An Alternative Preference Relation to Deal with Many-Objective Optimization Problems

Antonio López[2], Carlos A. Coello Coello[1], Akira Oyama[2], and Kozo Fuji[2]

[1] CINVESTAV-IPN⋆ (Evolutionary Computation Group)
Departamento de Computación, México D.F. 07300, MÉXICO
ccoello@cs.cinvestav.mx
[2] Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency,
Sagamihara 252-5210, JAPAN
{antonio,oyama,fujii}@flab.isas.jaxa.jp

Abstract. In this paper, we use an alternative preference relation that couples an achievement function and the ϵ-indicator in order to improve the scalability of a Multi-Objective Evolutionary Algorithm (MOEA) in many-objective optimization problems. The resulting algorithm was assessed using the Deb-Thiele-Laumanns-Zitzler (DTLZ) and the Walking-Fish-Group (WFG) test suites. Our experimental results indicate that our proposed approach has a good performance even when using a high number of objectives. Regarding the DTLZ test problems, their main difficulty was found to lie on the presence of dominance resistant solutions. In contrast, the hardness of WFG problems was not found to be significantly increased by adding more objectives.

1 Introduction

Since the first implementation of a Multi-Objective Evolutionary Algorithm (MOEA) in the mid 1980s, a wide variety of new MOEAs have been proposed, gradually improving in both their effectiveness and efficiency to solve Multiobjective Optimization Problems (MOPs) [1]. However, most of these algorithms have been evaluated in problems with only two or three objectives, in spite of the fact that many real-world problems have more than three objectives.

Recently, the Evolutionary Multiobjective Optimization community has devoted important efforts to investigate the performance of MOEAs in problems with a high number of objectives. These MOPs are usually known as Many-objective Optimization Problems (𝕄OPs). One of the first findings in this area [2, 3] is that MOEAs based on Pareto optimality scale poorly with respect to the number of objectives. Currently, two main difficulties that make a problem harder when the number of objectives is increased have been suggested:

– Increase of the proportion of nondominated solutions. Since in 𝕄OPs almost all solutions are equivalent in terms of Pareto optimality, many researchers have suggested [4–6] that in such problems, the selection of the appropriate

⋆ The second author acknowledges support from CONACyT project no. 103570.

individuals for steering the population towards the Pareto optimal set gets more difficult. However, as pointed out by Schütze et al. [7], the increase of the number of nondominated individuals is not a sufficient condition for an increase of the hardness of a problem. They found that in a class of uni-modal problems, the difficulty was marginally increased when more objectives are added.

– Effectiveness of crossover operators. In a combinatorial class of MOPs, Sato et al. [8] observed that solutions in decision variable space become more distant[3] from each other as more objectives are added. As a result, even if two parents close to the Pareto front are recombined, the generated offspring might be far from the Pareto front.

Although not related with the search ability of the MOEA, there are other important difficulties associated with a MOP. For example, the visualization of the Pareto front in high dimensional spaces, or the generation of an accurate sample of the Pareto front, since the required number of points increases exponentially with the number of objectives.

Although the rise of the proportion of incomparable solutions might not significantly determine the difficulty of a MOP *per se*, it seems that the addition of objectives aggravates some particular difficulties observed in the context of 2 or 3 objectives. This is the case of the so called Dominance Resistant Solutions (DRSs) or outliers [9–11]. DRSs are non Pareto optimal solutions with a poor value in at least one of the objectives, but with near optimal values in the others. These kinds of solutions represent potential difficulty since the number of DRSs grows as the number of objectives is increased.

In this paper, we propose the use of the recently introduced Chebyshev preference relation [12] in order to improve the scalability of a MOEA in MOPs. That new preference relation divides the objective space in two regions. In the region farther from the ideal point, the solutions are compared using an achievement scalarizing function, whereas in the region near the ideal point, solutions are compared using the usual Pareto dominance. The idea behind this proposal is to increase the selection pressure when the solutions are far from the Pareto front. This way, we have a discriminative criterion to evaluate nondominated solutions.

Additionally, we introduce the idea of coupling the Chebyshev relation with two preference relations based on the ϵ-indicator. These new preference relations show that a straightforward use of the ϵ-indicator produces a good approximation of the Pareto front.

The experiments are concentrated in evaluating the performance of the Chebyshev preference relation and also in the sources of difficulty when the number of objectives is increased. For the experiments we employed 5 problems from the DTLZ test suite, and 2 problems from the WFG test suite.

[3] In terms of Hamming distance between binary encoded solutions.

2 Basic Concepts and Notation

This section briefly presents the concepts and notation used throughout the rest of the paper.

2.1 Multiobjective Optimization Problems

Definition 1. *A MOP is defined as:*

$$\underset{\mathbf{x}\in\mathcal{X}}{\textit{Minimize}} \quad \mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})]. \tag{1}$$

The vector function $\mathbf{f} : \mathcal{X} \to \mathbb{R}^k$ is composed by $k \geq 2$ *objective functions* $f_i : \mathcal{X} \to \mathbb{R}$ $(i = 1, \ldots, k)$. The image of the *feasible set* $\mathcal{X} \subseteq \mathbb{R}^n$ under the function $\mathbf{f}$ is a subset of the objective function space denoted by $\mathcal{Z} = \mathbf{f}(\mathcal{X})$. The sets $\mathbb{R}^n$ and $\mathbb{R}^k$ are known as *decision variable space* and *objective function space*, respectively.

In multiobjective optimization, the *Pareto dominance relation* is usually adopted to compare vectors in $\mathbb{R}^k$.

Definition 2. *A vector* $\mathbf{z}^1 \in \mathbb{R}^k$ *is said to dominate vector* $\mathbf{z}^2 \in \mathbb{R}^k$ *(denoted* $\mathbf{z}^1 \prec_{\text{par}} \mathbf{z}^2$*) if and only if:* $z_i^1 \leq z_i^2$ $(i = 1, \ldots, k)$, *and* $\mathbf{z}^1 \neq \mathbf{z}^2$.

Definition 3. *A solution* $\mathbf{x}^* \in \mathcal{X}$ *is Pareto optimal if there is no solution* $\mathbf{x} \in \mathcal{X}$ *such that* $\mathbf{f}(\mathbf{x}) \prec_{\text{par}} \mathbf{f}(\mathbf{x}^*)$.

