Towards a Quick Computation of Well-Spread Pareto-Optimal Solutions

Kalyanmoy Deb, Manikanth Mohan, and Shikhar Mishra

Kanpur Genetic Algorithms Laboratory (KanGAL)
Indian Institute of Technology Kanpur
Kanpur, PIN 208016, India
deb@iitk.ac.in

Abstract. The trade-off between obtaining a good distribution of Pareto-optimal solutions and obtaining them in a small computational time is an important issue in evolutionary multi-objective optimization (EMO). It has been well established in the EMO literature that although SPEA produces a better distribution compared to NSGA-II, the computational time needed to run SPEA is much larger. In this paper, we suggest a clustered NSGA-II which uses an identical clustering technique to that used in SPEA for obtaining a better distribution. Moreover, we propose a steady-state MOEA based on ϵ-dominance concept and efficient parent and archive update strategies. Based on a comparative study on a number of two and three objective test problems, it is observed that the steady-state MOEA achieves a comparable distribution to the clustered NSGA-II with a much less computational time.

1 Introduction

It has been experienced in the recent past that the computation of a well-diverse set of Pareto-optimal solutions is usually time-consuming [4,8]. Some MOEAs use a quick-and-dirty diversity preservation operator, thereby finding a reasonably good distribution quickly, whereas some MOEAs use a more computationally expensive diversity preservation operator for obtaining a better distribution of solutions. For example, NSGA-II [2] uses a crowding approach which has a computational complexity of $O(N \log N)$, where N is the population size. On the other hand, SPEA [11] uses a clustering approach which has a computational complexity of $O(N^3)$ involving Euclidean distance computations. Although in two-objective problems, the difference in obtained diversity between these two MOEAs was not reported to be significant [2], the difference was clear while solving three-objective problems [4]. SPEA produced a much better distribution at the expense of a large computational effort.

In this paper, we address this trade-off and suggest a steady-state MOEA based on the ϵ-dominance concept suggested elsewhere [8]. The ϵ-dominance does not allow two solutions with a difference ϵ_i in the i-th objective to be non-dominated to each other, thereby allowing a good diversity to be maintained in a population. Besides, the method is quite pragmatic, because it allows the user to

C.M. Fonseca et al. (Eds.): EMO 2003, LNCS 2632, pp. 222–236, 2003.
© Springer-Verlag Berlin Heidelberg 2003

choose a suitable ϵ_i depending on the desired resolution in the i-th objective. In the proposed ϵ-MOEA, two populations (EA and archive) are evolved simultaneously. Using one solution each from both populations, two offspring solutions are created. Each offspring is then used to update both parent and archive populations. The archive population is updated based on the ϵ-dominance concept, whereas an usual domination concept is used to update the parent population. Since the ϵ-dominance concept reduces the cardinality of the Pareto-optimal set and since a steady-state EA is proposed, the maintenance of a diverse set of solutions is possible in a small computational time.

In the reminder of the paper, we briefly discuss a clustered version of NSGA-II and then present the ϵ-MOEA approach in details. Thereafter, these two MOEAs are compared along with the original NSGA-II on a number of two and three objective test problems. Finally, some useful conclusions are drawn from the study.

2 Distribution versus Computation Time

In addition to the convergence to the Pareto-optimal front, one of the equally important aspects of multi-objective optimization is to find a widely distributed set of solutions. Since the Pareto-optimal front can be a convex, non-convex, disconnected, piece-wise continuous hyper-surfaces, there are differences in opinion as to how to define a diversity measure denoting the true spread of a finite set of solutions on the Pareto-optimal front. Although the task is easier for a two-objective objective space, the difficulty arises in the case of higher-dimensional objective spaces. This is the reason why researchers have developed different diversity measures, such as the hyper-volume measure [11], the spread measure [9], the chi-square deviation measure [10], the R-measures [6], and others. In maintaining diversity among population (or archive) members, several researchers have used different diversity-preserving operators, such as clustering [11], crowding [2], pre-specified archiving [7], and others. Interestingly, these diversity-preserving operators produce a trade-off between the achievable diversity and the computational time.

The clustering approach of SPEA forms N clusters (where N is the archive size) from $N'(> N)$ population members by first assuming each of N' members forming a separate cluster. Thereafter, all $\binom{N'}{2}$ Euclidean distances in the objective space are computed. Then, two clusters with the smallest distance are merged together to form one cluster. This process reduces the number of clusters to $N' - 1$. The inter-cluster distances are computed again[1] and another merging is done. This process is repeated until the number of clusters is reduced to N. With multiple population members occupying two clusters, the average distance of all pair-wise distances between solutions of two clusters is used. If $(N' - N)$ is of the order of N (the archive size), the procedure requires $O(N^3)$ computations

[1] A special book-keeping procedure can be used to eliminate further computation of pair-wise Euclidean distances.

in each iteration. Since this procedure is repeated in every iteration of SPEA, the computational overhead as well as storage requirements for implementing the clustering concept are large.

