM using step size 10^{-2} for BNH.



(b) Approximate KKTPM using step size 10^{-3} for BNH.



(c) Approximate KKTPM using step size 10^{-4} for BNH.

Figure 6: Approximate KKTPM values with mathematical and numerical gradients for different step sizes on BNH.

In all our simulations, we run each problem and calculate the estimated KKTPM values using the mathematical and numerical differentiation methods.

Our bi-objective simulations include set of test problems: ZDT1, ZDT2, ZDT3, ZDT4, and ZDT6 [25]. Figures 7a to 7e present the KKTPM values for ZDT1 to ZDT6 problems solved using NSGA-II. Mathematical and numerical gradients of all objective and constraint functions are used to compute the approximate KKTPM values. We obtained similar results for exact KKTPM values as well, but for brevity, we do not show them here.

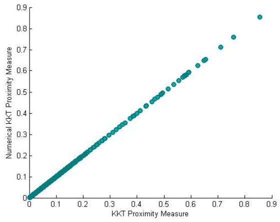
Figures 8 and 9 show similar plots for two other constrained problems: OSY and TNK, respectively. For most solutions of this problem, the correlation is high, except for some very small KKTPM values for which the numerical KKTPM values differ from the exact KKTPM values. To make the overall approach more accurate, the KKTPM can be computed with numerical gradients having a coarse step size early on (when the KKTPM values are large) and as the generations progress and KKTPM values get smaller, numerical gradients can be computed with a smaller step size.

It is clear from these figures that for all problems the numerical gradients produce very similar approximated KKTPM values compared to that obtained using mathematical gradients.

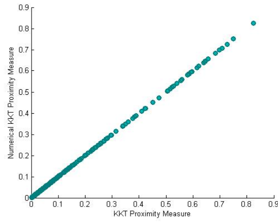
It is clear that the proposed approach achieves the very similar values for the KKTPM computing using both optimal and approximate estimated methods with a very high positive coefficient of correlation.

4.2 Three and Many-objective DTLZ Problems

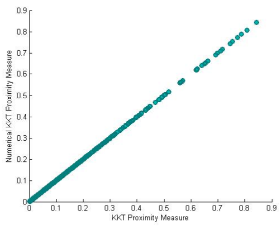
We now consider three and many-objective DTLZ test problems [13]. Figures 10a to 10c show the variation of KKTPM values on three-objective DTLZ1, DTLZ2, and DTLZ5



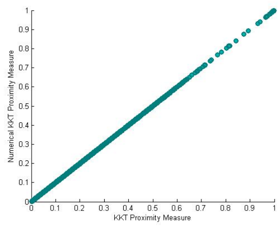
(a) Approximated KKTPM for ZDT1.



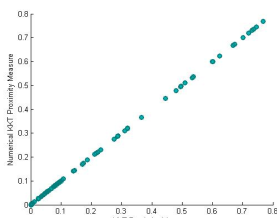
(b) Approximate KKTPM for ZDT2.



(c) Approximate KKTPM for ZDT3.



(d) Approximate KKTPM for ZDT4.



(e) Approximate KKTPM for ZDT6.

Figure 7: Approximate KKTPM values with mathematical and numerical gradients on bi-objective unconstrained ZDT problems.

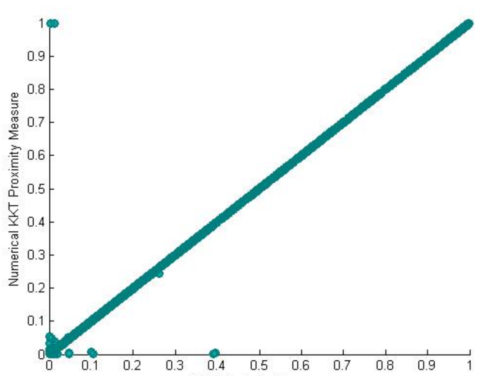


Figure 8: Approximate KKTPM values for OSY.

