The Importance of Diversity in the Variable Space in the Design of Multi-objective Evolutionary Algorithms

Carlos Segura[a], Joel Chacón Castillo[a], Oliver Schütze[b]

[a] *Center for Research in Mathematics (CIMAT), Computer Science Department, Callejón Jalisco s/n, Mineral de Valenciana, Guanajuato, Guanajuato 36240, Mexico*
[b] *Department of Computer Science, CINVESTAV-IPN, Mexico City, Mexico*

Abstract

Most current Multi-Objective Evolutionary Algorithms (MOEAs) do not directly manage the population's diversity in the variable space. Usually, these kinds of mechanisms are only considered when the aim is to attain diverse solutions in the variable space. This is a remarkable difference with respect to single-objective optimizers, where even when no diverse solutions are required, the benefits of diversity-aware techniques are well-known. The aim of this paper is to show that the quality of current MOEAs in terms of objective space metrics can be enhanced by integrating mechanisms to explicitly manage the diversity in the variable space. The key is to consider the stopping criterion and elapsed period in order to dynamically alter the importance granted to the diversity in the variable space and to the quality and diversity in the objective space, which is an important difference with respect to niching-based MOEAs. Specifically, more importance is given to

Email addresses: `carlos.segura@cimat.mx` (Carlos Segura), `joel.chacon@cimat.mx` (Joel Chacón Castillo), `shuetze@cs.cinvestav.mx` (Oliver Schütze)

the variable space in the initial phases, a balance that is shifted towards the objective space as the evolution progresses. This paper presents a novel MOEA based on decomposition (AVSD-MOEA/D) that relies on these principles by means of a novel replacement phase. Extensive experimentation shows the clear benefits provided by our design principle.

Keywords: Diversity, Decomposition, Multi-objective Optimization, Evolutionary Algorithms.

---

1. Introduction

Multi-Objective Evolutionary Algorithms (MOEAs) are one of the most popular approaches for dealing with Multi-Objective Optimization Problems (MOPs) [1, 2]. MOEAs are usually employed in complex problems where more traditional optimization techniques are not applicable [3]. A continuous box-constrained minimization MOP, which is the case addressed in this paper, involves two or more conflicting objectives as defined in (1)

$$\begin{aligned} min \quad & F(x) = (f_1(x), ..., f_M(x)) \\ s.t. \quad & x_i^{(L)} \leq x_i \leq x_i^{(U)} \end{aligned} \tag{1}$$

where $x \in \Re^D$, D is the number of variables, and each decision variable $x_i \in \Re$ is constrained by $x_i^{(L)}$ and $x_i^{(U)}$, i.e. the lower bound and the upper bound. The feasible space bounded by $x_i^{(L)}$ and $x_i^{(U)}$ is denoted by Ω. Each solution is mapped to the *objective space* Y with the function $F : \Omega \rightarrow Y \subseteq \Re^M$, which consists of M real-valued objective functions. The goal of most MOEAs is to find a proper approximation of the Pareto front, i.e., a set of solutions whose images are well-distributed and close to the Pareto front [4].

In recent years, the development of MOEAs has grown dramatically [5, 6], resulting in effective and broadly applicable algorithms. However, some function features provoke significant degradation of the performance of MOEAs [7], meaning that better design principles are still required. Regarding the design of MOEAs, several paths have been explored, resulting in diverse taxonomies [4]. For instance, principles related to decomposition, dominance and quality metrics are used to design MOEAs. Current state-of-the-art MOEAs consider in some way the diversity in the objective space. In some cases, this is done explicitly through density estimators [8], whereas in other cases, this is done indirectly [9]. Since optimizing most objective space quality indicators implies attaining a well-spread set of solutions in the objective space, not considering this kind of diversity would result in fairly ineffective optimizers. A quite different condition appears with respect to diversity in the variable space. Since objective space quality metrics do not consider at all the diversity in the variable space, most MOEA designers disregard this diversity.

Alternatively, several state-of-the-art single-objective methods introduce mechanisms to vary the trend of the diversity in the variable space, even if obtaining a diverse set of solutions is not the aim of the optimization [10]. Instead, this is done to induce a better balance between exploration and exploitation. In fact, the proper management of diversity is considered one of the cornerstones for proper performance [11]. Thus, these differences between the design principles applied in single-objective and multi-objective evolutionary algorithms are surprising. Moreover, practitioners have shown that modern MOEAs suffer some drawbacks involving stagnation and premature convergence in subsets of variables [12, 13, 14, 15]. As a result, this paper studies the hypothesis that incorporating mechanisms to manage the diversity in the variable space might yield important benefits to the field of multi-objective optimization. Note that unlike other proposals, such as niching-based MOEAs [16, 17], we are not interested in obtaining a diverse set of solutions in the variable space; rather, we state that the quality of the results in terms of objective-space indicators can be improved further with these kinds of mechanisms.

Since controlling diversity in the variable space is so important in single-objective domains to attain a proper balance between exploration and intensification [18], a large number of related methods have been devised [10]. Recent research on single-objective optimization has shown that important advances can be achieved when the balance between exploration and intensification is managed by relating the diversity of the population to the stopping criterion and the elapsed execution time. Specifically, these methods reduce the importance given to preserving diversity as the end of the optimization is approached. This principle has been used to find new best-known solutions for the Frequency Assignment Problem (FAP) [19], to improve further continuous optimizers [20] and to design the winning strategy in the extended round of Google Hash Code 2020[1], which featured over $100,000$ participants. Thus, we decided to explore the incorporation of this principle into the design of MOEAs.

One of the main problems in incorporating the above principle into the

[1]https://codingcompetitions.withgoogle.com/hashcode/

design of MOEAs is that measures of the variable and objective spaces must be considered simultaneously. This design principle is based on reducing the importance of diversity in the variable space as generations evolve, so we maintain this decision and indirectly increase the importance granted to diversity and quality in the objective space as the execution progresses. In order to show the validity of our hypothesis, this paper proposes the *Archived Variable Space Diversity MOEA based on Decomposition* (AVSD-MOEA/D). AVSD-MOEA/D simplifies MOEA/D-DE [21] by deactivating the dynamical resource allocation scheme and disregarding the notion of neighborhood; at the same time, it is extended by including a novel replacement strategy that applies the design principle discussed. Our proposal is validated by taking into consideration MOEA/D-DE [21], NSGA-II [22], R2-EMOA [23] and NSGA-III [24]. Remarkable benefits are achieved in terms of robustness and scalability.

The rest of this paper is organized as follows. Section 2 provides a review of MOEAs, diversity management and other related works. The AVSD-MOEA/D proposal is detailed in section 3. Section 4 is devoted to an extensive experimental validation of the novel proposal and design principle. Finally, conclusions and some lines of future work are presented in section 5.

2. Literature Review

MOEAs that take into account the diversity in the variable space are not new. These kinds of techniques have been especially useful for multimodal multi-objective optimization [25], i.e., for cases where distant solutions in terms of the variable space are desired [26, 27]. These types of MOEAs must maintain a large degree of diversity in the variable space during the entire optimization process. In contrast, our work analyzes the hypothesis that state-of-the-art techniques can be advanced further by explicitly managing the diversity in the variable space, even if the performance is measured in terms of metrics that only take into account the objective space. Since our aim is different, using multimodal multi-objective optimizers that maintain a large diversity for the entire optimization process is counter-productive, so more advanced ways of integrating diversity management are required. This section reviews some of the most important advances in diversity-aware techniques that motivated the design of AVSD-MOEA/D, which are a set of methods that were applied in single-objective optimization. Also summarized are MOEAs that in some way consider the diversity in the variable space, as well as state-of-the-art MOEAs.

2.1. Diversity-aware Single-objective Optimizers

Striking a proper balance between exploration and exploitation is one of the keys for successful single-objective EAs [18]. Several strategies for accomplishing this aim have been explored, and one of the most promising is to alter the trend in the diversity, whether explicitly or implicitly [10]. This principle has encouraged the development of a vast quantity of diversity management techniques [28]. A popular classification of these methods is based on the sort of components modified in the EA. It identifies the following groups [10]: *selection-based*, *population-based*, *crossover/mutation-based*, *fitness-based*, and *replacement-based*. Additionally, it defines *uniprocess-driven* and *multiprocess-driven* categories, depending on the number of components that are altered.

Replacement-based methods have yielded quite promising results in recent years. One of the most popular proposals belonging to this group is the *crowding* method, in which each new individual should replace similar individuals from the previous generation [29]. Several variants of this strategy have been devised, such as the *Restricted Tournament Selection* (RTS) [30]. A more recent approach that has been shown to be quite effective in a large variety of single-objective problems [19] is based on the principle of biasing the decisions made in the replacement phase by considering the stopping criterion and elapsed time. This design principle has been successful in continuous problems [20] and discrete problems [19, 31], including the problem proposed for Google Hash Code 2020, where the most effective optimizer was designed using this principle. Specifically, the initial phases of the optimization explicitly preserve a greater amount of diversity than the final phases, with a gradual shift between these behaviors. Given the success of this methodology in the single-objective case, we have adopted it to design our novel MOEA.

2.2. Diversity in the Variable Space in MOEA*s*

In spite of the vast quantity of MOEAs proposed, none of the state-of-the-art techniques involve explicitly managing the diversity in the variable space, which is a remarkable difference with respect to single-objective optimizers. One of the main reasons behind this difference is probably that there is a relationship between the diversity maintained in the objective space and the diversity maintained in the variable space, so even if the diversity in the variable space is not managed explicitly, any negative effects that do appear are usually not as bad as in single-objective optimization. However, the relationship between the two different diversities depend on each MOP [32], meaning that including the successful principles of single-objective optimizers might result in more robust MOEAs. Most current MOEAs that take the diversity in the variable space into account are devoted to multimodal multi-objective optimization [26, 27]. However, some attempts to apply these mechanisms in traditional multi-objective optimization have also been made.