On the other hand, NSGA-II used a crowding operator, in which N' (as large as $2N$, where N is the population size) solutions are processed objective-wise. In each objective direction, solutions are first sorted in ascending order of objective value. Thereafter, for each solution an objective-wise crowding distance is assigned equal to the difference between normalized objective value of the neighboring solutions. The overall crowding distance is equal to the sum of the crowding distances from all objectives. Once all distance computations are achieved, solutions are sorted in descending order of crowding distance and the first N solutions are chosen. This procedure requires $O(N \log N)$ computations. It is clear that although NSGA-II is faster than SPEA, the diversity in solutions achievable by NSGA-II is not expected to be as good as that achievable with SPEA.

3 Two Approaches for a Better Spread

Here, we propose a modification to the NSGA-II procedure and an ϵ-dominance based MOEA in order to achieve a better distribution of solutions.

3.1 Clustered NSGA-II

The first approach is a straightforward replacement of NSGA-II's crowding routine by the clustering approach used in SPEA. After the parent and offspring population are combined into a bigger population of size $2N$ and this combined population is sorted into different non-domination levels, only N good solutions are required to be chosen based on their non-domination levels and *nearness* to each other. This is where the original NSGA-II used a computationally effective crowding procedure. In the clustering NSGA-II approach suggested here, we replace the crowding procedure with the clustering approach. The solutions in the last permissible non-dominated level are used for clustering. Let us say that the remaining population slots are N' and the solutions in the last permissible non-domination level from the combined population is n'. By definition, $n' \geq N'$. To choose N' solutions from n', we form N' clusters from n' solutions and choose one representative solution from each cluster. The clustering algorithm used in this study is exactly the same as that used in SPEA [11]. Although this requires a larger computational time, the clustered NSGA-II is expected to find a better distributed set of Pareto-optimal solutions than the original NSGA-II.

3.2 A Steady-State ϵ-MOEA

Next, we suggest a steady-state MOEA based on the ϵ-dominance concept introduced in [8]. The search space is divided into a number of grids (or hyper-boxes) and diversity is maintained by ensuring that a grid or hyper-box can be occupied by only one solution. Although, PAES and its variants [7] are developed with the similar idea, ϵ-dominance is a more general concept.

Here, we suggest a steady-state MOEA based on the ϵ-dominance concept. In the proposed MOEA, we have two co-evolving populations: (i) an EA population $P(t)$ and (ii) an archive population $E(t)$ (where t is the iteration counter). The MOEA is started with an initial population $P(0)$. To start with, $E(0)$ is assigned with the ϵ-non-dominated solutions of $P(0)$. Thereafter, one solution each from $P(t)$ and $E(t)$ are chosen for mating. To choose a solution from $P(t)$, two population members from $P(t)$ are picked at random and a domination check (in the usual sense) is made. If one solution dominates the other, the former is chosen. Otherwise, it indicates that two picked solutions are non-dominated to each other and we simply choose one of the two solutions at random. Let us denote the chosen solution by p. To choose a solution e from $E(t)$, several strategies involving a certain relationship with p can be made. However, here we randomly pick a solution from $E(t)$. After this selection phase, solutions p and e are mated to create λ offspring solutions (c_i, $i = 1, 2, \ldots, \lambda$). Now, each of these offspring solutions are compared with the archive and the EA population for their possible inclusion. Different strategies are used in each case.

For its inclusion in the archive, the offspring c_i is compared with each member in the archive for ϵ-dominance. Every solution in the archive is assigned an identification array ($\mathbf{B}$) as follows:

$$B_j(\mathbf{f}) = \begin{cases} \lfloor (f_j - f_j^{\min})/\epsilon_j \rfloor, & \text{for minimizing } f_j, \\ \lceil (f_j - f_j^{\min})/\epsilon_j \rceil, & \text{for maximizing } f_j. \end{cases} \tag{1}$$

where $f_j^{\min}$ is the minimum possible value of the j-th objective and ϵ_j is the allowable tolerance in the j-th objective beyond which two values are significant to the user. This ϵ_j value is similar to the ϵ used in the ϵ-dominance definition [8]. The identification arrays make the whole search space into grids having ϵ_j size in the j-th objective. Figure 1 illustrates that the solution P ϵ-dominates the entire region ABCDA (in the minimization sense), whereas the original dominance definition allows P to dominate only the region PECFP. For brevity, the rest