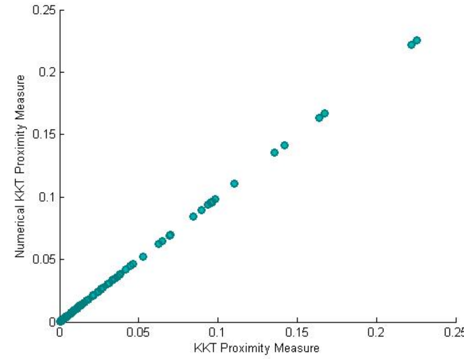


Figure 9: Approximate KKTPM values for TNK.

problems. Problems are solved using NSGA-III procedure. We observe that the same values for KKTPM with mathematical and numerical gradients are obtained.

To demonstrate the advantage of using numerical gradients with NSGA-III on many-objective problems, we further consider five and ten-objective DTLZ1 and DTLZ2 problems. Figures 11a, 11b, 11c and 11d show the relationships between estimated approximate KKTPM values with mathematical and numerical gradients on five and ten-objective DTLZ problems. The KKTPM values with numerical gradients is significantly close to KKTPM values with mathematical gradients on all DTLZ problems. In general, the numerical approach is able to get the very similar KKTPM values to the exact differentiation approach.

4.3 Engineering Design Problems

Finally, we include two engineering design problems: welded-beam design (WELD) [18] and car-side-impact (CAR) [17] problems. In Figures 12a and 12b, the numerical gradient based KKTPM values show a significant correlation with exact differentiation technique. As can be seen from the CAR problem, a difference between KKTPM values obtained using the numerical differentiation and exact differentiation methods exists for a few non-dominated solutions, but an overall correlation between the two quantities is very high.

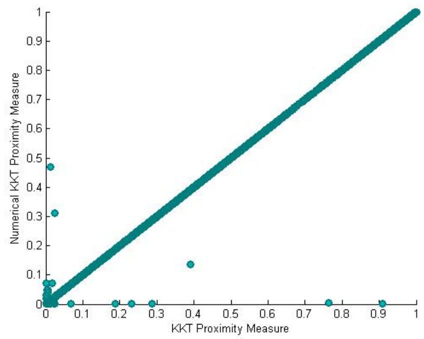
4.4 Accuracy of Numerical gradient KKTPM

The above description between the two sets of KKTPM values is shown graphically. Since the correlation is visually observed to be very high, we did not compute the correlation coefficient before. In this section, we present just that to quantify the relationships between the KKTPM values obtained by mathematical and numerical gradients.

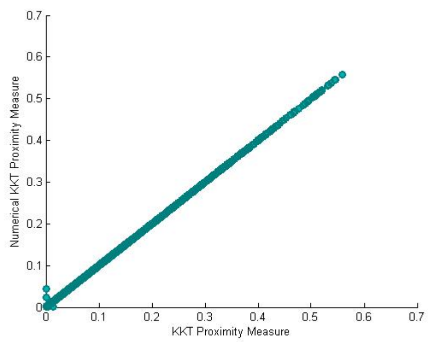
Table 2 presents the correlation coefficient between estimated approximate KKTPM values computed using mathematical and numerical gradients for all the problems considered in this study. A high correlation coefficient between these two measures is observed from the table. For most problems, a correlation coefficient of one is achieved, while the worst value of 0.991 is observed for DTLZ1-10 problem.

5. CONCLUSIONS

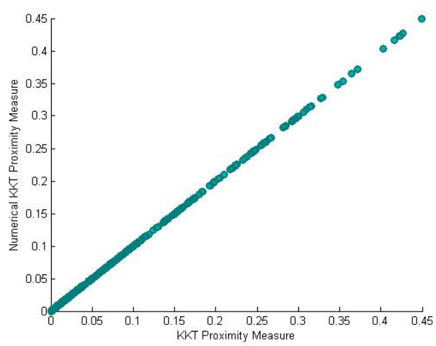
In this study, we have proposed a new numerical gradient based procedure for computing KKTPM values which can be used with any EMO algorithm as a termination condi-



(a) Approximate KKTPM values for DTLZ1.