Definition 4. *The Pareto optimal set,* P_{opt}*, is composed by all the Pareto optimal solutions.*

Definition 5. *The image of* P_{opt} *under the vector function* $\mathbf{f}(\mathbf{x})$ *is called the Pareto optimal front and is denoted by* PF_{opt}.

In practice, the goal of a MOEA is finding the best approximation set of the Pareto optimal front. We denote an approximation set by PF_{apx}. Currently, it is well accepted that the quality of an approximation set is determined by the closeness to the Pareto optimal front, and the spread over the entire Pareto optimal front.

In some cases it is useful to know the lower and upper bounds of the Pareto front. The *ideal point* represents the lower bound and is defined by $z_i^\star = \min_{z\in\mathcal{Z}}(z_i)$ for all $i = 1, \ldots, k$. In turn, the upper bound is defined by the *nadir point*, which is given by $z_i^{\text{nad}} = \max_{z\in PF_{\text{opt}}}(z_i)$ for all $i = 1, \ldots, k$.

2.2 Achievement Scalarizing Functions

The preference relation adopted in this paper is based on the achievement scalarizing function approach proposed by Wierzbicki [13].

Definition 6. *An achievement (scalarizing) function is a parameterized function* $s(\mathbf{z}|\mathbf{z}^{\text{ref}}) : \mathbb{R}^k \rightarrow \mathbb{R}$*, where* $\mathbf{z}^{\text{ref}} \in \mathbb{R}^k$ *is a reference point representing the desired aspiration levels.*

The augmented Chebyshev achievement function [14] is one of the most common achievement functions.

Definition 7. *The augmented Chebyshev achievement function is defined by*

$$s_\infty(\mathbf{z}|\mathbf{z}^{\text{ref}}) = \max_{i=1,\ldots,k}\{\lambda_i(z_i - z_i^{\text{ref}})\} + \rho\sum_{i=1}^{k}\lambda_i(z_i - z_i^{\text{ref}}), \tag{2}$$

where $\mathbf{z}^{\text{ref}}$ *is a reference point,* $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_k]$ *is a vector of weights such that* $\forall i \ \lambda_i \geq 0$ *and, for at least one* i*,* $\lambda_i > 0$*, and* $\rho > 0$ *is a sufficiently small augmentation coefficient.*

3 Related Work

In the current literature, some alternative preference relations have been used to deal with MOPs. However, the optimal solution set induced by these preference relations is a subset of PF_{opt}. As a consequence, when one of these preference relations is applied, for example, on the current population of a MOEA, the optimal solutions regarding the alternative preference relation would belong to a portion of PF_{opt}. Thus, some parts of the Pareto front will not be generated.

Among the alternative preference relations that have been proposed we can find the following. The Average Ranking and Maximum Ranking relations [15] which have the drawback of favoring extreme solutions. These preference relations have been used in [16] to deal with MOPs. Drechsler et al. [17] proposed the *favour relation* which also emphasizes extreme solutions.

The Preference Order Relation, developed by di Pierro [18], compares two solutions by discarding objectives until one of them dominates the other. The disadvantage of this approach is its high computational cost.

Sato et al. [19] proposed a preference relation to control the dominance area of a solution. This relation emphasizes solutions in the middle region of the Pareto front.

4 Solving MOPs Using an Alternative Preference Relation

In this section we first present the Chebyshev preference relation introduced in [12] and we describe how to use this relation to approximate the entire Pareto front.

4.1 The Chebyshev Preference Relation

The Chebyshev preference relation combines the Pareto dominance relation and an achievement function to compare solutions in objective function space. First, this relation defines a Region of Interest (RoI) with respect to a given reference point. This region contains all solutions with an achievement value $s_\infty(\mathbf{z}|\mathbf{z}^{\text{ref}}) \leq s^{\min} + \delta$, where $s^{\min} = \min_{\mathbf{z}\in\mathcal{Z}} s_\infty(\mathbf{z}|\mathbf{z}^{\text{ref}})$, and δ is a threshold that determines the size of the RoI. Fig. 1 shows the RoI defined by means of the achievement function. Solutions in this region are compared using the usual Pareto dominance relation, while solutions outside of the RoI are compared using their achievement function value.

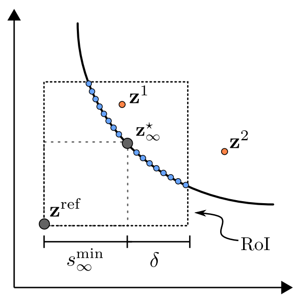

Fig. 1. Nondominated solutions with respect to the Chebyshev relation.

The Chebyshev preference relation is formally defined as follows:

Definition 8. *A solution* $\mathbf{z}^1$ *is preferred to solution* $\mathbf{z}^2$ *with respect to the Chebyshev relation (*$\mathbf{z}^1 \prec_{\text{ch}} \mathbf{z}^2$*), if and only if:*

1. $s_\infty(\mathbf{z}^1|\mathbf{z}^{\text{ref}}) < s_\infty(\mathbf{z}^2|\mathbf{z}^{\text{ref}}) \ \wedge \{\mathbf{z}^1 \notin R(\mathbf{z}^{\text{ref}},\delta) \vee \mathbf{z}^2 \notin R(\mathbf{z}^{\text{ref}},\delta)\}$*, or,*
2. $\mathbf{z}^1 \preceq_{\text{par}} \mathbf{z}^2 \wedge \{\mathbf{z}^1, \mathbf{z}^2 \in R(\mathbf{z}^{\text{ref}},\delta)\}$,

where $R(\mathbf{z}^{\text{ref}},\delta) = \{\mathbf{z} \,:\, s_\infty(\mathbf{z}|\mathbf{z}^{\text{ref}}) \leq s^{\min} + \delta\}$ *is the Region of Interest with respect to a given reference point* $\mathbf{z}^{\text{ref}}$*.*

As an illustration of the preference relation, consider solutions $\mathbf{z}^1$ and $\mathbf{z}^2$ presented in Fig. 1. Since $\mathbf{z}^2 \notin R(\mathbf{z}^{\text{ref}},\delta)$ and $s_\infty(\mathbf{z}^1,\mathbf{z}^{\text{ref}}) < s_\infty(\mathbf{z}^2,\mathbf{z}^{\text{ref}})$, then $\mathbf{z}^1 \prec_{\text{ch}} \mathbf{z}^2$.

