Collaborative and Adaptive Strategies of Different Scalarizing Functions in MOEA/D

Miriam Pescador-Rojas
ESCOM, Instituto Politécnico Nacional,
Computer Science Department
CINVESTAV-IPN Mexico City
07360
Email: pescador@computacion.cs.cinvestav.mx

Carlos A. Coello Coello
Computer Science Department
CINVESTAV-IPN
Mexico City
07360
Email: ccoello@cs.cinvestav.mx

***Abstract*—In recent years, the use of decomposition-based multi-objective evolutionary algorithms has been very successful in solving both multi- and many-objective optimization problems. In these algorithms, the adopted Scalarizing Functions (SFs) play a crucial role in their performance. Methods such as the Modified Weighted Chebyshev (MCHE), Penalty Boundary Intersection (PBI) and Augmented Achievement Scalarizing Function (AASF) have been found to be very effective for achieving both convergence to the true Pareto front and a uniform distribution of solutions along it. However, the choice of an appropriate model parameter is required for these SFs. Some studies have analyzed the impact of these parameter values on the performance of the best-known decomposition multi-objective evolutionary algorithm (MOEA/D). In this paper, we propose a strategy based on collaborative populations combining different SFs and model parameter values via an adaptive operator selection based on the multi-armed bandit technique. Our preliminary results give rise to some interesting observations regarding the way in which different SFs are combined and adapted during the evolutionary process of MOEA/D.**

I. Introduction

A Multi-objective Optimization Problem (MOP) has two or more (often conflicting) objective functions which must be optimized at the same time. Mathematically, it can be described as follows:

$$\begin{aligned} &\text{minimize } \mathbf{f}(\mathbf{x}) = \left[f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})\right]^T \\ &\text{subject to } \mathbf{x} \in \Omega \end{aligned} \tag{1}$$

where Ω is the feasible space of solutions and $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \Omega$, is the vector of decision variables, $f_i : \mathcal{R}^n \to \mathcal{R}$, $i \in \{1, \ldots, m\}$ are the objective functions.

Let us assume that we have two vectors $\mathbf{u}, \mathbf{v} \in \mathcal{R}^m$. Then, we say that $\mathbf{u}$ *dominates* $\mathbf{v}$ (denoted by $\mathbf{u} \prec \mathbf{v}$) if $u_i \leq v_i$ for every $i \in \{1, ..., m\}$, and $u_j \neq v_j$ for at least one index $j \in \{1, ..., m\}$. We say that a vector of decision variables $\mathbf{x}^* \in \Omega$ is *Pareto optimum* if there does not exist another $\mathbf{x} \in \Omega$ such that $\mathbf{f}(\mathbf{x}) \prec \mathbf{f}(\mathbf{x}^*)$. The *Pareto Optimal Set* is defined by: $\mathcal{PS}^* = \{\mathbf{x} \in \Omega | \nexists \mathbf{y} \in \Omega : \mathbf{y} \prec \mathbf{x}\}$. The vectors $\mathbf{x}^*$ corresponding to the solutions included in $\mathcal{PS}^*$ are called *nondominated* set. The *Pareto Optimal Front* is defined by: $\mathcal{PF}^* = \{\mathbf{f}(\mathbf{x}) \in \mathcal{R}^m | \mathbf{x} \in \mathcal{PS}^*\}$.

Multi-Objective Evolutionary Algorithms (MOEAs) have been able to solve complex MOPs properly. The first generation of MOEAs were algorithms that used Pareto dominance in their selection mechanism such as NSGA-II [1] and SPEA [2]. However, these MOEAs do not work properly in many-objective optimization (i.e., with problems having more than three objectives). The two most common approaches to deal with many-objective problems are: (1) indicator-based MOEAs and (2) decomposition-based MOEAs. The former refers to the use of methods to assess the quality of the approximation generated by the MOEA using a performance indicator. The best-known indicator-based MOEA is SMS-EMOA [2] which is based on the hypervolume indicator and is, therefore, computationally expensive in many-objective problems. Decomposition-based MOEAs have several advantages such as scalability to many-objective problems (MaOPs) [3], [4], a high search ability for combinatorial optimization [5], [6], [7], and a high compatibility with local search methods [8], [9], [10]. In this category, MOEA/D [11] is a computationally efficient algorithm that decomposes a MOP into a set of single-objective subproblems and applies its evolutionary operators on neighboring solutions. Several proposals [12], [13], [14], [15], [16] have been designed to enhance the original MOEA/D. In [12], the simulated binary crossover was replaced by the differential evolution operator. MOEAD/M2M [13] introduced a method to improve diversity through the generation of subpopulations. MOEA/DD [14] combined Pareto dominance with decomposition-based algorithms to maintain diversity.