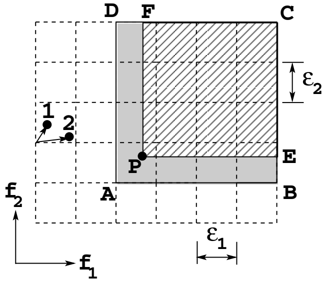

Fig. 1. The ϵ-dominance concept is illustrated (for minimizing f_1 and f_2)

of the discussion is made here for minimization cases only. However, a similar analysis can be followed for maximization cases as well. The identification array of P is the coordinate of point A in the objective space. With the identification arrays calculated for the offspring c_i and each archive member a, we use the following procedure. If the $\mathbf{B}_a$ of any archive member a dominates that of the offspring c_i, it means that the offspring is ϵ-dominated by at least one archive member and the offspring is not accepted. Otherwise, if $\mathbf{B}_c$ of the offspring dominates $\mathbf{B}_a$ of any archive member a, those archive members are deleted and the offspring is accepted. If neither of the above two cases occur, this means that the offspring is ϵ-non-dominated with the archive members. We separate this case into two. If the offspring shares the same $\mathbf{B}$ vector with an archive member (meaning that they belong to the same hyper-box), then they are checked for the usual non-domination. If the offspring dominates the archive member or the offspring is non-dominated to the archive member but is closer to the $\mathbf{B}$ vector (in terms of the Euclidean distance) than the archive member, the offspring is retained. The latter case is illustrated in Figure 1 with solutions 1 and 2. They occupy the same hyper-box (or have the same $\mathbf{B}$ vector) and they are non-dominated according to the usual definition. Since solution 1 has a smaller distance to the $\mathbf{B}$ vector, it is retained and solution 2 is deleted. In the event of an offspring not sharing the same $\mathbf{B}$ vector with any archive member, the offspring is accepted. It is interesting to note that the former condition ensures that only one solution with a distinct $\mathbf{B}$ vector exists in each hyper-box. This means that each hyper-box on the Pareto-optimal front can be occupied by exactly one solution, thereby providing two properties: (i) solutions will be well distributed and (ii) the archive size will be bounded. For this reason, no specific upper limit on the archive size needs to be pre-fixed. The archive will get bounded according to the chosen ϵ-vector.

The decision whether an offspring will replace any population member can be made using different strategies. Here, we compare each offspring with all population members. If the offspring dominates any population member, the offspring replaces that member. Otherwise, if any population member dominates the offspring, it is not accepted. When both the above tests fail, the offspring replaces one randomly chosen population member. This way the EA population size remains unchanged.

The above procedure is continued for a specified number of iterations and the final archive members are reported as the obtained solutions. A careful thought will reveal the following properties of the proposed algorithm:

1. It is a steady-state MOEA.
2. It emphasizes non-dominated solutions.
3. It maintains diversity in the archive by allowing only one solution to be present in each pre-assigned hyper-box on the Pareto-optimal front.
4. It is an elitist approach.

3.3 Differences with Past Studies

Although the above steady-state ϵ-MOEA may look similar to the multi-parent PAES [7], there is a subtle difference. The ϵ-dominance concept implemented here with the **B**-vector domination-check does not allow two non-dominated solutions with a difference less than ϵ_i in the i-th objective to be both present in the final archive. Figure 1 can be used to realize that fewer non-dominated solutions will be obtained with the proposed approach than PAES. The procedure will not only allow a reduction in the size of the Pareto-optimal solutions, but also has a practical significance. Since a user is not interested in obtaining solutions within a difference ϵ_i in the i-th objective, the above procedure allows the user to find solutions according to his/her desire. However, the actual number of solutions to be obtained by the ϵ-dominance procedure is unknown, but it is bounded. Because of this reason, the overall computation is expected to be faster.

The proposed archive update strategy is also similar to that proposed in [8], except in the case when two solutions have the same **B** vector. Here, a non-dominated solution to an existing archive member can still be chosen if it is closer to the **B** vector. The earlier study only accepted an offspring if it dominated an existing member. Moreover, here we have proposed a steady-state MOEA procedure with an EA population update strategy, an archive update strategy, and a recombination plan.

4 Simulation Results

In this section, we compare the three MOEAs discussed above on a number of two and three objective test problems. We have used a convergence measure for exclusively computing the extent of convergence to the Pareto-optimal front and a sparsity measure for exclusively computing the diversity in solutions. The hyper-volume measure [11] for computing a combined convergence and diversity estimate is also used. Since all problems considered in this paper are test problems, the exact knowledge of the Pareto-optimal front is available. We calculate H uniformly distributed (on the f_1-f_2-$\cdots$-f_{M-1}-plane) solutions P^* on the Pareto-optimal front. For each such point in the $(M-1)$-dimensional plane, f_M is calculated from the known Pareto-optimal front description. Then, the Euclidean distance of each obtained solution from the nearest solution in P^* is computed. The average of the distance value of all obtained solutions is defined as the convergence measure here. The sparsity measure is described in detail while discussing the three-objective optimization results. We also present the computational time needed to run each MOEA on the same computer.