Since, in general, the objective ranges of PF_{opt} might be different, the weight vector $\boldsymbol{\lambda}$ (Eq. 2) is used for normalizing each objective function. The weights are set as $\lambda_i = 1/(z_i^{\text{nad}} - z_i^\star)$, for all $i = 1,\ldots,k$. As the ideal and nadir points are not usually known in advance, these values are approximated using the current PF_{apx}. In order to approximate these bounding points, the Chebyshev relation always considers extreme solutions as nondominated in order to keep them in the population. This way, the approximation of the bounding points can be improved during the course of the search. To approximate $\mathbf{z}^\star$, the following set must be updated at each generation: $\Phi = \{\mathbf{z}^1, \ldots, \mathbf{z}^k \mid \mathbf{z}^i = \arg\min_{\mathbf{z}\in PF_{\text{apx}}}(z_i)\}$. That is, the solutions having the best value for each objective. The approximation of the ideal point is then $\dot{\mathbf{z}}^\star = \{z_i^i, \ldots, z_k^k\}$ with $\mathbf{z}^i \in \Phi$. Similarly, to approximate $\mathbf{z}^{\text{nad}}$, the following set is computed: $\Theta = \{\mathbf{z}^1, \ldots, \mathbf{z}^k \mid \mathbf{z}^i = \arg\max_{\mathbf{z}\in PF_{\text{apx}}}(z_i)\}$.

Thus, the normalized Chebyshev relation is defined by:

Definition 9. *A solution* $\mathbf{z}^1$ *is preferred to* $\mathbf{z}^2$ *with respect to the normalized Chebyshev preference relation (*$\mathbf{z}^1 \prec_{\text{n-ch}} \mathbf{z}^2$*) if and only if:* $\mathbf{z}^1 \prec_{\text{ch}} \mathbf{z}^2$*, and* $\mathbf{z}^2 \notin \{\Phi \cup \Theta\}$*.*

Additionally, the threshold δ can be normalized using the current range of the achievement function. Thus, the user can provide a normalized $\delta' \in [0, 1]$, and the actual value used for computing the Chebyshev relation is $\delta = \delta' \cdot (s^{\max} - s^{\min})$, where $s^{\max} = \max_{\mathbf{z}\in PF_{\text{apx}}} s_\infty(\mathbf{z}|\mathbf{z}^{\text{ref}})$ and $s^{\min} = \min_{\mathbf{z}\in PF_{\text{apx}}} s_\infty(\mathbf{z}|\mathbf{z}^{\text{ref}})$.

In order to incorporate the (normalized) Chebyshev relation into a MOEA we only have to replace the usual Pareto dominance checking procedure by the procedure that implements the new relation.

4.2 Using the Chebyshev Relation to Approximate the Entire Pareto Front in Many-objective Problems

Although the Chebyshev relation was proposed to guide the search towards a subset of PF_{opt}, in this section we propose the use of this relation to approximate the entire range of the Pareto front.

As previously mentioned, the Chebyshev relation ranks solutions outside the region of interest using the achievement function. This way it can help to rank solutions considered as incomparable by the Pareto dominance relation. In order to approximate the entire Pareto front we used as reference point the approximation of the ideal point maintained by the Chebyshev relation. In addition, we adopted a threshold $\delta' = 0.9$, comparing this way most of the solutions using Pareto dominance, while solutions far from the current PF_{apx} will be compared using their achievement function value. The basic idea is to use a stringent criterion for solutions far from the Pareto front for guiding the solutions towards the ideal point, and when the solutions are near to the Pareto front, then we use Pareto dominance to cover the entire Pareto front.

Furthermore, since the Chebyshev relation preserves the vectors that generate the approximations of $\mathbf{z}^\star$ and $\mathbf{z}^{\text{nad}}$ in the current population, the extreme solutions of the Pareto front will be found.

Additionally, since the relation used inside RoI is not essential for the mechanism of the Chebyshev relation, a different preference relation can be used as the second criteria. In this paper we investigate the performance of two preference relations derived from the additive ϵ-indicator [20]:

$$I_\epsilon(A, B) = \inf_{\epsilon\in\mathbb{R}}\{\forall \mathbf{z}^2 \in B\ \exists \mathbf{z}^1 \in A : z_i^1 \leq \epsilon + z_i^2 \text{ for } i = 1, \ldots, k\},$$

where A and B are two nondominated sets. In other words, $I_\epsilon(A, B)$ is the minimum ϵ value such that added to any vector in B, then $A \preceq B$. As shown in [21], the ϵ-indicator is dominance preserving since if $\mathbf{z}^1 \prec_{\text{par}} \mathbf{z}^2$, then $I_\epsilon(\{\mathbf{z}^1\}, \{\mathbf{z}^2\}) < I_\epsilon(\{\mathbf{z}^2\}, \{\mathbf{z}^1\})$.

In order to use the information provided by the ϵ-indicator we need to define a function for measuring the performance of a solution $\mathbf{z}^1 \in P$ with respect to the members in the population P. In this paper we have adopted two functions for this purpose. The first function uses the minimum value of $I_\epsilon(\{\mathbf{z}^2\}, \{\mathbf{z}^1\})$ among every $\mathbf{z}^2$ in the current population. That is, $F_\epsilon^{\min}(\mathbf{z}^1) = \min_{\mathbf{z}^2 \in P \setminus \{\mathbf{z}^1\}} I_\epsilon(\{\mathbf{z}^2\}, \{\mathbf{z}^1\})$. This function is also known as *maximin* fitness function[4] [22].

The second fitness function was proposed by Zitzler and Künzli [21] and it is defined by $F_\epsilon^{\text{sum}}(\mathbf{z}^1) = \sum_{\mathbf{z}^2 \in P \setminus \{\mathbf{z}^1\}} -\exp(-I_\epsilon(\{\mathbf{z}^2\}, \{\mathbf{z}^1\})/(c \cdot \kappa)$, where c is a normalizing factor given by $c = \max_{\mathbf{z}^1, \mathbf{z}^2 \in P} |I_\epsilon(\{\mathbf{z}^2\}, \{\mathbf{z}^1\})|$, and κ is a scaling factor that regulates the influence of the dominating solutions over dominated ones. In our computations we used $\kappa = 0.05$ since this value yielded good results in [21].