One of the most important components of MOEA/D is the choice of the Scalarizing Function (SF) since such function is known to play an important role in its performance. Moreover, some SFs require the definition of a parameter value to which they may be very sensitive. For example, the Penalty Boundary Intersection (PBI) function defines a parameter to balance convergence and uniformity along the true Pareto front. Some studies [5], [17], [6] have provided a sensitivity analysis of PBI, indicating that the choice of an appropriate parameter value depends on specific features such as the Pareto front geometry [18], [19], or the number of objective functions [11]. In Section II, we will detail additional observations of the PBI method and other SFs.

The two most common possibilities for dealing with the selection of parameters settings are (1) parameter tuning and (2) parameter control techniques. The first one refers to establishing *a priori* the parameters values and use them in all the iterations of an MOEA. In contrast, the second case includes adaptive mechanisms that modify the parameters values according to the information gathered during the search process. Here, we explore the second option to select the best model parameters of the SFs under study.

Ishibuchi et al. [20], [21] presented some mechanisms to use several SFs simultaneously in MOEA/D as an option to maintain diversity and to solve many-objective problems.

Diverse MOEA/D improvements have employed a bandit-based Adaptive Operator Selection (AOS) mechanism which is an upper confidence model that selects the most suitable operator from a pool of options according to some determinist rule based on multi-armed bandit algorithms. Bandit-based AOS establish two tasks: 1) to determine a reward for each operator via its record of fitness improvement rates, and 2) to select an operator for being used according to the current reward value.

The aim of this paper is to adopt a bandit-armed strategy, only used in evolutionary operators, to select an appropriate SF during the evolutionary search process of MOEA/D. Our proposed approach uses collaborative subpopulations to establish the best model parameter for each MOP. We also investigate the scalability of our proposed approach up to 10 objectives, using several benchmark problems.

The remainder of this paper is organized as follows. In Section II, we provide the mathematical definition of the SFs adopted. Section III presents the previous related work. In Section IV, we describe our proposal and we provide a study of the effect of combining different SFs via collaborative subpopulations and an adaptive selection strategy. Finally, Section V presents our conclusions and some possible paths for future work.

II. Weighted and unconstrained Scalarizing Functions

Weighted and unconstrained Scalarizing Functions (SFs) are methods that transform a Multi-objective Optimization Problem (MOP) into several single-objective subproblems. Let $\Lambda = \{\boldsymbol{\lambda}^1, \ldots, \boldsymbol{\lambda}^N\}$ be a set of N uniformly distributed weight vectors, where each $\boldsymbol{\lambda}^j = (\lambda_1^j, \ldots, \lambda_m^j)^T$ must satisfy $\sum_{i=1}^m \lambda_i^j = 1$ and $\lambda_i^j \geq 0$ for all $i \in \{1, \ldots, m\}$. Each component value in the weight vectors represents the relative importance assigned to each objective function. Then, the decomposition method associates each objective function with a weight coefficient to minimizes a SF (as a single-objective problem). This can be stated as:

$$\begin{aligned} &\text{minimize } g(\mathbf{x}|\boldsymbol{\lambda}^j) \\ &\text{subject to } \mathbf{x} \in \Omega \end{aligned} \tag{2}$$

Function $g(\mathbf{x}|\boldsymbol{\lambda}^j)$ minimizes the distance between a candidate solution and the reference point via a target direction (weight vector) in objective space. Next, we briefly describe some SFs and their additional model parameter.