4.1 Two-Objective Test Problems

ZDT1 Test Problem: The $n = 30$-variable ZDT1 problem has a convex Pareto-optimal front. We use a population size of 100 and real-parameter SBX recombination operator with $p_c = 1$ and $\eta_c = 15$ and a polynomial mutation operator with $p_m = 1/n$ and $\eta_m = 20$ [1]. In order to investigate the effect of ϵ used in the ϵ-MOEA, we use different ϵ values and count the number of solutions found in the archive in each case after 20,000 solution evaluations. Figure 2 shows that as ϵ increases, the number of obtained solutions varies proportional to $1/\epsilon$. For an equally-sloped Pareto-optimal straight line in the range $f_1 \in [0,1]$, we would expect $1/\epsilon$ solutions to appear in the archive. But with a non-linear Pareto-optimal front, we would expect smaller number of archive members with the ϵ-dominance concept. This will be clear from Figure 3, which shows the distribution of solutions obtained with $\epsilon = 0.05$. The boxes within which a solution lies are also shown in the figure. It is interesting to note that all solutions are ϵ-nondominated with respect to each other and in each of the expected boxes only one solution is obtained. In the boxes with $f_1 \in [0, 0.05]$, one Pareto-optimal

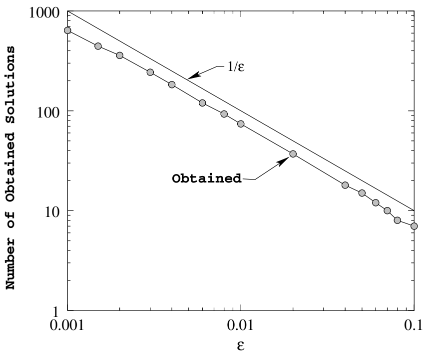

Fig. 2. The number of solutions versus ϵ on ZDT1

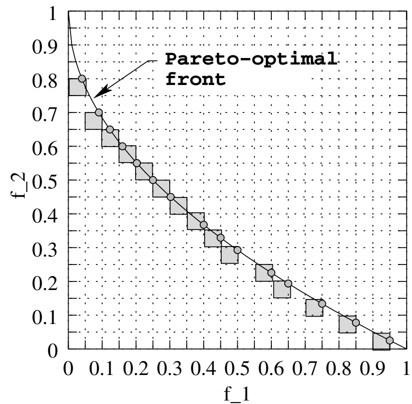

Fig. 3. ϵ-MOEA distribution on ZDT1 with $\epsilon = 0.05$

solution in the minimum f_2 box ($f_2 \in [0.75, 0.80]$) is obtained. Other four boxes on top of this box (with larger f_2) are ϵ-dominated and hence not retained in the final archive.

To compare the three MOEAs, we have used the convergence metric discussed above, the hyper-volume metric [11], and a sparsity measure. We use $\epsilon_i = 0.008$ in order to get roughly 100 solutions in the archive after 20,000 solution evaluations. The convergence metric is computed with $H = 1,000$ equi-spaced solutions on the Pareto-optimal front. Table 1 shows that the convergence of solutions achieved with ϵ-MOEA is the best. For the hyper-volume measure, we use the reference point at $(1.1, 1.1)^T$. Although the convergence of the ϵ-MOEA is better than that of the C-NSGA-II in this problem, the hyper-volume of ϵ-MOEA is worse. This is mainly due to the absence of extreme solutions in the Pareto-optimal front (refer Figure 3). Since the ϵ-dominance concept is used in ϵ-MOEA,

Table 1. Performance comparison of three MOEAs for ZDT1, ZDT4, and ZDT6. Best metric values are shown in a slanted font

Test Problem	MOEA	Convergence measure Avg.	Convergence measure Std. Dev.	Hyper-volume Measure	Sparsity Measure	Time (sec)
ZDT1	NSGA-II	0.00060	9.607e-05	0.87060	0.857	7.25
	C-NSGA-II	0.00054	1.098e-04	*0.87124*	0.997	1984.83
	ϵ-MOEA	*0.00041*	2.818e-05	0.87031	*1.000*	*1.20*
ZDT4	NSGA-II	0.01356	2.955e-04	0.85039	0.937	5.64
	C-NSGA-II	0.00884	2.314e-04	*0.85747*	*1.000*	99.01
	ϵ-MOEA	*0.00227*	6.019e-05	0.85723	0.943	*0.57*
ZDT6	NSGA-II	0.07104	1.186e-04	0.40593	0.803	3.93
	C-NSGA-II	*0.06507*	6.943e-04	*0.41582*	*1.000*	2365.95
	ϵ-MOEA	0.07280	1.777e-04	0.40174	0.997	*0.91*

the extreme solutions usually get dominated by solutions within ϵ and which are better in other objectives. Figure 3 illustrates this aspect. The presence or absence of the extreme solutions makes a significant difference in the hyper-volume metric, a matter we discuss further later. Thus, the hyper-volume measure may not be an ideal metric for measuring diversity and convergence of a set of solutions. We suggest and use a different diversity measure (which we have described in Section 4.2 in detail for any number of objectives). In short, the sparsity measure first projects the obtained solutions on a suitable hyper-plane (a unit vector $(1/\sqrt{2}, 1/\sqrt{2})^T$ is used here) and then computes the non-overlapping area occupied by the solutions on that plane. The higher the measure value, the better is the distribution. This measure is also normalized so that the maximum possible non-overlapping area is one, meaning that there is no overlap (or 100% non-overlap) among projected solutions. The measure involves a size parameter, which is adjusted in such a way that one of the competing MOEAs achieve a 100% non-overlapping area. For ZDT1, ϵ-dominance achieves the best sparsity measure.