The Non-Dominated Sorting Genetic Algorithm (NSGA) developed in 1995 [17] was one of the first MOEAs that employed diversity in the variable space. Specifically, it relies on fitness sharing to discriminate between solutions in the same front. In a way, this method was designed in an opposite way to current methods: the diversity in the variable space is considered, but at the cost of disregarding the information on the diversity in the objective space. The performance of NSGA is not even close to that of current MOEAs, and one of the reasons is precisely that it does not consider the diversity in the objective space.

In 2003, the GDEA [33] proposed by Toffolo and Benini integrated the diversity into the search as an additional objective, which modifies the ranking of the individuals and favors maintaining distant individuals during the entire optimization process. In 2005, Chan and Ray [34] proposed the application of selection operators to encourage the preservation of distant solutions in both the objective and variable spaces. Later, Deb and Tiwari proposed the Omni-optimizer [26]. This algorithm is designed as a generic multi-objective, multi-optima optimizer. In the multi-objective case, it is an extension of NSGA-II where the crowding distance considers both the objective and variable spaces. Since it first uses the typical rank procedure considering only the objective space, more importance is given to this space, and the diversity plays an inferior role. Unfortunately, there is no way to easily alter the importance given to each space. In 2009, Shir et al. showed that in many cases the diversity in the variable space can be significantly enhanced without hampering the convergence to a diverse Pareto front approximation. Following this insight, the CMA-ES niching framework was proposed [32]. Estimation of Distribution Algorithms have also considered the information in the variable space to attain better approximations of the Pareto set [35]. Finally, the Diversity Integrating Hypervolume-based Search Algorithm (DIVA) [36] weighs the contribution from the hypervolume against the diversity in the variable space.

Note that of the methods discussed, the only one that shows any improvements in terms of objective-space metrics is GDEA. However, the results attained by GDEA are not as competitive as those attained by modern solvers, so it is not currently considered as a state-of-the-art MOEA, but rather as an easily applicable and general approach. The remaining methods focus on showing that solutions with a higher level of diversity in the variable space can be obtained without an overly negative effect on the approximation of the Pareto front.

2.3. Multi-objective State-of-the-art Algorithms

Given the large number of MOEAs proposed in the last decades, several taxonomies have been defined [37]. Most current techniques consider in some way at least one of the following concepts [4]: Pareto dominance, indicators and/or decomposition. Note that several MOEAs use more than one of these principles, but most practitioners make the effort to classify proposals as domination-based, indicator-based or decomposition-based.

In domination-based MOEAs, the Pareto dominance relationship is applied to bias the decisions of the optimizer. Although this relationship stimulates convergence to the Pareto front, additional techniques to promote diversity in the objective space must be integrated. Several traditional MOEAs belong to this group, and while they face some difficulties for many-objective optimization, they are considered quite effective for optimization problems with two and three objectives. Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) is the most popular technique belonging to this group [22]. NSGA-III [24] extends NSGA-II by replacing the crowding selection with a strategy based on generating distributed reference points that implicitly decompose the objective space [4]. While this method is aimed at many-objective optimization, it also provides some benefits for problems with three objectives.

Indicator-based MOEAs incorporate a measure of quality over sets of solutions to alter some components, such as the selection and/or replacement [38]. Most indicators take into account both the convergence and coverage, so no additional techniques to promote diversity in the objective space are required. A highly effective MOEA that belongs to this category is the R2-Indicator-Based Evolutionary Multi-objective Algorithm (R2-EMOA) [23]. SMS-EMOA [8] is also quite effective but it uses the computationally expensive hypervolume indicator, meaning that it is not as generally applicable as R2-EMOA.

Finally, decomposition-based MOEAs [39] incorporate scalarizing functions to transform the MOP into several single-objective optimization subproblems. Those sub-problems are then solved simultaneously and collaboratively. The weighted Tchebycheff scalarizing and its variant, the achievement scalarizing function (ASF) [40, 41], have shown remarkable performance. The most popular MOEA belonging to this category is MOEA/D, which was proposed by Zhang et al. [9]. A distinctive feature of MOEA/D is the application of neighborhoods at the level of each sub-problem. The mating selection and replacement operator takes into account this neighborhood to promote collaboration between similar subproblems. MOEA/D has gained a significant popularity in the last decade, and many extensions have been devised as a result. Specifically, the MOEA/D-DE [42] — winner of some optimization competitions — provides important advances by incorporating DE operators, polynomial mutation, a dynamic computational resource allocation strategy, mating restrictions and a modified replacement operator to prevent the excessive replication of individuals.

Since practitioners compare algorithms under different conditions, it is not easy to clearly identify the most effective ones. Given the complementary properties contributed by the different methodologies discussed, our comparisons include MOEAs belonging to each category. Specifically, NSGA-II, NSGA-III, R2-EMOA and MOEA/D-DE are the set of state-of-the-art algorithms used to validate our proposal.

3. Proposal

This section describes our proposal, the *Archived Variable Space Diversity* MOEA *based on Decomposition* (AVSD-MOEA/D)[2]. The main novelty and motivation behind AVSD-MOEA/D is the incorporation of an explicit way to

[2]The source code in C++ is freely available at https://drive.google.com/drive/folders/1JAIrOybzyafxo2BUW99DnsUiqzZ3Wwco

manage diversity in the variable space, the goal being to improve the behavior in terms of objective space metrics, especially in long-term executions, which is the environment where diversity-aware techniques have excelled [43]. Although AVSD-MOEA/D is inspired by MOEA/D, it has been simplified, so in some ways it resembles more mature decomposition-based MOEAs, such as WBGA [44]. For instance, the notion of subproblem neighborhood is not used and the dynamic resource allocation usually applied in modern variants of MOEA/D is deactivated. The main reason for the simplification is to show that even a simple MOEA incorporating our design principles can improve further more complex state-of-the-art algorithms.

Our proposal decomposes the MOP into several single-objective problems. Notwithstanding that any scalarization approach can be employed, our strategy applies the ASF, which has reported some of the most effective results in recent years [40]. Let $\lambda_1, ..., \lambda_N$ be a set of weight vectors and z^* a reference vector. The MOP is decomposed into N scalar optimization sub-problems as shown in (2).

$$g^{asf}(x|\lambda_j, z^*) = \max_{1 \leq i \leq M} \left\{ \frac{|f_i(x) - z_i^*|}{\lambda_{j,i}} \right\} \tag{2}$$

The main novelty of AVSD-MOEA/D appears in the survivor selection scheme. In keeping with some of the most successful single-objective diversity-aware algorithms [19], the replacement strategy relates the degree of diversity in the variable space to the stopping criterion and elapsed generations. The aim of this relationship is to gradually bias the search from exploration to exploitation as the optimization evolves. Specifically, diversity is explicitly promoted less and less until half of the total generations. Then, in the remain-

Algorithm 1 Main procedure of AVSD-MOEA/D

Input: N (Population and Archive Size), λ (a set of N weight vectors for decomposition in the population), Λ (a set of weight vectors for the R2-based archive), D_I (Initial Penalty Threshold), CR (Crossover Rate), F (Scale Factor), p_m (mutation probability)
Output: A^t (Archive with N solutions)
Initialization: Generate an initial population P^0 with N individuals
Evaluation: Evaluate each individual in P^0 and assemble the reference vector z^* with the best objective values
Archive Initialization: $A^0 = P^0$
Assign $t = 0$
while (not stopping criterion) **do**
    for each individual $P_i^t \in P^t$ **do**
        DE variation: Generate solution Q_i^t by applying DE/rand/1/bin with F as the mutation scale factor and CR as the crossover rate, using P_i^t as the target vector
        Mutation: Apply polynomial mutation to Q_i^t with probability p_m
        Evaluation: Evaluate the new individual Q_i^t and update the reference vector z^* with the best objective values.
    Survivor selection: Generate P^{t+1} by applying the replacement scheme described in Algorithm 3, using P^t, Q^t, λ, z^* and D_I as input
    Update Archive: Create A^{t+1} by applying Algorithm 2 using A^t, Q^t, Λ, and z^* as input.
    $t = t + 1$
return A^t

ing generations, AVSD-MOEA/D behaves similarly to most popular MOEAs, i.e. the diversity in the variable space is not considered explicitly.

The main procedure of AVSD-MOEA/D is shown in Algorithm 1. Its general template is quite standard. The variation step is similar to those used in typical MOEAs, meaning it is based on crossover and mutation and any operators might be used. Specifically, N individuals are created in each generation by randomly selecting at each step three individuals to apply the $DE/rand/1/bin$ operator. Note that each member of the population is used as the target vector once. Then, polynomial mutation is applied to the output of the DE operator. As in most current MOEA/D variants, the initial population is generated randomly, the number of weight vectors ($|\lambda|$) is equal to the population size, and the reference vector z^* used for ASF consists of the best objective values achieved. The weight vectors used in this paper are based on an uniform design strategy and are specified in the experimental validation section. Finally, the survivor selection stage is applied. This is quite different from traditional techniques, in the sense that P^t and Q^t are merged, meaning that unlike in MOEA/D, the position of each individual is not important, and a diversity-aware selection is performed. Since this is the novelty of the paper, its working operation is given in detail.

Note that AVSD-MOEA/D incorporates an external archive to store the best solutions. While in MOEA/D this is considered optional, in our approach it is quite important because the penalty approach included to pick up the survivors of the population in the replacement phase might discard elite solutions for some weight vectors. Since methods based on the R2 indicator [23] have reported quite high-quality solutions, our archive is based on

Algorithm 2 Procedure to update the R2-based archive

1: Input: A^t (External archive in the current generation), Q^t (Offspring of current generation), Λ (weight vectors for $R2$), z^* (Reference vector)
2: Output: A^{t+1}
3: $R^t = A^t \cup Q^t$
4: $A^{t+1} = \emptyset$
5: **while** $|A^{t+1}| < N$ **do**
6: $\quad \forall x \in R^t : r(x) = R2(A^{t+1} \cup \{x\}; \Lambda, z^*)$
7: $\quad x^* = argmin(r(x) : x \in R^t)$
8: $\quad A^{t+1} = A^{t+1} \cup \{x^*\}$
9: $\quad R^t = R^t \setminus \{x^*\}$

the R2 indicator applying the ASF and the weights Λ to generate the utility functions. In each iteration, the archive selects N candidate solutions by combining its contents with the offspring of each generation (see line 13 in Algorithm 1). This is done by following a simple greedy approach (Algorithm 2). Specifically, it iteratively selects the individual that minimizes the $R2$ (lines 6-9) considering the individuals already selected with ties broken randomly. Note that since $R2$ is weakly Pareto-compliant, it might happen that some non-dominated individuals do not contribute to $R2$, so incorporating more complex archiving strategies might be helpful. However, our preliminary experiments have shown that this it not overly important for proper performance, so we decided to maintain its simplicity.