(b) Approximate KKTPM values for DTLZ2.

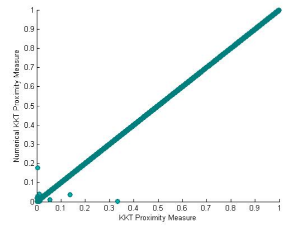


(c) Approximate KKTPM values for DTLZ5.

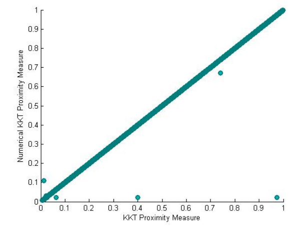
Figure 10: Approximate KKTPM values with and without numerical gradient on three-objective DTLZ problems.

Table 2: Correlation coefficient between approximate KKTPM values computed using mathematical and numerical gradients for multi and many-objective optimization problems of this study.

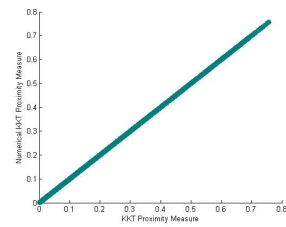
ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
1.000	1.000	1.000	1.000	1.000
BNH	TNK	SRN	OSY	CAR
1.000	1.000	1.000	0.997	0.999
WELD	DTLZ1-3	DTLZ1-5	DTLZ1-10	DTLZ2-3
1.000	0.995	1.000	0.991	0.999
DTLZ2-5	DTLZ2-10	DTLZ5		
1.000	1.000	1.000		



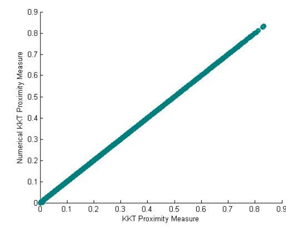
(a) Approximate KKTPM values for five-objective DTLZ1.



(b) Approximate KKTPM values for 10-objective DTLZ1.

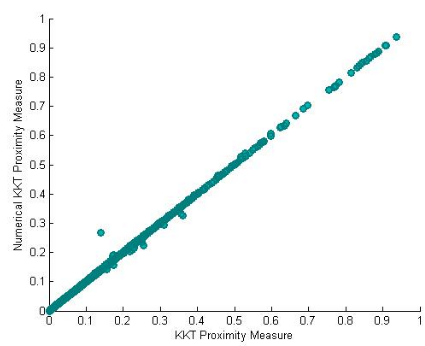


(c) Approximate KKTPM values for five-objective DTLZ2.

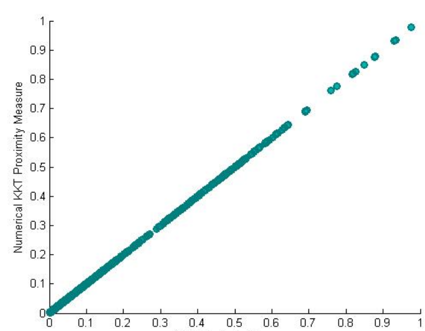


(d) Approximate KKTPM values for 10-objective DTLZ2.

Figure 11: Approximate KKTPM values with mathematical and numerical gradients on five and 10-objective DTLZ problems.



(a) Estimated KKTPM for CAR (3 obj.)



(b) Estimated KKTPM for WELD (2 obj.)

Figure 12: Estimated KKTPM values with and without numerical gradient on two engineering design problems.

tion or in improving its operators for a better convergence. The KKTPM provides the convergence property of a non-dominated point without knowing the true Pareto-optimal front. It has been emphasized that a highly correlated KKTPM computed using a numerical gradients with the mathematically computed gradients will make the KKTPM approach more widely applicable.

Simulation results on a number of two-objective, three-objective, and many-objective unconstrained and constrained problems have demonstrated clearly that a numerical gradient computation (with a step size of 1% of the variable range) produces accurate KKTPM values in most problems. While most problems has shown to have a correlation coefficient of one, the worst correlation of 99.1% has been reported.