Finally, the table presents the actual computational time needed by each MOEA. It is clear that the ϵ-MOEA obtained a better convergence and sparsity measure in a much smaller computational time than the other two MOEAs.

ZDT4 Test Problem: Table 1 also shows the performance measures on the 10-variable ZDT4. This problem has a number of local Pareto-optimal fronts. All chosen parameters used here are identical to that used in ZDT1, except that $\epsilon_i = 0.006$ are used to get about 100 solutions in the final archive (after 20,000 solution evaluations). The table shows that ϵ-MOEA is better than the other two MOEAs in terms of convergence and computational time, and is the second-best in terms of diversity among obtained solutions.

ZDT6 Test Problem: The 10-variable ZDT6 problem has a non-uniform density of solutions on the Pareto-optimal front. Here, we use $\epsilon_i = 0.0067$ to get about 100 solutions in the archive after 20,000 solution evaluations. Here, the convergence measures of C-NSGA-II and NSGA-II are slightly better than that of ϵ-MOEA. However, the diversity measures (Table 1) obtained in ϵ-MOEA and C-NSGA-II are comparable. But, ϵ-MOEA is superior in terms of the computational time.

Thus, from the two-objective problems studied above, we can conclude that ϵ-MOEA produces a good convergence and diversity with a smaller (at least an order of magnitude smaller) computational time than the other two MOEAs.

4.2 Three-Objective Test Problems

Now, we consider a few three-objective test problems developed elsewhere [4].

DTLZ1 Test Problem: First, we consider the $n = 7$ variable, three-objective DTLZ1 test problem. The Pareto-optimal solutions lie on a three-dimensional plane satisfying: $f_3 = 0.5 - f_1 - f_2$ in the range $f_1, f_2 \in [0, 0.5]$. For each algorithm, we use a population size of 100, real-parameter representation with the SBX recombination (with $\eta_c = 15$ and $p_c = 1$), the polynomial mutation operator (with $\eta_m = 20$ and $p_m = 1/n$) [1], and a maximum of 30,000 function evaluations. For ϵ-MOEA, we have chosen $\epsilon = [0.02, 0.02, 0.05]^T$. The solutions obtained in each case are shown on the objective space in Figures 4, 5, and 6, respectively. It is clear from the figures that the distribution of solutions

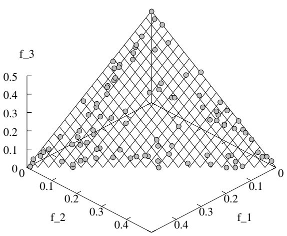

Fig. 4. NSGA-II distribution on DTLZ1

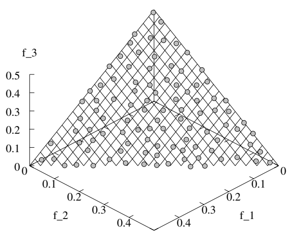

Fig. 5. C-NSGA-II distribution on DTLZ1

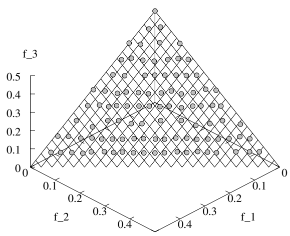

Fig. 6. ϵ-MOEA distribution on DTLZ1

with the original NSGA-II is poor compared to the other two MOEAs. Using $H = 5,000$ solutions, we tabulate the average and standard deviation of the convergence measure in columns 2 and 3 in Table 2, respectively. It is observed that the convergence of the ϵ-MOEA is relatively better than that of the other two MOEAs with identical function evaluations.

Column 4 of the table shows the hyper-volume measure calculated with a reference solution $f_1 = f_2 = f_3 = 0.7$. Since for minimization problems, a larger hyper-volume is better, the table indicates that the C-NSGA-II is best,

Table 2. Comparison of three MOEAs in terms of their convergence and diversity measures on DTLZ1

MOEA	Convergence measure Avg.	Std. Dev.	Hyper-volume Measure	Sparsity Measure	Time (sec)
NSGA-II	0.00267	1.041e-04	0.313752	0.867	8.46
C-NSGA-II	0.00338	8.990e-05	*0.314005*	*1.000*	3032.38
ϵ-MOEA	*0.00245*	9.519e-05	0.298487	0.995	*2.22*

followed by the NSGA-II. The ϵ-MOEA performs the worst. However, by visually investigating solutions displayed in Figures 4 and 6 for NSGA-II and ϵ-MOEA, respectively, it can be seen that the distribution of ϵ-MOEA is better. The hyper-volume measure fails to capture this aspect, despite a better convergence and sparsity of ϵ-MOEA solutions. The extent of solutions obtained by these two MOEAs are shown below:

	f_1	f_2	f_3
NSGA-II:	0–0.500525	0–0.501001	0–0.500763
ϵ-MOEA:	0.000273–0.435053	0.000102–0.434851	0.044771–0.499745