3.1. Novel Replacement Phase of AVSD-MOEA/D

The purpose of the replacement phase (see Algorithm 3) is to select the set of survivors of the next generation. The survivor selection described in this

Algorithm 3 Replacement Phase of AVSD-MOEA/D

1: Input: P^t (Population of current generation), Q^t (Offspring of current generation), λ (a set of weight vectors), z^* (Reference vector), and D_I (Initial Penalty Threshold).
2: Output: P^{t+1}
3: $R^t = P^t \cup Q^t$
4: $P^{t+1} = \emptyset$
5: $Penalized = \emptyset$
6: $r\lambda = \lambda$
7: $D^t = D_I - D_I * \frac{G_{Elapsed}}{0.5 * G_{End}}$
8: **while** $|P^{t+1}| < N$ **do**
9:     Compute DCS in R^t using P^{t+1} as reference set
10:     Move the individuals in R^t with $DCS < D^t$ to $Penalized$
11:     **if** R^t is empty **then**
12:         Compute DCS of each individual in $Penalized$ set employing P^{t+1} as reference set
13:         Move the individual in $Penalized$ with the largest DCS to R^t
14:     Identify the pair of non-penalized individual R_i^t and weight vector $r\lambda_j$ with the best scalarizing function value according to $g^{asf}(R_i^t|r\lambda_j, z^*)$
15:     Move individual R_i^t to P^{t+1}
16:     Erase the weight vector rλ_j from rλ
17: **return** P^{t+1}

work incorporates similar design principles to those applied in the diversity-aware single-objective optimizer DE-EDM [20]. It operates as follows. First, the parent and offspring populations are merged in a multi-set to establish the candidate set R^t (line 3). A key of the scheme is to promote the selection of individuals with a large enough contribution to diversity in the variable space. Specifically, the contribution of an individual x is calculated as $\min_{p \in P^{t+1}} Distance(x, p)$, where P^{t+1} is the multi-set of the already picked survivors and the normalized Euclidean distance specified in (3) is applied. Note that in the pseudocode, the tag DCS (Distance to Closest Solution) is used to denote the contribution to diversity.

$$Distance(A, B) = \left(\frac{1}{D} \sum_{i=1}^{D} \left(\frac{A_i - B_i}{x_i^{(U)} - x_i^{(L)}} \right)^2 \right)^{1/2} \tag{3}$$

In order to promote the selection of distant individuals, a threshold D^t is dynamically calculated (line 7) and individuals with a DCS value lower than the threshold are considered as undesirable individuals. Note that the calculation of D^t depends on an initial penalty threshold (D_I), which is a parameter of our proposal, on the number of generations that have evolved ($G_{Elapsed}$) and on the stopping criterion (G_{End}), i.e., the number of generations to evolve. Specifically, the value is decreased linearly as generations evolve. Since survivors with larger DCS values provoke exploration steps, while survivors with short DCS values promote intensification steps, this linear decrease promotes a gradual transition from exploration to exploitation. Also note that after 50% of the total number of generations, the D^t value is below 0, meaning that no penalties are applied and a more traditional strategy focused only on the objective values is used to perform the selection steps.

The strategy iteratively selects an individual from the candidate set (R^t) to enter the new population (P^{t+1}) until it is filled with N individuals (lines

8-16). In particular, the aim is to select a proper individual for each weight vector, while at the same time fulfilling the condition imposed for the contribution to diversity in the variable space. In order to satisfy this last condition, non-selected individuals with a DCS lower than D^t are moved from R^t to the *Penalized* set (lines 9-10), and in each iteration an individual belonging to R^t is selected to survive. The set of weight vectors considered by our strategy are initially placed in $r\lambda$. In each iteration, the individual in R^t with the best scalarizing function for any of the weight vectors in $r\lambda$ is identified (line 14). Then, this individual is selected as a survivor (line 15) and the weight vector used is erased from $r\lambda$ (line 16). Note that N individuals are selected, meaning that each weight vector is used to select exactly one individual. Also note that it might happen that R^t is empty prior to selecting N individuals. This means that the diversity is lower than expected, so in order to increase the amount of exploration, the individual with the largest DCS value in the *Penalized* set is selected to survive (lines 11 - 13).

4. Experimental Validation

In this section, we provide the validation of our proposal against state-of-the-art MOEAs and explain the reasons behind the superiority of AVSD-MOEA/D. Since methods that include strategies to explicitly delay convergence usually require additional computational resources to excel, long-term analyses are presented. In order to draw proper conclusions, four experiments were carried out. First, a comparison between AVSD-MOEA/D and four state-of-the-art MOEAs is presented. This comparison focuses on showing the benefits of AVSD-MOEA/D for benchmarks configured in standard ways. Second, an analysis to test the scalability on the number of decision variables is carried out. Third, the robustness of AVSD-MOEA/D in terms of the initial penalty threshold (D_I) is analyzed. Finally, the ability of dealing with different levels of difficulty in biased test problems is studied. Note that in order to test the quality of the approximations, the hypervolume is used. Additionally, our analyses also include some studies to better understand the implications of AVSD-MOEA/D on the diversity in the variable space. These analyses provide a better understanding of the reasons behind the proper performance of AVSD-MOEA/D and highlight the significant differences between AVSD-MOEA/D and other strategies in terms of the dynamics of the population.

In the first three analyzes our validation takes into account three of the most popular benchmarks in multi-objective optimization: WFG [7], DTLZ [45], and UF [46]. The remaining one experiment takes into the BT problems [47]. Unless otherwise stated, standard configurations were used. Specifically, the WFG test problems were used with two and three objectives with 24 parameters[3], 20 of them corresponding to distance parameters and 4 to position parameters. The DTLZ test problems were also used with two and three objectives, and the number of variables in each case was set to $D = M + r - 1$, where $r = \{5, 10, 20\}$ for DTLZ1, DTLZ2 to DTLZ6 and DTLZ7, respectively. The UF benchmark comprises seven problems with two objectives (UF1-7) and three problems with three objectives (UF8-10). Similarly,

[3]In the WFG context, the term *parameter* is equivalent to variable.

Table 1: Parameterization of the variation phase applied in each MOEA

	2 objectives		3 objectives	
	CR	F	CR	F
AVSD-MOEA/D	0.0	0.75	0.0	0.75
MOEA/D-DE	0.75	0.75	0.5	0.5
R2-EMOA	0.75	0.5	0.5	0.5
NSGA-II	0.75	0.5	0.0	0.25
NSGA-III	0.75	0.25	0.5	0.75

the BT benchmark includes eight problems with two objectives (BT1-8) and one problem with three objectives (BT9). These two last groups of problems were configured with 30 variables.

Regarding our comparisons, the set of state-of-the-art MOEAs used to validate our proposals is comprised of four popular and complementary MOEAs: NSGA-II [22], MOEA/D-DE [21], R2-EMOA [23] and NSGA-III [24]. Given that all the algorithms are stochastic, each execution was repeated 35 times in every experiment. The hypervolume indicator (HV) is used to compare the various schemes. The reference point used to calculate the HV is chosen to be a vector whose values are sightly larger (ten percent) than the nadir point, as suggested in [48]. The value reported is computed as the ratio between the normalized HV attained [49] and the maximum attainable normalized HV. This way, a value equal to one means a perfect approximation. Note that this value is not attainable because MOEAs yield discrete approximations. Finally, to statistically compare the HV ratios, a guideline similar to that proposed in [50] was used, which entails the use of the Shapiro-Wilk, Levene, ANOVA, Welch and Kruskal-Wallis tests. An algorithm $A1$ is said to beat an algorithm $A2$ when the differences between the HV ratios attained are statistically significant, and the mean and median HV ratios obtained by $A1$ are higher than the mean and median achieved by $A2$.

An important step to perform fair comparisons is the parameterization of algorithms. Note that the variation operators used in each algorithm in their original variants differ. Using the original variation operators to perform comparisons is not fair, and probably would offer more conclusions about the effectiveness of the operators than about the general framework proposed in each MOEA. However, there might also be a dependency between the general framework and the proper variation operators. We thus decided to use a common simple framework for the variation step, but to allow a different parameterization for each algorithm. Specifically, the variation phase first applies the classic DE scheme known as DE/rand/1/bin with parameters F and CR, and then it applies polynomial mutation with probability p_m and a distribution index equal to 50. Note that the use of the additional mutation in variants based on MOEA/D is quite important [21]. The additional common parameter is the population size. Since the hypervolume is highly dependent on the number of solutions used to approximate the Pareto front, all the MOEAs were configured with a common population size equal to 100 individuals.

In order to set the CR, F and p_m parameters, 40 parameterizations were tested for each algorithm. They were generated by combining four values of F (0.25, 0.5, 0.75 and 1.0), five values of CR (0.0, 0.25, 0.5, 0.75, 1.0) and two values of p_m (0.0, $\frac{1}{D}$). These configurations were executed by setting the stopping criterion to 2.5×10^6 function evaluations, with all the aforementioned benchmarks. The mean of the resulting hypervolume ratios were calculated independently for the problems with two and three objectives. Then, in the experiments that followed, the parameter configuration that attained the largest mean was used. Table 1 shows the configuration of CR and F selected for each MOEA. Note that all of them yielded better results when mutation was enabled, so the benefits reported in [21] also appeared for other MOEAs. Note that in the case of AVSD-MOEA/D, CR was set to 0, which reduces the strength of the perturbation performed by DE. Since AVSD-MOEA/D maintains a larger degree of diversity than other methods, low disruptive operators seem to be more helpful.