Using these different fitness functions we can define appropriate preference relations in order to integrate them into the Chebyshev preference relation.

Definition 10. *A solution* $\mathbf{z}^1$ *is preferred to solution* $\mathbf{z}^2$ *with respect to the* I_ϵ^{sum}*-relation (*$\mathbf{z}^1 \prec_\epsilon^{\text{sum}} \mathbf{z}^2$*), if and only if:* $F_\epsilon^{\text{sum}}(\mathbf{z}^1) > F_\epsilon^{\text{sum}}(\mathbf{z}^2)$.

Definition 11. *A solution* $\mathbf{z}^1$ *is preferred to solution* $\mathbf{z}^2$ *with respect to the* $I_\epsilon^{\min}$*-relation (*$\mathbf{z}^1 \prec_\epsilon^{\min} \mathbf{z}^2$*), if and only if:* $F_\epsilon^{\min}(\mathbf{z}^1) > F_\epsilon^{\min}(\mathbf{z}^2)$.

5 Experimental Evaluation and Analysis

In this section, we analyze the Chebyshev relation coupled with each preference relation derived from the ϵ-indicator, i.e., solutions outside the RoI are compared using their achievement value, while solutions inside the RoI are compared employing the relations $I_\epsilon^{\min}$ or I_ϵ^{sum}, respectively.

5.1 Algorithms and Parameter Settings

The experiments presented in this section were designed with two goals in mind. First, to investigate whether the Chebyshev relation is able to improve the scalability of Nondominated Sorting Genetic Algorithm II (NSGA-II) w.r.t. the number of objectives. Secondly, to analyze the effect of DRSs on the performance of Pareto-based MOEAs.

For the first goal, we compare the performances of NSGA-II using three different preference relations, namely: usual Pareto dominance, Chebyshev relation with I_ϵ^{sum}, and Chebyshev relation with $I_\epsilon^{\min}$. We evaluated the cases with 3, 4, 6, 8, 10, 12 and 14 objectives.

[4] Since the maximin fitness is to be minimized, the value $-F_\epsilon^{\min}(\mathbf{z}^1)$ is used instead.

The Chebyshev relation relies on two key elements: the evaluation of the solutions far from the Pareto front using the achievement function, and the approximation of the ideal point and the nadir point. Therefore, the selection of the test problems was made in order to evaluate whether the pressure selection biased towards solutions near the ideal point might lead to premature convergence in problems with several local Pareto fronts. Besides, we want to test the quality of the approximation of the bounding points in problems with disconnected Pareto fronts and different objective ranges.

We adopted 7 test problems presented in Table 1 taken from the DTLZ [10], and WFG [11] test suites . The variables of these problems are divided in position-related and distance-related parameters.

For the second goal of the experiments we kept the same number of distance-related variables for any number of objectives in order to isolate the effect of the number of objectives, namely, $k-1$ position-related variables and we fixed the number of distance-related variables to 5 for DTLZ1, and for the other test problems to 20. Similarly, we carried out the same number of function evaluations in every problem in order to observe variations in performance when more objectives are added. In Table 2, we can see the standard parameter values used for NSGA-II. For all the configurations we carried out 30 runs for each MOEA. The results presented were averaged over the total of this number of runs.

Problem	Features
DTLZ1, DTLZ3	Multiple local Pareto fronts.
DTLZ4	Nonuniform solution density.
DTLZ7, WFG2	Disconnected PF_{opt}.
WFG6	Nonseparable MOP.

Table 1. Adopted MOPs.

Parameter	Value
Population size	200
Generations	200
Crossover rate	0.9
Mutation rate	$1/n$
Crossover index	20
Mutation index	20

Table 2. NSGA-II parameters.

Another reason for our selection of MOPs is that the generational distance (GD) can be computed without the need of having a discrete representation of the Pareto optimal front. For these problems we took advantage of their geometrical shape or their known Pareto optimal set.

For computing GD for DTLZ1 we used $GD = (\|\mathbf{z}\|/|P|) - 0.5$ since its Pareto front is a hyperplane that intersects each axis in 0.5, while for DTLZ2, DTLZ3 and DTLZ4 we used $GD = (\|\mathbf{z}\|_2/|P|) - 1$ since its Pareto front is a sphere of radius 1. In DTLZ7, we used the value of the auxiliary function $g(\mathbf{x}) \geq 1$ (see [10] for details). The Pareto optimal front of DTLZ7 is achieved when $g(\mathbf{x}) = 1$. Thus, we use this function to compute a variant of GD, defined by $GD_g = g(\mathbf{x}) - 1$. Since the optimal solutions of WFG2 and WFG6 are those for which the distance-related variables are equal to 0.35, we adopted another variant of GD, denoted by $GD_{\mathbf{x}}$, which measures distance in decision variable space. For the sake of clarity, in the following discussion we refer to all these variants just as GD.

Additionally, to evaluate distribution we employed the inverted generational distance (IGD). As reference set, we used the nondominated set resulting from the union of all the PF_{apx} sets generated in the experiments for each problem.

In order to directly compare the performance of the MOEAs we used the additive ϵ-indicator previously presented. Roughly speaking, A is better than B if $I_\epsilon(A,B) < I_\epsilon(B,A)$.

5.2 Discussion of the results

From observing the GD values obtained (Table 3 and Fig. 4) we can confirm that the convergence ability of the original NSGA-II deteriorates as the number of objectives is increased. In contrast, when the Chebyshev relation is employed the performance is degraded by some small degree. In particular, the performance achieved by using I_ϵ^{sum}-relation or $I_\epsilon^{\min}$-relation is very similar in most of the test problems. Only on DTLZ1 (Fig. 4) we can see that $I_\epsilon^{\min}$-relation achieved a bad GD on some objectives. This suggests that $I_\epsilon^{\min}$-relation can lead to get stuck in local Pareto fronts in some runs. The results obtained using the ϵ-indicator confirm that the performance of NSGA-II is greatly improved by introducing the Chebyshev relation (see Fig. 5 for problem DTLZ1). Although not shown here, the results for the other DTLZ problems showed a tendency similar to that of DTLZ1. Specifically, in all the DTLZ problems we observed that $I_\epsilon(\text{NSGA2-I}_\epsilon, \text{NSGA2}) < I_\epsilon(\text{NSGA2}, \text{NSGA2-I}_\epsilon)$.