The **Weighted Norm (WN)** [22] or weighted L_p- metric function is stated as follows:

$$\min_{\mathbf{x}\in\Omega} g^{wn}(\mathbf{x}|\boldsymbol{\lambda}^j) = \big(\sum_{i=1}^{m} \lambda_i^j (f_i'(\mathbf{x})^p)\big)^{1/p}, \tag{3}$$

where $p \in [1, \infty)$, $f_i'(\mathbf{x}) = f_i(\mathbf{x}) - z_i^*$, and $z_i^* = \arg\min\{f_i(x)|x \in \Omega\}$ also known as the ideal vector which is not available *a priori* in a MOP. Thus, z_i^* is updated during the execution of a decomposition-based MOEA, via the best values for each objective function from a set of non-dominated solutions. If $p = 1$ the formulation is known as the **Weighted Sum (WS) function** [23]. It is well known that WS is not able to solve problems with nonconvex Pareto front shapes.

Some studies have shown that the WN function is very sensitive to the model parameter p. In [24], different ways to adapt dynamically the parameter p were presented. The $p-$value was set in the range of 1 to 4, and it was changed during the search process. In [25], a selection strategy was introduced to determine the $p-$value based on the curvature estimation of local regions of the Pareto front.

The **Chebyshev or Tchebycheff (CHE) function** [26] is a derivative formulation of WN when $p = \infty$, and is given by:

$$\min_{\mathbf{x}\in\Omega} g^{che}(\mathbf{x}|\boldsymbol{\lambda}^j) = \max_{1\leq i\leq m} \Big\{\lambda_i^j \big|f_i'(\mathbf{x})\big|\Big\}. \tag{4}$$

The CHE function does not require additional parameters and was widely used in different improved versions of MOEA/D [11] such as [12], [27], [28], [29], [30], [31].

The **Augmented Chebyshev (ACHE) function** [32] is a modified CHE version defined by equation (5), where an extra parameter (α) is considered to avoid the generation of *weak Pareto optimal solutions*[1].

$$\min_{\mathbf{x}\in\Omega} g^{ache}(\mathbf{x}|\boldsymbol{\lambda}^j) = \max_{1\leq i\leq m} \Big\{\lambda_i^j \big|f_i'(\mathbf{x})\big|\Big\} + \alpha \sum_{i=1}^{m} f_i'(\mathbf{x}). \tag{5}$$

In [32], the α parameter adopted small values defined in the range from 0.001 to 0.01. However, in [21], a large value of α provided a better performance within the MOEA/D framework.

Another variant of CHE is the so-called **Modified Chebyshev (MCHE) function** [33], and is formulated as:

$$\min_{\mathbf{x}\in\Omega} g^{mche}(\mathbf{x}|\boldsymbol{\lambda}^j) = \max_{1\leq i\leq m} \left\{\lambda_i^j \Big(\big|f_i'(\mathbf{x})\big| + \alpha \sum_{i=1}^{m} f_i'(\mathbf{x})\Big)\right\}, \tag{6}$$

where the α parameter should be set to small positive values. The differences between ACHE and MCHE were discussed in [34, p. 101].

[1]Let us assume $\mathbf{x}, \mathbf{y} \in \mathcal{S}$. We say that $\mathbf{x}$ is *weakly Pareto optimal* if there is no $\mathbf{y}$ such that $\forall i \in \{1, \ldots, m\}$, $f_i(\mathbf{y}) < f_i(\mathbf{x})$.

The **Achievement Scalarizing Function (ASF)** [35] is a modified version of CHE that uses the inverse of the weight vectors. It is stated as:

$$\min_{\mathbf{x}\in\Omega} g^{asf}(\mathbf{x}|\boldsymbol{\lambda}^j) = \max_{1\leq i\leq m}\left\{\frac{|f_i'(\mathbf{x})|}{\lambda_i^j}\right\}. \quad (7)$$

Contrary to the CHE function, the ASF function generates optimal solutions in the same target directions (see [3]).