Since the ϵ-dominance does not allow two solutions with a difference of ϵ_i in the i-th objective to be mutually non-dominated to each other, it will be usually not possible to obtain the extreme corners of the Pareto-optimal front. However, the diversity of solutions elsewhere on the Pareto-optimal front is ensured by the archive update procedure of ϵ-MOEA. It is then trivial to realize that without the extreme solutions the hyper-volume measure fails to add a large portion of hyper-volume calculated from the extreme points. The advantage in diversity metric obtained from the well-distributed solutions on the interior of the Pareto-optimal front was not enough to account for the loss in diversity due to the absence of boundary solutions. We argue that in this sense the hyper-volume measure is biased towards the boundary solutions. Thus, we do not use this measure for the rest of the test problems used in this paper.

Instead, we define a *sparsity* measure, which is similar to the entropy measure [5] or the grid diversity measure [3]. The Pareto-optimal solutions are first projected on a suitable hyper-plane (with a unit vector $\boldsymbol{\eta}$). Figure 7 illustrates the calculation procedure of this measure. A hyper-box of certain size d is centered around each projected solution. The total hyper-volume covered by these hyper-boxes is used as the measure of sparsity of solutions. If a solution set has many clustered points, their hyper-boxes will overlap with each other and the obtained sparsity measure will be a small number. On the other hand, if solutions are well distributed, the hyper-boxes do not overlap and a large overall measure will be obtained. To normalize the measure, we divide the total hyper-volume by the total expected hyper-volume calculated with a same-sized solution set having no overlap between the hyper-boxes. Thus, the maximum sparsity achievable is one and the larger the sparsity measure the better is the distribution. However, the choice of the parameter d is important here. A too small value of d will make any

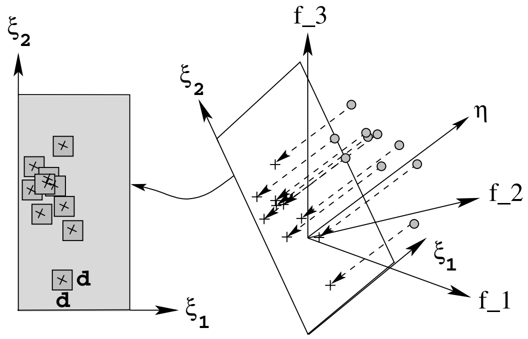

Fig. 7. The sparsity measure is illustrated

distribution to achieve the maximum sparsity measure of one, whereas a very large value of d will make every distribution to have a small sparsity measure. We solve this difficulty of choosing a suitable d value by finding the smallest possible value which will make one of the competing distributions to achieve the maximum sparsity value of one and use the same d for computing sparsity for other distributions. In all case studies here, points are projected on an equally-inclined plane to coordinate axes ($\boldsymbol{\eta} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^T$). The column 5 of Table 2 shows that C-NSGA-II attains the best distribution, followed by ϵ-MOEA, and then NSGA-II. The distribution of solutions plotted in Figures 4 to 6 also support this observation visually.

Although the C-NSGA-II achieves the best distribution, it is also computationally the slowest of the three MOEAs. Since the clustering algorithm requires comparison of each population member with each other for computing adequate number of clusters in every generation of C-NSGA-II, the table shows that it takes three orders of magnitude more time than the other two MOEAs. Based on both convergence and diversity measures and the required computational time tabulated in the table, the ϵ-MOEA emerges out as a good compromised algorithm.

DTLZ2 Test Problem: Next, we consider the 12-variable DTLZ2 test problem having a spherical Pareto-optimal front satisfying $f_1^2 + f_2^2 + f_3^2 = 1$ in the range $f_1, f_2 \in [0, 1]$. Identical parameters to those used in DTLZ1 are used here. About $H = 8,000$ Pareto-optimal solutions are considered as P^* for the convergence metric computation. For the ϵ-MOEA, we have used $\epsilon = [0.06, 0.06, 0.066]^T$. This produces about 100 solutions on the Pareto-optimal front. Table 3 shows the comparison of performance measures of three MOEAs.