Note also that there are some additional parameters that are specific to some of the MOEAs. They were set to typical values used in literature. Table 2 shows this additional parameterization. Note also that scalarization functions are used in MOEA/D-DE, R2-EMOA, NSGA-III and AVSD-MOEA/D. In all those cases, the ASF approach is used. However, the weight vectors employed in R2-EMOA are different from those in the remaining algorithms because in R2-EMOA, using a larger number of weight vectors than the population size is beneficial. As in the official code, R2-EMOA was applied with 501 and 496 weight vectors for two and three objectives, respectively [23]. In the remaining cases — including AVSD-MOEA/D — the number of weight vectors was equal to the population size, and they were generated with the uniform design strategy described in [51]. Note that in the case of AVSD-MOEA/D, a second set of weight vectors (Λ) was considered for the external archive. Since the archive is based on $R2$, it considers the same weight vectors as R2-EMOA.

Table 2: Configuration of the specific parameters of each MOEA

Algorithm	**Configuration**
MOEA/D-DE	Max. updates by sub-problem (η_r) = 2, tour selection = 10, neighbor size = 20, period utility updating = 50 generations, local selection probability (δ) = 0.9
R2-EMOA	$\rho = 1$, offspring by iteration = 1
AVSD-MOEA/D	$D_I = 0.4$

4.1. Performance of MOEA*s in long-term executions*

One of the aims behind the design of AVSD-MOEA/D is to profit from long-term executions. Therefore, in this section we present the results attained by the different algorithms when setting the stopping criterion to 2.5×10^7 function evaluations. Table 3 shows the HV ratios obtained for the benchmark functions with two objectives. Note that the same results can be drawn with the IGD+ metric [52] and can be inspected in the supplementary material. For each method and problem, the best, mean and standard deviation of the HV ratio values are reported. Furthermore, in order to summarize the results attained by each method, the last row shows the mean for the whole set of problems. For each test problem, the method that yielded the largest mean and those that were not statistically inferior to the best are shown in **boldface**. Similarly, the method that yielded the best HV value among all the runs is <u>underlined</u>. From here on, the methods shown in **boldface** for a given problem are referred to as the winning methods. AVSD-MOEA/D, R2-EMOA, MOEA/D-DE, NSGA-III and NSGA-II belonged to the winning methods in 17, 6, 2, 2 and 0 problems, respectively. The superiority of AVSD-MOEA/D is clear both in terms of this metric and in terms of the mean HV. Particularly, AVSD-

Table 3: Summary of the hypervolume ratios attained for problems with two objectives

	AVSD-MOEA/D			**MOEA/D-DE**			**NSGA-II**			**NSGA-III**			**R2-EMOA**		
	Best	Mean	Std	Best	Mean	Std	Best	Mean	Std	Best	Mean	Std	Best	Mean	Std
WFG1	0.995	0.982	0.020	0.957	0.842	0.058	0.994	0.966	0.026	0.993	**0.989**	0.011	0.993	0.921	0.039
WFG2	0.999	**0.999**	0.000	0.996	0.996	0.000	0.998	0.998	0.000	0.997	0.990	0.013	0.998	0.998	0.000
WFG3	0.993	**0.993**	0.000	0.992	0.992	0.000	0.980	0.978	0.001	0.992	0.992	0.000	0.992	0.991	0.000
WFG4	0.991	**0.991**	0.000	0.988	0.988	0.000	0.979	0.975	0.002	0.988	0.986	0.003	0.988	0.973	0.007
WFG5	0.933	**0.905**	0.008	0.891	0.882	0.004	0.883	0.878	0.002	0.895	0.888	0.003	0.890	0.885	0.003
WFG6	0.959	0.922	0.020	0.988	0.963	0.019	0.977	0.974	0.001	0.956	0.934	0.013	0.991	**0.990**	0.001
WFG7	0.991	**0.991**	0.000	0.988	0.988	0.000	0.980	0.977	0.001	0.988	0.988	0.000	0.991	**0.991**	0.000
WFG8	0.963	**0.954**	0.004	0.846	0.833	0.004	0.825	0.815	0.003	0.829	0.826	0.001	0.835	0.832	0.001
WFG9	0.978	**0.976**	0.002	0.974	0.954	0.039	0.941	0.873	0.071	0.798	0.796	0.001	0.975	0.939	0.051
DTLZ1	0.993	**0.993**	0.000	0.993	**0.993**	0.000	0.992	0.991	0.000	0.993	**0.993**	0.000	0.992	0.992	0.000
DTLZ2	0.991	0.991	0.000	0.989	0.989	0.000	0.987	0.986	0.001	0.989	0.989	0.000	0.992	**0.992**	0.000
DTLZ3	0.991	0.991	0.000	0.989	0.989	0.000	0.989	0.989	0.000	0.989	0.989	0.000	0.992	**0.992**	0.000
DTLZ4	0.991	**0.991**	0.000	0.989	0.989	0.000	0.987	0.903	0.231	0.989	0.989	0.000	0.992	0.803	0.320
DTLZ5	0.991	0.991	0.000	0.989	0.989	0.000	0.987	0.986	0.001	0.989	0.989	0.000	0.992	**0.992**	0.000
DTLZ6	0.991	**0.991**	0.000	0.989	0.986	0.014	0.989	0.989	0.000	0.989	0.989	0.000	0.992	0.985	0.021
DTLZ7	0.997	**0.997**	0.000	0.996	0.996	0.000	0.997	0.996	0.000	0.996	0.996	0.000	0.997	**0.997**	0.000
UF1	0.995	**0.995**	0.000	0.987	0.986	0.001	0.989	0.988	0.001	0.992	0.991	0.001	0.992	0.992	0.000
UF2	0.995	**0.995**	0.000	0.990	0.988	0.001	0.984	0.982	0.001	0.986	0.981	0.003	0.988	0.987	0.001
UF3	0.938	0.906	0.016	0.991	**0.990**	0.001	0.988	0.985	0.004	0.985	0.968	0.019	0.991	0.982	0.005
UF4	0.979	**0.977**	0.001	0.914	0.904	0.006	0.892	0.882	0.005	0.880	0.876	0.003	0.902	0.893	0.003
UF5	0.990	**0.975**	0.009	0.715	0.439	0.137	0.792	0.734	0.087	0.777	0.654	0.067	0.792	0.733	0.092
UF6	0.962	**0.938**	0.013	0.928	0.748	0.175	0.870	0.720	0.069	0.820	0.708	0.043	0.827	0.691	0.091
UF7	0.993	**0.993**	0.000	0.991	0.990	0.001	0.980	0.976	0.002	0.983	0.975	0.002	0.992	0.982	0.006
Mean	0.983	**0.976**	0.004	0.960	**0.931**	0.020	0.956	**0.937**	0.022	0.948	**0.934**	0.008	0.960	**0.936**	0.028

MOEA/D attained a value equal to 0.976, while all the remaining methods attained values between 0.931 and 0.937. A careful inspection of the data shows that in those cases where AVSD-MOEA/D loses, the difference with respect to the best method is low. In fact, the difference between the mean HV ratio attained by the best method and by AVSD-MOEA/D is never greater than 0.1. However, in all the other methods, there were several problems where the distance with respect to the best approach was greater than 0.1. Specifically, it happened in 4, 4, 4 and 5 problems for R2-EMOA, MOEA/D-DE, NSGA-II and NSGA-III, respectively. This means that AVSD-MOEA/D wins in most cases and that when it loses, the difference is always small. Note also

Table 4: Statistical Tests and Deterioration Level of the HV ratio for problems with two objectives

	↑	↓	↔	**Score**	**Deterioration**
AVSD-MOEA/D	78	13	1	65	0.160
MOEA/D-DE	41	50	1	-9	1.181
NSGA-II	21	66	5	-45	1.057
NSGA-III	35	52	5	-17	1.119
R2-EMOA	47	41	4	6	1.066

that in terms of standard deviation, AVSD-MOEA/D yields much lower values than all the other algorithms, meaning it is quite robust.

In order to better clarify these findings, pair-wise statistical tests were applied between each method tested in each test problem. For the two-objective cases, Table 4 shows the number of times that each method statistically won (column ↑), lost (column ↓) or tied (column ↔). The **Score** column shows the difference between the number of times that each method won and the number of times that each method lost. Additionally, for each method M, we calculated the sum of the differences between the mean HV ratio attained by the best method (the ones with the highest mean) and method M, for each problem where M was not in the group of winning methods. This value is shown in the *Deterioration* column. The data confirm that although AVSD-MOEA/D loses in some pair-wise tests, the overall numbers of wins and losses clearly favor AVSD-MOEA/D. More importantly, the total deterioration is much lower in the case of AVSD-MOEA/D, confirming that when AVSD-MOEA/D loses, the differences are low.