With respect to the distribution, the results of IGD suggest that the Chebyshev relation was able to cover the full range of the Pareto front in all the test problems considered in this paper. In Table 4 we show a representative selection of the obtained results.

The results obtained in problems WFG2 and WFG6 deserve a more detailed analysis since according to the ϵ-indicator (Fig. 5), the incorporation of the Chebyshev relation yielded a small improvement for NSGA-II. However, by inspecting the GD values of the WFG problems (Table 3 and right plot of Fig. 4), NSGA-II's performance is not as remarkably deteriorated as we observed in the DTLZ problems, especially in problem WFG2.

By analyzing some plots and performance indicator results we hinted that the divergence problems of the Pareto-based MOEAs when the number of objectives increases was due to the so-called DRSs. Fig. 2 shows an example of DRSs generated by NSGA-II in problem DTLZ3. Although the pointed DRSs in the figure have poor values in objective f_3, for example, they are nondominated solutions because they have values close to zero in objectives f_1 and f_2.

As Figs. 2 and 3 suggest, an important source of the scalability issues observed in the DTLZ test problems might be due to the generation of DRSs. Other DTLZ test problems not included in this paper have a similar feasible search space to that of DTLZ2 or DTLZ3. Therefore, we can expect that other DTLZ test problems will also have DRSs.

In order to evaluate in a quantitative manner the effect of DRSs in problems DTLZ and WFG we suggest using the distribution of the *maximum tradeoff* of the solutions of PF_{apx}. We define the maximum tradeoff of solution $\mathbf{z}$ as $\Lambda^{\max}(\mathbf{z}) =$

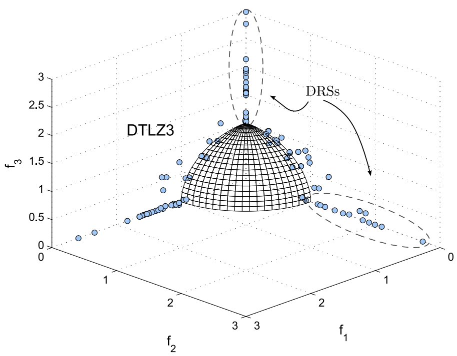

Fig. 2. Illustration of dominance resistant solutions in DTLZ3 using NSGA-II (objs. values are divided by 20 to observe the distribution wrt the Pareto front).

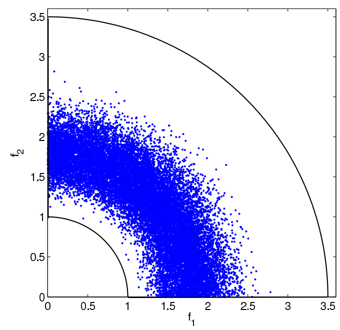

Fig. 3. Feasible objective function space of 2-objective DTLZ2 and 20 000 solutions generated at random.

$\max_{i=1,\ldots,k}(z_i)/(\min_{i=1,\ldots,k}(z_i)+1)$. By using this value, DRSs would receive a very large $\Lambda^{\max}$ value since they have in at least one objective a small value and in at least another objective a large value. It is worth noting that solutions far from the Pareto front but located in the middle region of the objective space would not obtain a large $\Lambda^{\max}$, since they have large values in all the objectives.

For computing this measure, we used the exact $\mathbf{z}^\star$ and $\mathbf{z}^{\text{nad}}$ points for normalizing the achieved PF_{apx} by each MOEA. Therefore, for every MOP, we have that $\max_{\mathbf{z}\in PF_{\text{opt}}}\{\Lambda^{\max}(\mathbf{z})\} = 1$. Any solution with $\Lambda^{\max} > 1$ is a potential DRS.

In Figs. 6–8 we show the distribution of $\Lambda^{\max}$ for the solutions generated by NSGA-II and NSGA-II with I_ϵ^{sum}-relation. For DTLZ1 (Fig. 6) we can clearly see that the proportion of DRSs not removed by NSGA-II is very high when the number of objectives is large. In contrast, by using the Chebyshev relation almost all DRSs are eliminated from the population even for 12 objectives. In the case of DTLZ2 (Fig. 7) the effect of the number of objectives on NSGA-II is more clear since the number of DRSs drastically increases with the number of objectives. WFG6 is an interesting test problem (Fig. 8), since in this case, regardless of the number of objectives, DRSs are not maintained by NSGA-II.

NSGA-II is specially sensitive to DRSs since they are spread in a very large space and, therefore, their crowding distance is larger compared to that of solutions nearby the Pareto front. As a consequence, DRSs are preferred over good solutions to compose the next generation.

On the other hand, when the Chebyshev relation is used, solutions far from the Pareto front are compared using the achievement function value. Thus, although DRSs are equally ranked by the Pareto relation, the Chebyshev relation ranks DRSs worse than other nondominated solutions located nearby the Pareto front. As a result, as it was shown in the experiments using DTLZ test problems, the Chebyshev relation can effectively discard dominance resistant solutions.

Finally, the results suggest that WFG2 and WFG6 do not induce the rise of dominance resistant solutions.