Figures 8 and 9 shows the distribution of solutions obtained by C-NSGA-II and ϵ-MOEA, respectively. Although both approaches produced a very similar sparsity measure, they produced them in different ways. In the C-NSGA-II, since

Table 3. Performance comparison of three MOEAs for DTLZ2

MOEA	Convergence measure Avg.	Convergence measure Std. Dev.	Sparsity Measure	Time (sec)
NSGA-II	0.01606	0.00112	0.821	6.87
C-NSGA-II	*0.00986*	0.00088	0.998	7200.51
ϵ-MOEA	0.01160	0.00119	*1.000*	*1.86*

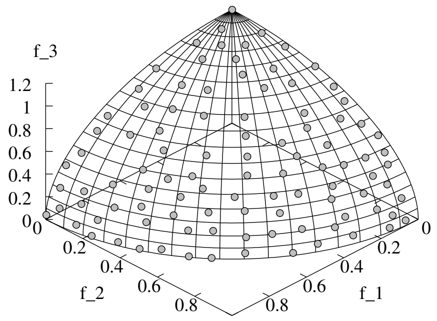

Fig. 8. C-NSGA-II distribution on DTLZ2

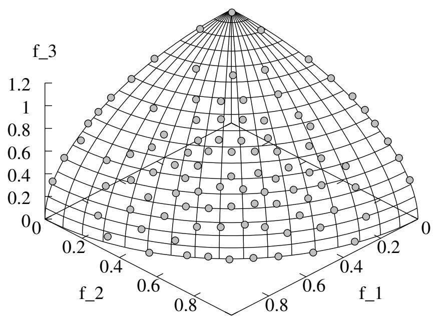

Fig. 9. ϵ-MOEA distribution on DTLZ2

the Euclidean distance measure is used in the clustering approach for maintaining diversity, a uniform spread of solutions on the front is observed. In the case of ϵ-MOEA, there seems to be a considerable amount of gap observed between the boundary solutions and their nearest neighbors. This happens because of the fact that there is a gentle slope near the boundary solutions on a spherical surface and the ϵ-dominance consideration does not allow any solution to be non-dominated within an ϵ_i in the i-th objective. But, wherever there is a considerable change of slope, more crowded solutions are found. It is argued earlier that such a set of solutions with a minimum pre-specified difference in objective values has a practical significance and hence we advocate the use of ϵ-dominance in this paper.

It is clear from the Table 3 that ϵ-MOEA achieves a good convergence and diversity measure with a much less computational time than C-NSGA-II.

DTLZ4 Test Problem: The 12-variable DTLZ4 test problem introduces a non-uniform density of solution on the Pareto-optimal front. The first two variables are raised to the power of 100, thereby making the $x_1 = 0$ and $x_2 = 0$ solutions to have a larger probability to be discovered. In such a problem, a uniform distribution of Pareto-optimal solutions is difficult to obtain. Table 4 shows the performance of three MOEAs. For ϵ-MOEA, we have used $\epsilon_i = 0.062$ to obtain

Table 4. Performance comparison of three MOEAs for DTLZ4

MOEA	Convergence measure Avg.	Convergence measure Std. Dev.	Sparsity Measure	Time (sec)
NSGA-II	0.03658	0.00172	0.339	10.61
C-NSGA-II	0.01235	0.00115	0.999	5041.85
ϵ-MOEA	*0.00938*	0.00094	*1.000*	*3.86*

100 solutions on the Pareto-optimal front. Once again, the sparsity measure of C-NSGA-II and ϵ-MOEA are much better than that of NSGA-II. However, the ϵ-MOEA achieves this diversity with a much smaller computational time.

DTLZ5 Test Problem: The DTLZ5 is a 12-variable problem having a Pareto-optimal curve: $f_3^2 = 1 - f_1^2 - f_2^2$ with $f_1 = f_2 \in [0, 1]$. Table 5 shows the performance measures. Here, we use $\epsilon_i = 0.005$ for the ϵ-MOEA. It is also clear

Table 5. Performance comparison of three MOEAs for DTLZ5

MOEA	Convergence measure Avg.	Convergence measure Std. Dev.	Sparsity Measure	Time (sec)
NSGA-II	0.00208	0.00038	0.970	5.23
C-NSGA-II	0.00268	0.00034	0.990	1636.96
ϵ-MOEA	*0.00091*	0.00018	*1.000*	*1.58*

from the table that ϵ-MOEA is the quickest and best in terms of achieving convergence and diversity.

DTLZ8 Test Problem: The 30-variable DTLZ8 test problem is chosen next. The overall Pareto-optimal region is a combination of two fronts: (i) a line and (ii) a plane. Figure 10 shows the distribution of points obtained using the C-NSGA-II after 100,000 evaluations. As discussed elsewhere [4], the domination-based MOEAs suffers from the so-called 'redundancy problem' in this test problem. For two distinct solutions on the line portion of the Pareto-optimal front, many other non-Pareto-optimal solutions appear as non-dominated. In the figure, solutions on the adjoining sides (shown shaded) of the Pareto-optimal line are these redundant solutions and the clustered NSGA-II is unable to get rid of these solutions, because these solutions are non-dominated to some Pareto-optimal solutions. But with ϵ-dominance, many of these redundant solutions get ϵ-dominated by the Pareto-optimal solutions. Figure 11 shows the solutions obtained with ϵ-MOEA having $\epsilon = [0.02, 0.02, 0.04]^T$. With a 30-variable decision