Tables 5 and 6 shows the same information for the problems with three objectives. In this case, the number of times that each method belonged to the

Table 5: Summary of the hypervolume ratios attained for problems with three objectives

	AVSD-MOEA/D			**MOEA/D-DE**			**NSGA-II**			**NSGA-III**			**R2-EMOA**		
	Best	Mean	Std	Best	Mean	Std	Best	Mean	Std	Best	Mean	Std	Best	Mean	Std
WFG1	0.985	**0.982**	0.007	0.972	0.937	0.030	0.966	0.959	0.008	0.970	0.967	0.008	0.981	0.965	0.017
WFG2	0.991	**0.991**	0.000	0.981	0.979	0.001	0.974	0.967	0.003	0.972	0.971	0.001	0.963	0.963	0.000
WFG3	0.995	**0.994**	0.000	0.990	0.990	0.000	0.986	0.975	0.006	0.966	0.954	0.007	0.992	0.992	0.000
WFG4	0.943	**0.941**	0.001	0.899	0.898	0.001	0.892	0.876	0.008	0.897	0.897	0.000	0.911	0.906	0.002
WFG5	0.901	**0.872**	0.011	0.831	0.831	0.000	0.828	0.812	0.009	0.833	0.827	0.003	0.849	0.846	0.001
WFG6	0.912	0.888	0.011	0.887	0.862	0.013	0.851	0.822	0.013	0.880	0.858	0.012	0.902	**0.893**	0.006
WFG7	0.943	**0.942**	0.001	0.899	0.898	0.001	0.892	0.865	0.010	0.897	0.897	0.000	0.906	0.904	0.001
WFG8	0.910	**0.902**	0.003	0.816	0.812	0.003	0.759	0.748	0.006	0.810	0.807	0.002	0.827	0.824	0.001
WFG9	0.910	**0.894**	0.006	0.875	0.862	0.005	0.819	0.732	0.019	0.858	0.749	0.027	0.886	0.880	0.002
DTLZ1	0.967	**0.967**	0.000	0.953	0.953	0.000	0.950	0.941	0.004	0.953	0.953	0.000	0.942	0.941	0.001
DTLZ2	0.945	**0.944**	0.000	0.914	0.914	0.000	0.905	0.892	0.008	0.913	0.913	0.000	0.916	0.915	0.001
DTLZ3	0.945	**0.944**	0.000	0.914	0.914	0.000	0.901	0.883	0.009	0.913	0.913	0.000	0.916	0.915	0.001
DTLZ4	0.945	**0.944**	0.000	0.914	0.914	0.000	0.908	0.813	0.238	0.913	0.903	0.059	0.916	0.893	0.127
DTLZ5	0.985	0.985	0.000	0.979	0.979	0.000	0.986	0.984	0.001	0.967	0.959	0.005	0.986	**0.986**	0.000
DTLZ6	0.985	**0.985**	0.000	0.979	0.959	0.038	0.984	0.955	0.127	0.958	0.948	0.007	0.986	0.985	0.008
DTLZ7	0.970	**0.968**	0.001	0.922	0.922	0.000	0.941	0.924	0.025	0.929	0.912	0.008	0.907	0.848	0.020
UF8	0.922	**0.916**	0.003	0.891	0.862	0.032	0.747	0.695	0.035	0.890	0.835	0.101	0.893	0.877	0.016
UF9	0.957	**0.951**	0.003	0.947	0.813	0.071	0.822	0.735	0.069	0.954	0.936	0.043	0.942	0.862	0.077
UF10	0.831	**0.787**	0.041	0.681	0.435	0.147	0.543	0.483	0.084	0.624	0.458	0.127	0.579	0.561	0.042
Mean	0.944	**0.937**	0.005	0.908	**0.881**	0.018	0.877	**0.845**	0.036	0.900	**0.877**	0.022	0.905	**0.892**	0.017

winning groups were 17, 2, 0, 0 and 0 for AVSD-MOEA/D, R2-EMOA, MOEA/D-DE, NSGA-III and NSGA-II, respectively. Thus, AVSD-MOEA/D yielded quite superior results. Considering the whole set of problems, AVSD-MOEA/D obtained a much larger mean HV ratio than the other ones. Moreover, the difference between the mean HV ratio obtained by the best method and by AVSD-MOEA/D was never greater than 0.1. However, all the other methods exhibited a deterioration in excess of 0.1 in several cases. In particular, this happened in 2, 2, 2 and 6 problems for MOEA/D-DE, R2-EMOA, NSGA-III and NSGA-II respectively. Remarkably, AVSD-MOEA/D is quite superior in both the total deterioration and in the score generated from the pair-wise statistical tests. In fact, its deterioration for the entire problem set is just 0.006. Beating all the state-of-the-art algorithms in such a large number of prob-

Table 6: Statistical Tests and Deterioration Level of the HV ratio for problems with three objectives

	↑	↓	↔	**Score**	**Deterioration**
AVSD-MOEA/D	74	2	0	72	0.006
MOEA/D-DE	33	38	5	-5	1.075
NSGA-II	9	64	3	-55	1.745
NSGA-III	22	50	4	-28	1.149
R2-EMOA	45	29	2	16	0.851

lem benchmarks is a quite significant achievement, and shows the robustness of AVSD-MOEA/D. Our results show that the superiority of AVSD-MOEA/D persists, and even increases, when problems with three objective functions are considered. For a better comprehension of the strenghts and weakness of the algorithms, in the Figure 1 is shown the 50% attainment surfaces for WFG8 and UF5. An attainment surface approximation can be interpreted as the spatial region that is statistically attained among all the runs that were carried out by an algorithm [53, 54]. In other words, it can be understood as *the spatial region that is achieved by the k% among all the runs by one algorithm.* The most challenging characteristic of these problems are that WFG8 has strong dependencies among all the parameters, and UF5 is a multi-modal biased problem whose Pareto optimal front is discrete and consists of 21 points. In both problems AVSD-MOEA/D was the only one that converged adequately to the Pareto front at least 50% among all the runs. Even more, given that the standard deviation is too low it can be though that all the runs converged similarly well.

We can better understand the reasons behind the benefits of AVSD-MOEA/D against the state-of-the-art MOEAs by analyzing the evolution of the HV val-

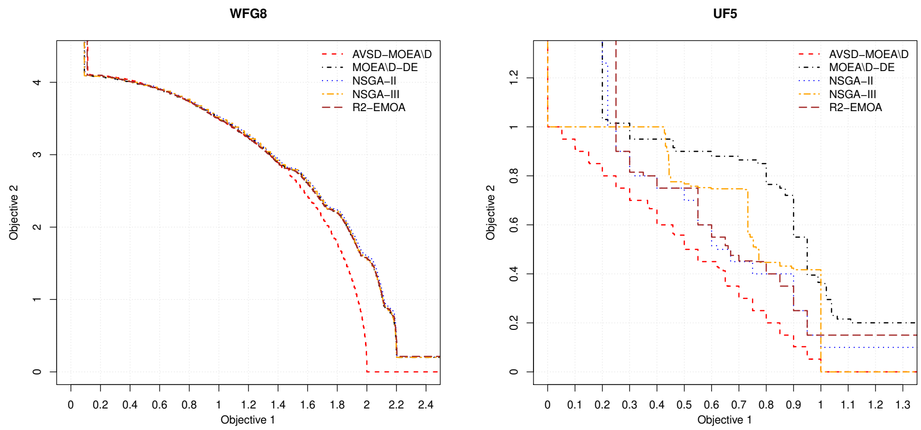

Figure 1: 50% attainment surfaces achieved for WFG8 and UF5 test problems.

ues and the diversity. Note that in some MOPs, variables can be classified into two types: distance variables and position variables. A variable x_i is a distance variable when for all x, modifying x_i results in a new solution that dominates x, is equivalent to x, or is dominated by x. Differently, if x_i is a position variable, modifying x_i in x always results in a vector that is incomparable or equivalent to x [7]. This is important because in some cases, MOEAs do not maintain a large enough diversity in the distance variables [13], so analyzing the diversity trend for these kinds of variables provides an useful insight into the dynamics of the population.

In order to show the behavior of the different schemes, we selected WFG5 and UF5. They are complementary in the sense that in WFG5, all the Pareto solutions exhibit constant values for the distant variables, which is not the case in UF5. Moreover, in UF5, the optimal regions are isolated in the variable space, meaning that more diversity is required. For each algorithm, the diversity is calculated as the average Euclidean distance between individuals

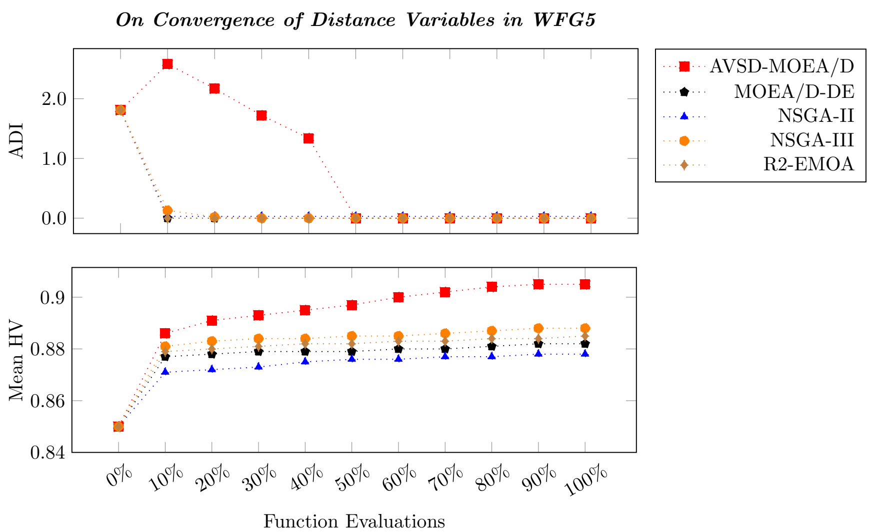

Figure 2: Diversity of distance variables (top) and mean of HV ratios (bottom) vs. elapsed time in the bi-objective WFG5 test problem. The results reported were taken from 35 runs.

(ADI) in the population by considering only the distance variables. Figures 2 and 3 show the evolution of the ADI (top) and the mean of HV (bottom) for WFG5 and UF5, respectively. In the WFG5 problem, the distance variables quickly converged to a small region in state-of-the-art MOEAs. Thus, the differential evolution operator loses it exploring power and as a result, those MOEAs were unable to significantly improve the quality of the approximations as the evolution progresses. By contrast, in the case of AVSD-MOEA/D, the decrease in ADI is quite linear until the midpoint of the execution, and the increase in HV is gradual. The final HV attained by AVSD-MOEA/D is the largest one, which shows the important benefit of gradually decreasing the diversity.