MOP	MOEA	3	4	6	8	10	12	14
DTLZ1	NSGA-II	0.0095	5.2103	248.836	359.990	406.337	425.471	436.947
		0.0058	3.3177	20.6124	17.4954	16.5937	13.9222	12.5937
	NSGA-II-I_ϵ^{sum}	**0.0034**	**0.0069**	**0.0106**	0.0122	0.0120	**0.0266**	**0.0223**
		0.0018	0.0028	0.0090	0.0121	0.0107	0.0925	0.0742
	NSGA-II-I_ϵ^{min}	0.1397	0.0265	0.0249	**0.0088**	**0.0107**	0.0444	0.0264
		0.4940	0.0897	0.0890	0.0079	0.0087	0.1201	0.0898
DTLZ2	NSGA-II	**0.0085**	0.0290	0.7404	3.3684	3.9733	4.1150	4.1956
		0.0008	0.0030	0.1456	0.2080	0.0965	0.0832	0.0657
	NSGA-II-I_ϵ^{sum}	0.0104	**0.0275**	0.0500	**0.0575**	**0.0661**	0.0735	0.0778
		0.0011	0.0026	0.0044	0.0063	0.0067	0.0094	0.0078
	NSGA-II-I_ϵ^{min}	0.0105	0.0276	**0.0482**	0.0577	0.0671	**0.0722**	**0.0759**
		0.0010	0.0027	0.0063	0.0064	0.0061	0.0082	0.0097
DTLZ3	NSGA-II	39.41	193.98	1436.66	2557.56	3058.35	3310.95	3409.70
		11.5519	29.1555	117.9495	146.3092	76.1268	68.8680	57.1938
	NSGA-II-I_ϵ^{sum}	**10.9479**	**11.0799**	**14.5876**	**17.8609**	**23.6319**	**28.3081**	29.1092
		3.4523	3.8608	4.8746	6.6429	7.9697	8.2998	12.3618
	NSGA-II-I_ϵ^{min}	55.0847	70.4200	17.1429	19.5645	24.7109	28.4519	**28.5141**
		26.6606	16.3303	4.5782	7.2080	7.9627	8.2112	9.3669
DTLZ4	NSGA-II	0.0063	0.0257	1.4083	3.9525	4.2261	4.3055	4.3426
		0.0031	0.0075	0.3612	0.1322	0.0687	0.0717	0.0562
	NSGA-II-I_ϵ^{sum}	**0.0041**	**0.0120**	0.0305	0.0337	0.0416	0.0440	0.0491
		0.0044	0.0099	0.0086	0.0073	0.0066	0.0058	0.0078
	NSGA-II-I_ϵ^{min}	0.0060	0.0190	**0.0230**	**0.0309**	**0.0377**	**0.0434**	**0.0464**
		0.0046	0.0082	0.0112	0.0081	0.0056	0.0077	0.0091
DTLZ7	NSGA-II	0.0145	0.0534	0.2565	0.8974	1.7709	2.3915	2.7311
		0.0018	0.0042	0.0297	0.1338	0.1491	0.1747	0.1816
	NSGA-II-I_ϵ^{sum}	**0.0094**	**0.0198**	0.0343	**0.0428**	0.0506	0.0694	**0.0718**
		0.0009	0.0017	0.0034	0.0033	0.0050	0.0062	0.0086
	NSGA-II-I_ϵ^{min}	0.0099	**0.0198**	**0.0332**	0.0433	**0.0484**	**0.0680**	0.0748
		0.0011	0.0023	0.0032	0.0042	0.0040	0.0078	0.0079
WFG2	NSGA-II	0.0524	0.0722	0.0999	0.1192	0.1351	0.1353	0.1280
		0.0307	0.0178	0.0157	0.0128	0.0254	0.0322	0.0278
	NSGA-II-I_ϵ^{sum}	0.0374	**0.0564**	**0.0605**	**0.0509**	0.0490	0.0451	**0.0469**
		0.0065	0.0079	0.0130	0.0124	0.0158	0.0128	0.0135
	NSGA-II-I_ϵ^{min}	**0.0372**	0.0578	0.0634	0.0514	**0.0448**	**0.0431**	0.0471
		0.0066	0.0084	0.0132	0.0101	0.0095	0.0126	0.0118
WFG6	NSGA-II	0.6961	0.7180	0.7265	0.7761	0.7454	0.8848	0.8664
		0.1967	0.1534	0.1686	0.1428	0.1921	0.1478	0.1510
	NSGA-II-I_ϵ^{sum}	0.6793	0.6545	**0.6939**	0.7310	0.7048	**0.6851**	0.7055
		0.1895	0.1457	0.1415	0.1176	0.1094	0.1157	0.1471
	NSGA-II-I_ϵ^{min}	**0.6780**	**0.6498**	0.7085	**0.6842**	**0.6774**	0.7183	**0.6884**
		0.1405	0.1318	0.1229	0.1186	0.1331	0.1218	0.1253

Table 3. Results of *GD* for 3 to 14 objectives. The first line for each MOEA is the mean of *GD* (best values are shown in bold type) and the second line, the standard deviation.

6 Conclusions and Future Work

In this paper we replaced the Pareto dominance by a new preference relation that couples the Chebyshev relation and an ϵ-indicator based relation. The new preference relation improves drastically the scalability of NSGA-II with respect

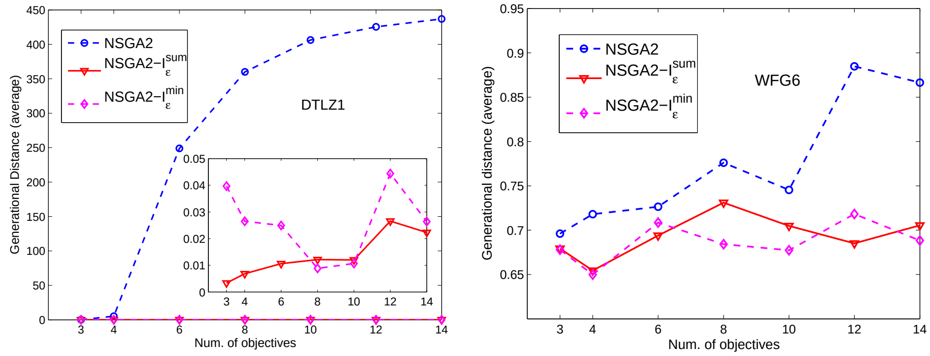

Fig. 4. GD values for DTLZ1 and WFG6 varying the number of objectives from 3 to 14.