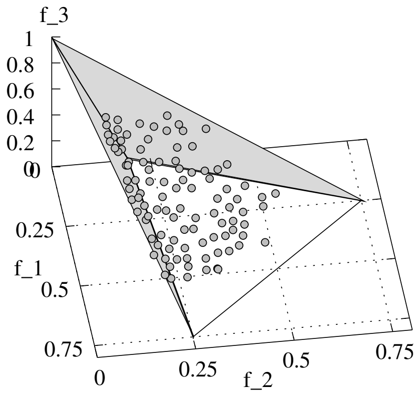

Fig. 10. C-NSGA-II distribution on DTLZ8

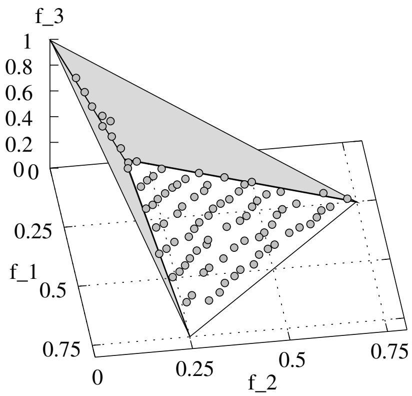

Fig. 11. ϵ-MOEA distribution on DTLZ8

space, the density of solutions near the Pareto-optimal line and near the $f_3 = 0$ part of the Pareto-optimal plane are very small. Thus, it may be in general difficult to find solutions at these portions in the Pareto-optimal front. For ϵ-MOEA, we have used $\eta_c = 2$ and $\eta_m = 5$, however for C-NSGA-II $\eta_c = 15$ and $\eta_m = 20$ produced better results. It is clear from the plot that the ϵ-MOEA is able to find a reasonable distribution of solutions on the line and the plane. Although ϵ-MOEA is able to get rid of most redundant solutions, the problem is not entirely eradicated by this procedure. However, the number of redundant solutions is much smaller than a procedure which uses the original dominance criterion.

5 Conclusions

Achieving a good distribution of solutions in a small computational time is a dream to an MOEA researcher or practitioner. Although past studies have either demonstrated a good distribution with a large computational overhead or a not-so-good distribution quickly, in this paper we have suggested an ϵ-dominance based steady-state MOEA to achieve a good distribution with a computationally fast procedure. Careful update strategies for archive and parent populations ensure a continuous progress towards the Pareto-optimal front and maintenance of a good diversity. The use of ϵ-dominance criterion has been found to have two advantages: (i) it helps in reducing the cardinality of Pareto-optimal region and (ii) it ensures that no two obtained solutions are within an ϵ_i from each other in the i-th objective. The first aspect is useful in using the proposed ϵ-MOEA to higher-objective problems and to somewhat lessen the 'redundancy' problem [4]. The second aspect makes the approach highly suitable for practical problem solving, particularly in making the MOEA approach interactive with a decision-maker.

On a number of two and three-objective test problems, the proposed ϵ-MOEA has been successful in finding well-converged and well-distributed solutions with a much smaller computational effort than the original NSGA-II and a clustered yet computationally expensive NSGA-II. The consistency in the solution accuracy and the requirement of a fraction of computational effort needed in other MOEAs suggest the use of the proposed ϵ-MOEA to more complex and real-world problems. The study recommends a more rigorous comparison of ϵ-MOEA with other competing MOEAs, such as the new versions of SPEA and PAES. By choosing ϵ_i as a function of f_i, the method can also be used to find a biased distribution of solutions on the Pareto-optimal front, if desired.

References

1. K. Deb. *Multi-objective optimization using evolutionary algorithms.* Chichester, UK: Wiley, 2001.
2. K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.
3. K. Deb and S. Jain. Running performance metrics for evolutionary multi-objective optimization. In *Simulated Evolution and Learning (SEAL-02)*, in press.
4. K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable multi-objective optimization test problems. In *Proceedings of the Congress on Evolutionary Computation (CEC-2002)*, pages 825–830, 2002.
5. A. Farhang-Mehr and S. Azarm. Diversity assessment of pareto-optimal solution sets: An entropy approach. In *Proceedings of the World Congress on Computational Intelligence*, pages 723–728, 2002.
6. M. P. Hansen and A. Jaskiewicz. Evaluating the quality of approximations to the non-dominated set. Technical Report IMM-REP-1998-7, Lyngby: Institute of Mathematical Modelling, Technical University of Denmark, 1998.
7. Joshua D. Knowles and David W. Corne. Approximating the non-dominated front using the Pareto archived evolution strategy. *Evolutionary Computation Journal*, 8(2):149–172, 2000.
8. M. Laumanns, L. Thiele, K. Deb, and Eckart Zitzler. Combining convergence and diversity in evolutionary multi-objective optimization. *Evolutionary Computation*, 10(3):263–282, 2002.
9. J. R. Schott. Fault tolerant design using single and multi-criteria genetic algorithms. Master's thesis, Boston, MA: Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 1995.
10. N. Srinivas and K. Deb. Multi-objective function optimization using non-dominated sorting genetic algorithms. *Evolutionary Computation Journal*, 2(3):221–248, 1994.
11. E. Zitzler and L. Thiele. Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, 1999.