As expected, explicitly promoting diversity is also beneficial for problems

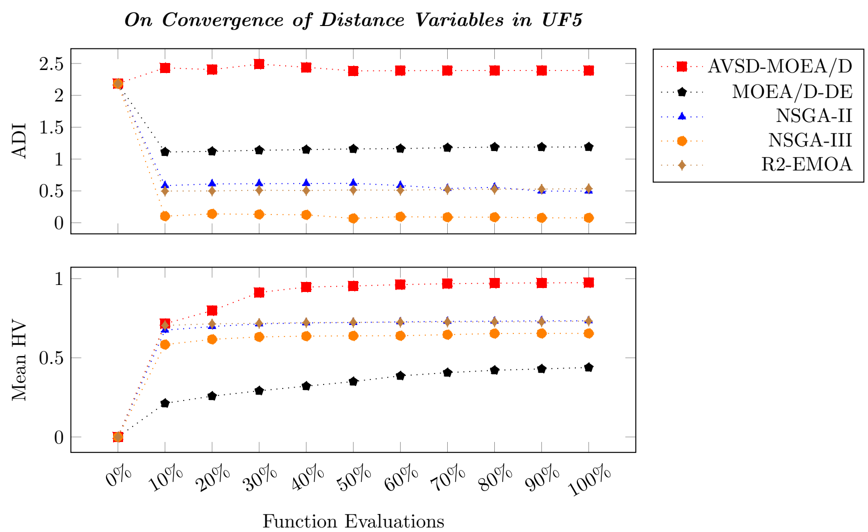

Figure 3: Diversity of distance variables (top) and mean of HV ratios (bottom) vs. elapsed time in the bi-objective UF5 test problem. The results reported were taken from 35 runs.

with disconnected optimal regions. As the data in Figure 3 show, the advantage of promoting diversity in the UF5 test problem is clear. In this case, state-of-the-art algorithms maintain some degree of diversity in the distance variables for the entire search. However, a large degree of diversity is required to obtain the 21 optimal solutions, and these MOEAs do not maintain the required amount of diversity, and as a result, they miss many of the solutions. In the case of AVSD-MOEA/D, enforcing a large degree of diversity in the initial phases promotes more exploration, which makes it possible to find additional optimal regions. Once these regions are located, they are not discarded, meaning that a larger level of diversity is maintained throughout the execution. This way, AVSD-MOEA/D not only attained better HV values for the first 10% of the total function evaluations, but it also kept looking for promising regions. In fact, its HV values improved significantly until the

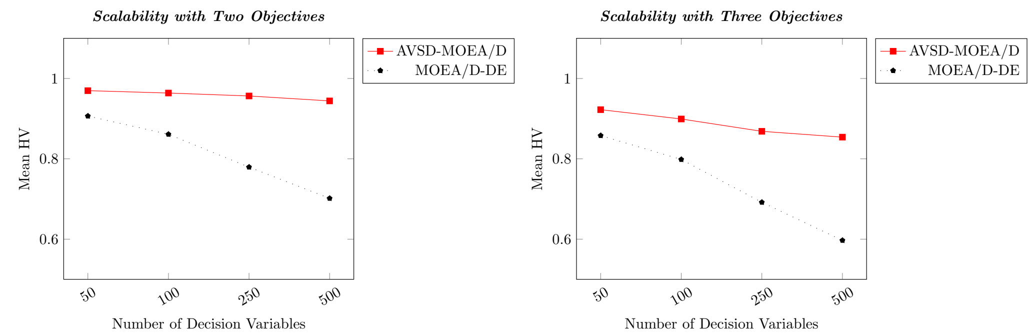

Figure 4: Mean of the HV ratio for 35 runs of the two-objective and three-objective problems for different numbers of variables

midpoint of the execution period i.e., the final moment when diversity was explicitly promoted. Then, an additional increase was obtained due to intensification in the regions identified. This analysis shows that the dynamic of the population depends on the problem at hand. The behavior of AVSD-MOEA/D with all the problems tested was similar to those already presented. Scenaries where the optimal regions consists of constant values for the distance variables behave like WFG5, whereas the behavior in those cases where the optimal regions consist of non-constant values for the distance variables is more similar to the UF5 case. Note, however, that in these cases, different levels of diversity are required, so the behavior is not as homogeneous.

4.2. Analysis of Scalability in the Decision Variables

In order to gain a better insight into the benefits of our proposal, we present an analysis of the scalability in terms of the number of decision variables. Given the computational cost associated with this experiment, it was only performed for the decomposition-based algorithms. AVSD-MOEA/D and MOEA/D-DE were applied to the same benchmark problems as in the previous experiment, but considering 50, 100, 250 and 500 variables. Note that in the WFG test problems, the number of position variables and distance variables must be specified. The number of distance variables was set to 42, 84, 210 and 418 when using 50, 100, 250 and 500 variables, respectively. The remaining decision variables were position variables, meaning there were 8, 16, 40 and 82 such variables, respectively. Thus, the relationship between the number of position and distance variables was kept fixed. In addition, the stopping criterion was set to 2.5×10^7 function evaluations. Figure 4 shows the mean HV ratio for the selected algorithms, considering the problems with two and three objectives, respectively. As expected, the HV ratio decreased as the number of variables increased. However, the performance of AVSD-MOEA/D is quite robust, and its decrease is less aggressive than the one in MOEA/D-DE, meaning that AVSD-MOEA/D is more helpful as the complexity increases. In fact, in our previous analyses, AVSD-MOEA/D also stood out in the most complex cases, such as WFG8 and UF5.

4.3. Analysis of the Initial Penalty Threshold

One of the main potential downsides of including a strategy to explicitly promote diversity is that this is usually at the cost of incorporating additional parameters. In the case of AVSD-MOEA/D, it requires setting the initial penalty threshold (D_I). Given that in single-objective cases, values close to 0.4 have yielded proper performance [31, 20], $D_I = 0.4$ was used in the above experiment. This section provides a more detailed study of the implications of this parameter.

In order to better understand the importance of D_I, the entire set of benchmark problems was tested with different values of D_I. As in previous

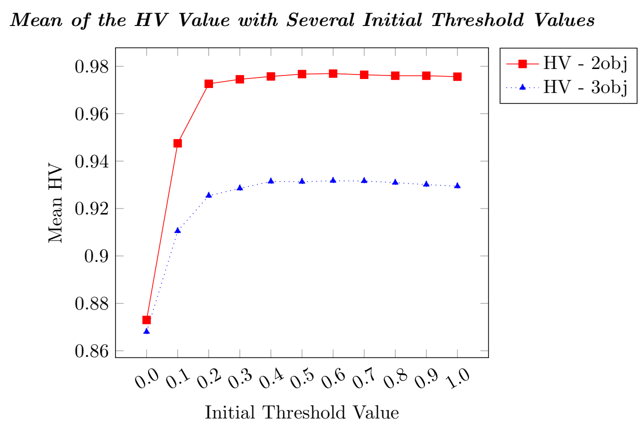

Figure 5: Mean of HV values for all the problems with several initial threshold values

experiments, the stopping criterion was set to 2.5×10^7 function evaluations. Since normalized distances are used, the maximum attainable distance between pairs of individuals is 1.0. Also note that setting D_I to 0 implies not promoting diversity in the variable space. Thus, several values in this range were considered. Specifically, the values $D_I = \{0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$ were tested. Figure 5 shows the mean HV ratio obtained for both the two-objective and the three-objective case with the D_I values tested. The AVSD-MOEA/D performed worst when D_I was set to 0. The HV ratio quickly increased as higher D_I values up to 0.2 were used. Larger values yielded quite similar performances. Thus, a wide range of values (from 0.2 to 1.0) exhibited very good performance, meaning that the behavior of AVSD-MOEA/D is quite robust. Thus, properly setting this parameter is not a complex task.

In order to better understand the implications of D_I on the dynamics of the population, Figure 6 shows, for AVSD-MOEA/D, the evolution of diversity in the distance variables in the WFG9 case for three different values of D_I. When setting $D_I = 0$, the diversity is reduced quite quickly, which results

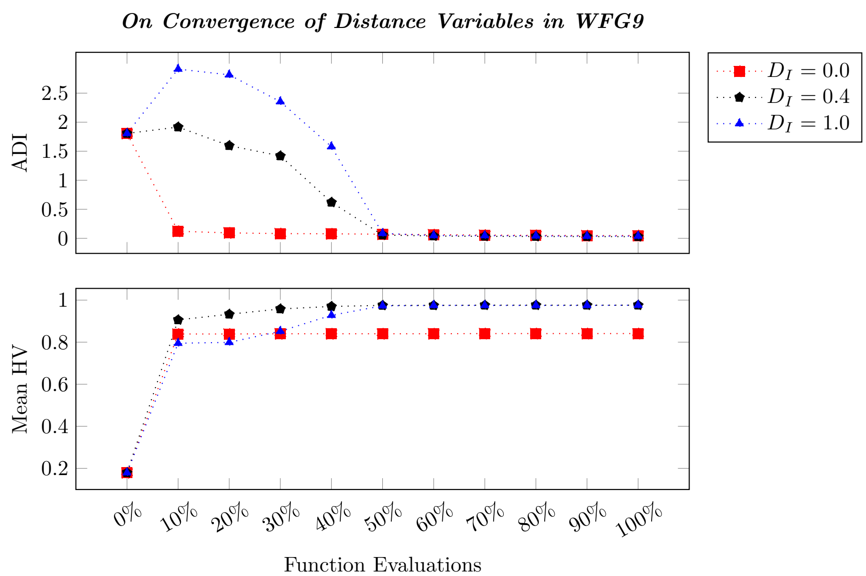

Figure 6: Diversity of distance variables (top) and mean of HV ratios (bottom) vs. elapsed time in the two-objective WFG9 test problem. The results reported were taken from 35 runs.

in premature convergence. The result is a hypervolume that is not too high. However, when $D_I = 0.4$ and $D_I = 1$ are used, the loss of diversity is slowed down, and the resulting hypervolume is quite large. Note that setting $D_I = 1$ promotes greater diversity, so the hypervolume increases slower than when $D_I = 0.4$. However, the degree of exploration in both cases is enough to yield high-quality solutions. The behavior is quite similar in every problem, which explains the stability of the algorithms for different values of D_I. Note that for shorter periods, setting a proper D_I value is probably much more important. However, for long-term executions at least, practically any value higher than 0.2 yields similar solutions, which we regard as a highly positive feature.