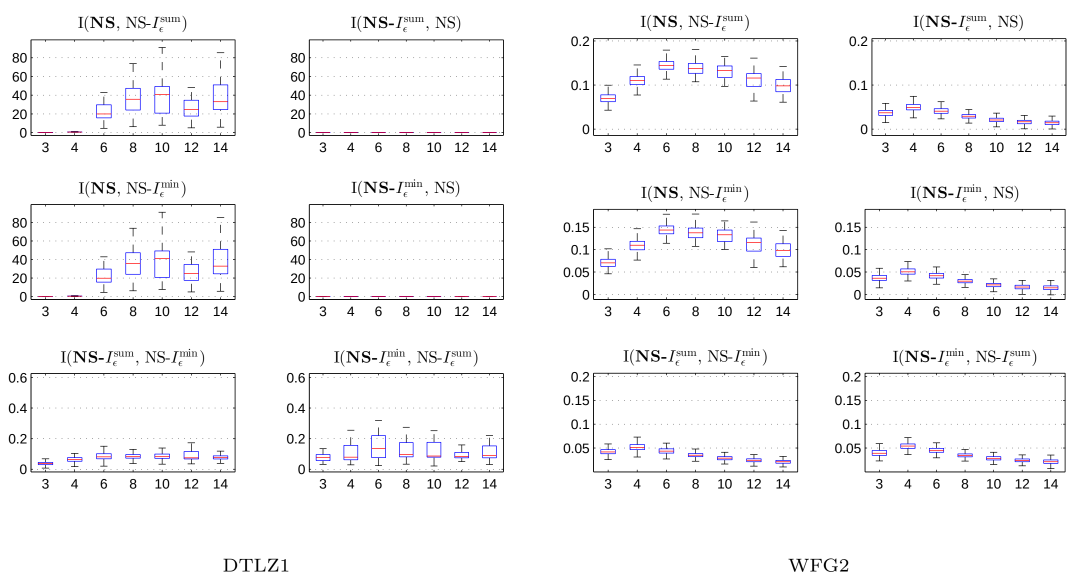

Fig. 5. Results of the ϵ-indicator for DTLZ1 and WFG2. Each subplot presents the values for 3 to 14 objectives (NS is the short for NSGA-II). Hint: A is better than B if $I_\epsilon(A, B) < I_\epsilon(B, A)$.

MOP	MOEA	3	4	6	8	10	12	14
DTLZ2	NSGA-II	0.0151	0.0163	0.0239	0.0380	0.0469	0.0421	0.0443
		0.0010	0.0010	0.0030	0.0077	0.0113	0.0111	0.0104
	NSGA-II-I_ϵ^{sum}	**0.0079**	0.0105	0.0132	**0.0148**	0.0183	0.0230	0.0331
		0.0025	0.0036	0.0047	0.0043	0.0091	0.0112	0.0185
	NSGA-II-I_ϵ^{min}	0.0087	**0.0104**	**0.0121**	0.0159	**0.0165**	**0.0189**	**0.0220**
		0.0027	0.0039	0.0028	0.0060	0.0070	0.0071	0.0099
DTLZ7	NSGA-II	0.0896	0.0759	0.3300	2.3732	12.7540	23.0939	38.0094
		0.2740	0.0071	0.0830	0.6462	4.8488	7.9811	9.5044
	NSGA-II-I_ϵ^{sum}	**0.0138**	**0.0295**	**0.2226**	1.2982	**4.9511**	**11.3958**	**9.1685**
		0.0029	0.0038	0.0083	0.0323	0.3701	0.1847	0.3093
	NSGA-II-I_ϵ^{min}	0.0150	0.0304	0.2229	**1.2806**	4.9978	11.4871	9.2333
		0.0061	0.0053	0.0073	0.0250	0.2918	0.1453	0.1852
WFG6	NSGA-II	0.0075	0.0081	0.0120	0.0146	0.0187	0.0230	0.0271
		0.0034	0.0023	0.0049	0.0047	0.0052	0.0066	0.0071
	NSGA-II-I_ϵ^{sum}	**0.0073**	**0.0075**	0.0091	0.0119	0.0141	0.0189	0.0232
		0.0024	0.0024	0.0020	0.0051	0.0038	0.0093	0.0240
	NSGA-II-I_ϵ^{min}	0.0075	0.0079	**0.0087**	**0.0114**	**0.0132**	**0.0180**	**0.0209**
		0.0015	0.0029	0.0025	0.0051	0.0050	0.0075	0.0085

Table 4. Results of *IGD* for 3 to 14 objectives. The first line for each MOEA is the mean of *IGD* (best values are shown in bold type) and the second line, the standard deviation.

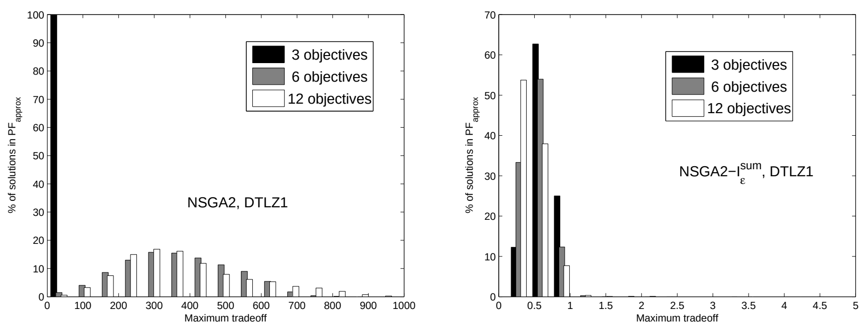

Fig. 6. Maximum tradeoff distribution for NSGA-II and NSGA-II-I_ϵ^{sum} in DTLZ1.

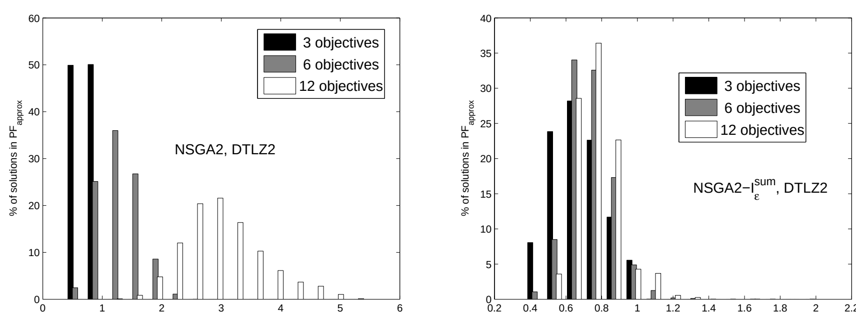

Fig. 7. Maximum tradeoff distribution for NSGA-II and NSGA-II-I_ϵ^{sum} in DTLZ2.