These results are encouraging and give us motivation to launch a more detailed study so that the proposed methodology can now be applied to a wide variety of problems. For this purpose, the use of sub-differentials [4] and other KKT optimality theories [14] can be used with our proposed numerical KKTPM construction methodology and applied to non-differentiable problems. We are also currently pursing an adaptive step-size reduction strategy for numerical gradient computations based on the current level of KKTPM values for improving the accuracy of KKTPM computation. The study should encourage researchers to pay more attention to KKTPM and other theoretical optimality properties of solutions in arriving at a faster and more theoretically-sound multi-objective optimization algorithms.

6. REFERENCES

[1] R. Andreani, J. M. Martínez, and B. F. Svaiter. A new sequential optimality condition for constrained optimization and algorithmic consequences. *SIAM Journal on Optimization*, 20(6):3533–3554, 2010.

[2] C. Bector, S. Chandra, and J. Dutta. *Principles of optimization theory*. Alpha Science International, Ltd, 2005.

[3] D. P. Bertsekas, A. Nedi, A. E. Ozdaglar, et al. *Convex analysis and optimization*. Athena Scientific, 2003.

[4] F. H. Clarke. *Methods of dynamic and nonsmooth optimization*. SIAM, 1989.

[5] C. A. C. Coello, D. A. Van Veldhuizen, and G. B. Lamont. *Evolutionary algorithms for solving multi-objective problems*, volume 242. Springer, 2002.

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

[7] K. Deb and M. Abouhawwash. A computationally fast and approximate method for Karush-Kuhn-Tucker proximity measure. Technical Report COIN Report No. 2015015, Department of Electrical and Computer Engineering, Michigan State University, East Lansing, USA, 2015.

[8] K. Deb and M. Abouhawwash. An optimality theory based proximity measure for set based multi-objective optimization. *IEEE Transactions on Evolutionary Computation*, in press.

[9] K. Deb, M. Abouhawwash, and J. Dutta. An optimality theory based proximity measure for evolutionary multi-objective and many-objective optimization. In *Evolutionary Multi-Criterion Optimization*, pages 18–33. Springer, 2015.

[10] K. Deb and R. B. Agrawal. Simulated binary crossover for continuous search space. *Complex Systems*, 9(2):115–148, 1995.

[11] K. Deb and H. Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints. *Evolutionary Computation, IEEE Transactions on*, 18(4):577–601, 2014.

[12] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, 6(2):182–197, 2002.

[13] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. *Scalable test problems for evolutionary multiobjective optimization*. Springer, 2005.

[14] 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.

[15] D. E. Golberg. Genetic algorithms in search, optimization, and machine learning. *Addion wesley*, 1989, 1989.

[16] G. Haeser and M. L. Schuverdt. Approximate KKT conditions for variational inequality problems. *Optimization Online*, 2009.

[17] H. Jain and K. Deb. An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, Part II: Handling constraints and extending to an adaptive approach. *Evolutionary Computation, IEEE Transactions on*, 18(4):602–622, 2014.

[18] J. Knowles, D. Corne, and K. Deb. *Multiobjective problem solving from nature: from concepts to applications*. Springer Science & Business Media, 2007.

[19] K. Miettinen. *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media, 2012.

[20] W. H. Press, S. A. Teuklosky, W. T. Vellerling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge, UK: Cambridge University Press, 1988.

[21] G. Reklaitis, A. Ravindran, and K. Ragsdell. Engineering optimization methods and applications. 1983.

[22] R. T. Rockafellar. Convex analysis (princeton mathematical series). *Princeton University Press*, 46:49, 1970.

[23] R. Tulshyan, R. Arora, K. Deb, and J. Dutta. Investigating EA solutions for approximate KKT conditions in smooth problems. In *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, pages 689–696. ACM, 2010.

[24] A. P. Wierzbicki. The use of reference objectives in multiobjective optimization. In *Multiple criteria decision making theory and application*, pages 468–486. Springer, 1980.

[25] E. Zitzler, K. Deb, and L. Thiele. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary computation*, 8(2):173–195, 2000.