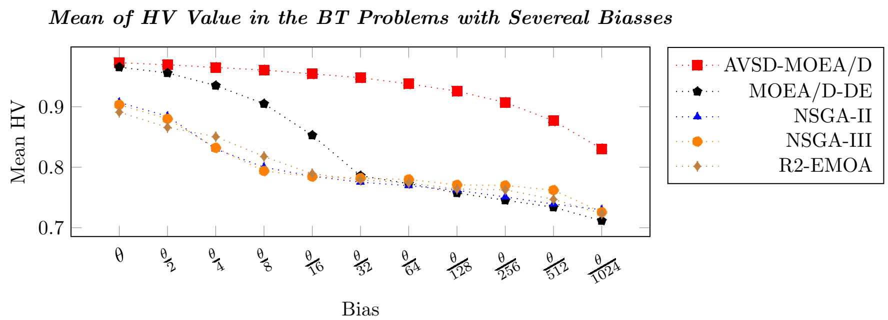

Figure 7: Mean of HV values for eight BTs problems (y-axis) against several biasses ratios (x-axis). The BT2 problem is not taken into consideration due that it suffers of numerical stability.

4.4. On the Convergence of MOEAs *in Test Problems with Bias Features*

As pointed out in [47, 55, 7], the bias feature is one of the most challenging difficulties that MOEAs might face. Recently, the BTs test problems were proposed to facilitate the study of the ability of MOEAs for dealing with biases. In this context bias means that small variations in the decision space around the Pareto set cause significant changes in vicinities of some Pareto front solutions [7]. Particularly, those problems are built with transformations that induce position-related bias and distance-related bias. While the former means that a small change on the position-related variables of one solution in the Pareto set projects a significant change along the Pareto front. The later imposes that a small variation on the distance-related variables of one solution in the Pareto set causes a significant deterioration on the convergence towards the Pareto front.

In order, to analyze the capability of the MOEAs to deal with bias features the BTs problems are taken into account. Specifically, this section analyses the sensitivity of the algorithms imposing several levels of bias in the distance-

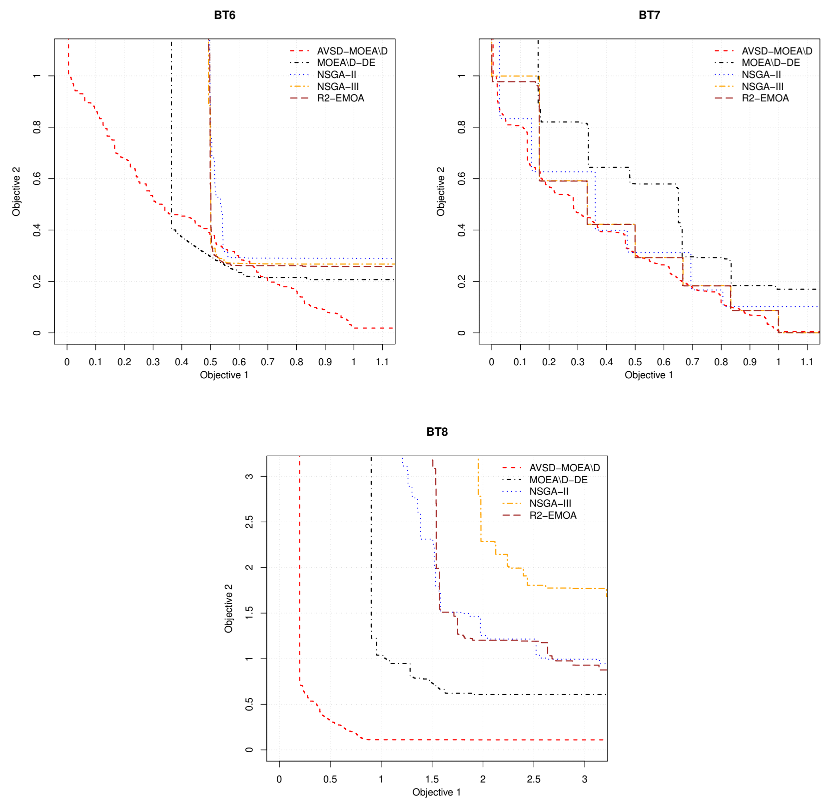

Figure 8: 50% attainment surfaces achieved for BT6 and BT8 test problems with a bias of $\frac{\theta}{32}$

related variables. Initially, for each problem the position-related bias and distance-related bias (θ) are kept exactly as the one proposed in the original work [47]. Then, for each problem its initial distance-related bias value (θ) is iterativelly decreased by a factor of two. Specifically, the distance-related bias taken into account are $\{\theta, \frac{\theta}{2}, \frac{\theta}{4}, \frac{\theta}{8}, \frac{\theta}{16}, \frac{\theta}{32}, \frac{\theta}{64}, \frac{\theta}{128}, \frac{\theta}{256}, \frac{\theta}{512}, \frac{\theta}{1028}\}$. Figure 7 shows the mean HV ratio obtained with several distance-related biasses. Also note that the BT2 problem is not taken into consideration due that increasing its bias values provokes numerical instability since that it incorporates a different bias transformation, nevertheless all the results can be consulted in the supplementary document. Taking exactly the original configuration (bias of θ) [47] AVSD-MOEA/D is sigthly better than MOEA/D-DE, but as soon as the bias is decreased to $\frac{\theta}{32}$ the performance of MOEA/D-DE decays aggressively. Furthermore, the performance of AVSD-MOEA/D is superior than 0.9 with biasses values upper or equal to $\frac{\theta}{256}$ which is quite superior than the state-of-the-art MOEAs whose values at that point are approximately of 0.75. Figure 8 shows the 50% of attainment surface of BT6, BT7 and BT8 with a bias of $\frac{\theta}{32}$. BT6 and BT8 have simple nolinear Pareto set while BT7 has a complicated nolinear Pareto set. BT8 is multimodal. Although that MOEA/D-DE converged to a region of the Pareto front with BT6 AVSD-MOEA/D covered a huge region of the Pareto front, in fact this shows that for this problem promoting diversity in the decision space results in diversity in the objective space. In addition, AVSD-MOEA/D converges quite well in complicates nonlinear Pareto sets shown in the 50% attained surface of BT7 (Figure 8). Finally but not less important AVSD-MOEA/D shows a superior behaviour with biased and multimodal problems as is the case of BT8 whose attainment surfaces have converged much better to the Pareto front.

5. Conclusion

Premature convergence is one of the most typical drawbacks of EAs. MOEAs indirectly promote the preservation of diversity in the variable space because of the implicit relationship between the diversity maintained in the objective space and the one maintained in the variable space. However, for many problems the degree of diversity maintained is not sufficient to ensure the exploratory power of genetic operators and locate the optimal regions. In single-objective optimization, many of the state-of-the-art algorithms explicitly manage diversity. Specifically, those schemes that relate the degree of diversity to the elapsed period of execution and to the stopping criterion have excelled. This paper shows that this design principle is also helpful in the area of multi-objective optimization, where the optimization of many of the most complex popular benchmark problems can be improved further by applying this design principle.

In order to prove this hypothesis, a novel replacement operator based on the aforementioned design principle is applied to generate a decomposition-based MOEA that takes into account the diversity in both the variable and objective spaces. This is done using a dynamic penalty method. Note that since the aim of the approach is to improve the results when considering metrics in the objective space, the importance given to the diversity in the variable space is reduced as the evolution progresses, meaning that in the later phases, our proposal behaves more similarly to traditional MOEAs. Additionally, taking into account recent advances, and to ensure that our proposal maintains high-quality solutions despite the penalty scheme, an external archive based on the R2-indicator is incorporated. Because of this, we refer to our proposal as *Archived Variable Space Diversity MOEA based on Decomposition* (AVSD-MOEA/D).

The experimental validation carried out shows the remarkable improvement provided by AVSD-MOEA/D in comparison to state-of-the-art MOEAs with both two-objective and three-objective problems. The scalability analyses show that as the number of decision variables increases, the benefits of including proper diversity management are even more important, so the differences in performance increase. In fact, the most remarkable benefits emerge for the most complex cases. Moreover, the analysis of the initial penalty threshold, which is an additional parameter required by AVSD-MOEA/D, shows that the method is quite robust, which makes finding a proper parameter value an easy task. Finally, in order to better understand the reasons behind the huge superiority of our proposal, some analyses involving the dynamics of the populations are provided. In comparison to state-of-the-art algorithms, our proposal clearly slows down convergence.

In the future, we plan to apply the principles studied in this paper to other categories of MOEAs. For instance, including the diversity management method presented in this paper in indicator-based MOEAs seems plausible. Additionally, in order to obtain even better results, these strategies are going to be incorporated with continuation and/or individual improvement methods.

Acknowledgments

Authors acknowledge the financial support from CONACyT through the "Ciencia Básica" project no. 285599 and the support from "Laboratorio de Supercómputo del Bajio" through the project 300832 from CONACyT.

References

[1] S. Das, S. Maity, B.-Y. Qu, P. N. Suganthan, Real-parameter evolutionary multimodal optimization—a survey of the state-of-the-art, Swarm and Evolutionary Computation 1 (2011) 71–88.

[2] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, Q. Zhang, Multiobjective evolutionary algorithms: A survey of the state of the art, Swarm and Evolutionary Computation 1 (2011) 32–49.

[3] F. A. Lootsma (Ed.), Multi-Objective Linear Programming, Springer US, Boston, MA, 1999, pp. 229–257. URL: https://doi.org/10.1007/978-0-585-28008-0_10. doi:10.1007/978-0-585-28008-0_10.

[4] A. Trivedi, D. Srinivasan, K. Sanyal, A. Ghosh, A survey of multiobjective evolutionary algorithms based on decomposition, IEEE Transactions on Evolutionary Computation 21 (2016) 440–462.

[5] D. A. Van Veldhuizen, G. B. Lamont, Multiobjective evolutionary algorithm research: A history and analysis, Technical Report, Citeseer, 1998.