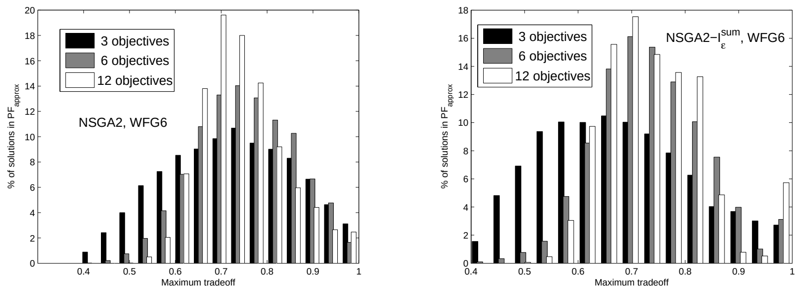

Fig. 8. Maximum tradeoff distribution for NSGA-II and NSGA-II-I_ϵ^{sum} in WFG6.

to the number of objectives. One important finding is that the main source of difficulty of DTLZ problems is the presence of dominance resistant solutions whose proportion increases with the number of objectives. Nonetheless, the Chebyshev relation was able to successfully eliminate these solutions preserving this way, the search ability in MOPs. On the other hand, since WFG problems do not induce DRSs, even the standard NSGA-II was able to maintain the same level of performance despite the number of objectives. Although these problems are hard for other reasons (e.g., nonseparability, multimodality), it seems that the number of objectives does not significantly affect their difficulty.

We are aware that there are other sources of difficulty for MOPs. However, since DRSs might be present in other problems, we suggest that the development of a MOEA integrates mechanisms to overcome these types of solutions.

In the future we plan to apply the Chebyshev preference relation in real-world problems in order to investigate if DRSs are also present. Finally, we will compare the performance of the Chebyshev relation against other optimization techniques that have shown good scalability in MOPs.

References

1. Coello Coello, C.A., Lamont, G.B., Van Veldhuizen, D.A.: Evolutionary Algorithms for Solving Multi-Objective Problems. Second edn. Springer, New York (September 2007) ISBN 978-0-387-33254-3.
2. Hughes, E.J.: Evolutionary Many-Objective Optimisation: Many Once or One Many? In: CEC'2005. Volume 1., Edinburgh, Scotland (September 2005) 222–227
3. Wagner, T., Beume, N., Naujoks, B.: Pareto-, Aggregation-, and Indicator-Based Methods in Many-Objective Optimization. In Obayashi, S., Deb, K., Poloni, C., Hiroyasu, T., Murata, T., eds.: EMO 2007, Matshushima, Japan (March 2007) 742–756
4. Purshouse, R.C., Fleming, P.J.: Evolutionary Multi-Objective Optimisation: An Exploratory Analysis. In: CEC'2003. Volume 3., Canberra, Australia (December 2003) 2066–2073
5. Knowles, J., Corne, D.: Quantifying the Effects of Objective Space Dimension in Evolutionary Multiobjective Optimization. In Obayashi, S., Deb, K., Poloni,

C., Hiroyasu, T., Murata, T., eds.: EMO 2007, Matshushima, Japan (March 2007) 757–771
6. Ishibuchi, H., Tsukamoto, N., Nojima, Y.: Evolutionary many-objective optimization: A short review. In: CEC'2008, Hong Kong (June 2008) 2424–2431
7. Schütze, O., Lara, A., Coello Coello, C.A.: On the Influence of the Number of Objectives on the Hardness of a Multiobjective Optimization Problem. IEEE Transactions on Evolutionary Computation **15**(4) (August 2011) 444–455
8. Sato, H., Aguirre, H.E., Tanaka, K.: Genetic Diversity and Effective Crossover in Evolutionary Many-Objective Optimization. In Coello Coello, C.A., ed.: LION 5, Rome, Italy, Springer. LNCS Vol. 6683 (January 17-21 2011) 91–105
9. Ikeda, K., Kita, H., Kobayashi, S.: Failure of Pareto-based MOEAs: Does Nondominated Really Mean Near to Optimal? In: CEC'2001. Volume 2., Piscataway, New Jersey (May 2001) 957–962
10. Deb, K., Thiele, L., Laumanns, M., Zitzler, E.: Scalable Multi-Objective Optimization Test Problems. In: CEC'2002. Volume 1., Piscataway, New Jersey (May 2002) 825–830
11. Huband, S., Barone, L., While, L., Hingston, P.: A Scalable Multi-objective Test Problem Toolkit. In Coello Coello, C.A., Hernández Aguirre, A., Zitzler, E., eds.: EMO 2005, Guanajuato, México, Springer. Lecture Notes in Computer Science Vol. 3410 (March 2005) 280–295
12. López Jaimes, A., Arias-Montaño, A., Coello Coello, C.A.: Preference Incorporation to Solve Many-Objective Airfoil Design Problems. In: CEC'2011, New Orleans, USA (June 2011)
13. Wierzbicki, A.: The use of reference objectives in multiobjective optimisation. In G., F., T., G., eds.: Multiple Criteria Decision Making Theory and Application. Number 177 in Lecture notes in economics and mathematical systems, Springer Verlag (1980) 468–486
14. Ehrgott, M.: Multicriteria Optimization. Second edn. Springer, Berlin (2005)
15. Bentley, P.J., Wakefield, J.P.: Finding Acceptable Solutions in the Pareto-Optimal Range using Multiobjective Genetic Algorithms. In Chawdhry, P.K., Roy, R., Pant, R.K., eds.: Soft Computing in Engineering Design and Manufacturing, London, Springer Verlag (June 1997) 231–240
16. Kukkonen, S., Lampinen, J.: Ranking-Dominance and Many-Objective Optimization. In: CEC'2007, Singapore (September 2007) 3983–3990
17. Drechsler, N., Drechsler, R., Becker, B.: Multi-Objected Optimization in Evolutionary Algorithms Using Satisfyability Classes. In Reusch, B., ed.: International Conference on Computational Intelligence, Theory and Applications, 6th Fuzzy Days, Dortmund, Germany (1999) 108–117
18. di Pierro, F., Khu, S.T., Savić, D.A.: An Investigation on Preference Order Ranking Scheme for Multiobjective Evolutionary Optimization. IEEE Transactions on Evolutionary Computation **11**(1) (February 2007) 17–45
19. Sato, H., Aguirre, H.E., Tanaka, K.: Controlling Dominance Area of Solutions and Its Impact on the Performance of MOEAs. In: EMO 2007, Matshushima, Japan (March 2007) 5–20
20. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., da Fonseca, V.G.: Performance Assessment of Multiobjective Optimizers: An Analysis and Review. IEEE Transactions on Evolutionary Computation **7**(2) (April 2003) 117–132
21. Zitzler, E., Künzli, S.: Indicator-based selection in multiobjective search. In: PPSN VIII, Springer (2004) 832–842
22. Balling, R.: The maximin fitness function: multi-objective city and regional planning. In: EMO'03. (2003) 1–15