[6] C. A. C. Coello, G. B. Lamont, D. A. V. Veldhuizen, Mop evolutionary algorithm approaches, Evolutionary Algorithms for Solving Multi-Objective Problems: Second Edition (2007) 61–130.

[7] S. Huband, P. Hingston, L. Barone, L. While, A review of multiobjective test problems and a scalable test problem toolkit, IEEE Transactions on Evolutionary Computation 10 (2006) 477–506.

[8] N. Beume, B. Naujoks, M. Emmerich, SMS-EMOA: Multiobjective selection based on dominated hypervolume, European Journal of Operational Research 181 (2007) 1653–1669.

[9] Q. Zhang, H. Li, Moea/d: A multiobjective evolutionary algorithm

based on decomposition, IEEE Transactions on evolutionary computation 11 (2007) 712–731.

[10] M. Črepinšek, S.-H. Liu, M. Mernik, Exploration and Exploitation in Evolutionary Algorithms: A Survey, ACM Computing Surveys 45 (2013) 35:1–35:33.

[11] A. Herrera-Poyatos, F. Herrera, Genetic and memetic algorithm with diversity equilibrium based on greedy diversification, CoRR abs/1702.03594 (2017). URL: `http://arxiv.org/abs/1702.03594`. `arXiv:1702.03594`.

[12] H. Ishibuchi, Y. Nojima, T. Doi, Comparison between single-objective and multi-objective genetic algorithms: Performance comparison and performance measures, in: 2006 IEEE International Conference on Evolutionary Computation, IEEE, 2006, pp. 1143–1150.

[13] J. C. Castillo, C. Segura, A. H. Aguirre, G. Miranda, C. León, A multi-objective decomposition-based evolutionary algorithm with enhanced variable space diversity control, in: Proceedings of the Genetic and Evolutionary Computation Conference Companion, ACM, 2017, pp. 1565–1571.

[14] D. Büche, S. Müller, P. Koumoutsakos, Self-adaptation for multi-objective evolutionary algorithms, in: International Conference on Evolutionary Multi-criterion Optimization, Springer, 2003, pp. 267–281.

[15] H. Lu, G. G. Yen, Dynamic population size in multiobjective evolutionary algorithms, in: Proceedings of the 2002 Congress on Evolutionary

Computation. CEC'02 (Cat. No. 02TH8600), volume 2, IEEE, 2002, pp. 1648–1653.

[16] S. W. Mahfoud, Niching methods for genetic algorithms, Ph.D. thesis, Citeseer, 1995.

[17] N. Srinivas, K. Deb, Muiltiobjective optimization using nondominated sorting in genetic algorithms, Evolutionary computation 2 (1994) 221–248.

[18] L. Lin, M. Gen, Auto-tuning strategy for evolutionary algorithms: balancing between exploration and exploitation, Soft Computing 13 (2009) 157–168.

[19] C. Segura, A. Hernández-Aguirre, F. Luna, E. Alba, Improving diversity in evolutionary algorithms: New best solutions for frequency assignment, IEEE Transactions on Evolutionary Computation 21 (2016) 539–553.

[20] J. C. Castillo, C. Segura, Differential evolution with enhanced diversity maintenance, Optimization Letters (2019) 1–20.

[21] Q. Zhang, W. Liu, H. Li, The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances, in: Evolutionary Computation, 2009. CEC'09. IEEE Congress on, IEEE, 2009, pp. 203–208.

[22] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE transactions on evolutionary computation 6 (2002) 182–197.

[23] H. Trautmann, T. Wagner, D. Brockhoff, R2-EMOA: Focused multi-objective search using r2-indicator-based selection, in: International Conference on Learning and Intelligent Optimization, Springer, 2013, pp. 70–74.

[24] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints, IEEE transactions on evolutionary computation 18 (2013) 577–601.

[25] J. J. Liang, C. T. Yue, B. Y. Qu, Multimodal multi-objective optimization: A preliminary study, in: 2016 IEEE Congress on Evolutionary Computation (CEC), 2016, pp. 2454–2461.

[26] K. Deb, S. Tiwari, Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization, European Journal of Operational Research 185 (2008) 1062–1087.

[27] O. Cuate, O. Schütze, Variation rate to maintain diversity in decision space within multi-objective evolutionary algorithms, Mathematical and Computational Applications 24 (2019).

[28] H. M. Pandey, A. Chaudhary, D. Mehrotra, A comparative review of approaches to prevent premature convergence in ga, Applied Soft Computing 24 (2014) 1047–1077.

[29] O. J. Mengshoel, S. F. Galán, A. De Dios, Adaptive generalized crowding for genetic algorithms, Information Sciences 258 (2014) 140–159.

[30] G. R. Harik, Finding multimodal solutions using restricted tournament selection., in: ICGA, 1995, pp. 24–31.

[31] E. Romero Ruiz, C. Segura, Memetic algorithm with hungarian matching based crossover and diversity preservation, Computación y Sistemas 22 (2018).

[32] O. M. Shir, M. Preuss, B. Naujoks, M. Emmerich, Enhancing decision space diversity in evolutionary multiobjective algorithms, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2009, pp. 95–109.

[33] A. Toffolo, E. Benini, Genetic diversity as an objective in multi-objective evolutionary algorithms, Evolutionary computation 11 (2003) 151–167.

[34] K. Chan, T. Ray, An evolutionary algorithm to maintain diversity in the parametric and the objective space, in: International Conference on Computational Robotics and Autonomous Systems (CIRAS), Centre for Intelligent Control, National University of Singapore, 2005.

[35] A. Zhou, Q. Zhang, Y. Jin, Approximating the set of pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm, IEEE transactions on evolutionary computation 13 (2009) 1167–1189.

[36] T. Ulrich, J. Bader, E. Zitzler, Integrating decision space diversity into hypervolume-based multiobjective search, in: Proceedings of the 12th annual conference on Genetic and evolutionary computation, ACM, 2010, pp. 455–462.

[37] S. Bechikh, R. Datta, A. Gupta, Recent advances in evolutionary multi-objective optimization, volume 20, Springer, 2016.

[38] E. Zitzler, S. Künzli, Indicator-based selection in multiobjective search, in: International Conference on Parallel Problem Solving from Nature, Springer, 2004, pp. 832–842.

[39] H. Ishibuchi, T. Murata, A multi-objective genetic local search algorithm and its application to flowshop scheduling, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 28 (1998) 392–403.

[40] R. Hernández Gómez, C. A. Coello Coello, Improved metaheuristic based on the r2 indicator for many-objective optimization, in: Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, 2015, pp. 679–686.

[41] M. Pescador-Rojas, R. Hernández Gómez, E. Montero, N. Rojas-Morales, M.-C. Riff, C. A. Coello Coello, An overview of weighted and unconstrained scalarizing functions, in: H. Trautmann, G. Rudolph, K. Klamroth, O. Schütze, M. Wiecek, Y. Jin, C. Grimme (Eds.), Evolutionary Multi-Criterion Optimization, Springer International Publishing, Cham, 2017, pp. 499–513.

[42] H. Li, Q. Zhang, Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II, IEEE Transactions on evolutionary computation 13 (2009) 284–302.

[43] C. Segura, C. A. C. Coello, E. Segredo, A. H. Aguirre, A novel diversity-based replacement strategy for evolutionary algorithms, IEEE Transactions on Cybernetics 46 (2015) 3233–3246.

[44] P. Hajela, E. Lee, C.-Y. Lin, Genetic Algorithms in Structural Topology Optimization, Springer Netherlands, Dordrecht, 1993, pp. 117–133. URL: https://doi.org/10.1007/978-94-011-1804-0_10. doi:10.1007/978-94-011-1804-0_10.

[45] K. Deb, L. Thiele, M. Laumanns, E. Zitzler, Scalable test problems for evolutionary multiobjective optimization, in: Evolutionary multiobjective optimization, Springer, 2005, pp. 105–145.

[46] Q. Zhang, A. Zhou, S. Zhao, P. N. Suganthan, W. Liu, S. Tiwari, Multiobjective optimization test instances for the CEC 2009 special session and competition (2008).

[47] H. Li, Q. Zhang, J. Deng, Biased multiobjective optimization and decomposition algorithm, IEEE transactions on cybernetics 47 (2016) 52–66.

[48] H. Ishibuchi, R. Imada, Y. Setoguchi, Y. Nojima, Reference point specification in hypervolume calculation for fair comparison and efficient search, in: Proceedings of the Genetic and Evolutionary Computation Conference, 2017, pp. 585–592.

[49] K. Li, K. Deb, Q. Zhang, S. Kwong, An evolutionary many-objective optimization algorithm based on dominance and decomposition, IEEE Transactions on Evolutionary Computation 19 (2014) 694–716.

[50] J. J. Durillo, A. J. Nebro, C. A. C. Coello, J. García-Nieto, F. Luna, E. Alba, A study of multiobjective metaheuristics when solving parameter scalable problems, IEEE Transactions on Evolutionary Computation 14 (2010) 618–635.

[51] T. Wagner, H. Trautmann, D. Brockhoff, Preference articulation by means of the r2 indicator, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2013, pp. 81–95.

[52] H. Ishibuchi, H. Masuda, Y. Tanigaki, Y. Nojima, Modified distance calculation in generational distance and inverted generational distance, in: International conference on evolutionary multi-criterion optimization, Springer, 2015, pp. 110–125.

[53] J. Knowles, A summary-attainment-surface plotting method for visualizing the performance of stochastic multiobjective optimizers, in: 5th International Conference on Intelligent Systems Design and Applications (ISDA'05), IEEE, 2005, pp. 552–557.

[54] C. M. Fonseca, P. J. Fleming, On the performance assessment and comparison of stochastic multiobjective optimizers, in: International Conference on Parallel Problem Solving from Nature, Springer, 1996, pp. 584–593.

[55] K. Deb, Multi-objective genetic algorithms: Problem difficulties and construction of test problems, Evolutionary computation 7 (1999) 205–230.