Additionally, we study two versions of ASF defined as: Augmented ASF (AASF) [34] and Modified ASF (MASF) given by equations (8) and (9).

$$\min_{\mathbf{x}\in\Omega} g^{aasf}(\mathbf{x}|\boldsymbol{\lambda}^j) = \max_{1\leq i\leq m}\left\{\frac{|f_i'(\mathbf{x})|}{\lambda_i^j}\right\} + \alpha\sum_{i=1}^{m} f_i'(\mathbf{x}). \quad (8)$$

$$\min_{\mathbf{x}\in\Omega} g^{masf}(\mathbf{x}|\boldsymbol{\lambda}^j) = \max_{1\leq i\leq m}\left\{\frac{|f_i'(\mathbf{x})| + \alpha\sum_{i=1}^{m} f_i'(\mathbf{x})}{\lambda_i^j}\right\}. \quad (9)$$

The **Penalty Boundary Intersection (PBI)** method was introduced in the original MOEA/D [11]. Its formulation is defined by the following equation:

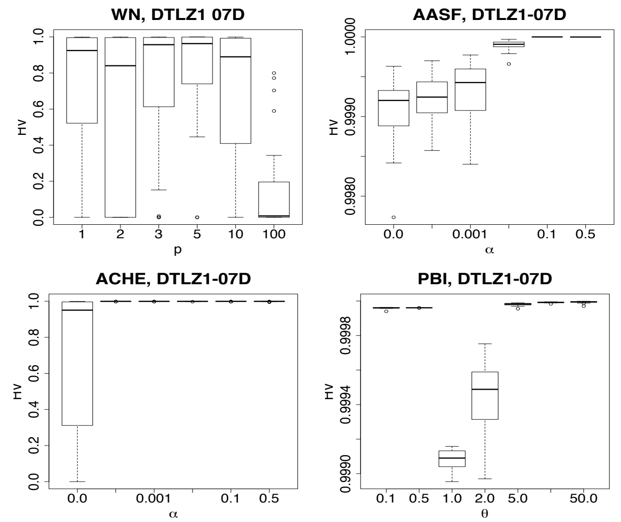
$$\min_{\mathbf{x}\in\Omega} g^{pbi}(\mathbf{x}|\lambda^j) = d_1 + \theta d_2 \quad (10)$$

where $d_1 := \left|\mathbf{f}'(\mathbf{x}) \cdot \frac{\lambda_j}{\|\lambda_j\|}\right|$ and $d_2 := \left\|\mathbf{f}'(\mathbf{x}) - d_1\frac{\lambda_j}{\|\lambda_j\|}\right\|$.

where d_1 represents the distance between the reference point and an optimal solution to measure convergence. Similarly, d_2 defines the perpendicular distance between the reference vector and an optimal point to measure uniformity. θ is a penalty parameter that handles the balance between d_1 and d_2. In different works, PBI has been considered one of the most suitable scalarizing functions for its ability to generate uniformly distributed Pareto optimal points. However, some studies [5], [17], [6] have shown a high sensitivity of this θ-penalty parameter.

Figure 1 illustrates an example of model parameter sensitivity for the MOEA/D framework using the AASF, ACHE, PBI and WN functions. Here, DTLZ1 with seven objectives is solved. In this case, we performed 30 independent runs and compute the normalized hypervolume for different model parameter values (WN with $p = \{1, 2, 3, 5, 10, 100\}$, PBI with $\theta = \{0.1, 0.5, 1, 2, 5, 10, 50\}$, AASF and ACHE with $\alpha = \{0.0, 0.0001, 0.001, 0.1, 0.5\}$). We can observe that ACHE reaches the maximum hypervolume value in most of the α values tested, but WN had the worst performance and high variance. PBI reached good hypervolume values with $\theta = \{0.1, 0.5, 5.0, 10.0, 50.0\}$ and AASF reached good hypervolume values with $\alpha = \{0.1, 0.5\}$. [36] showed evidence about the influence of the Pareto front shape and the number of objective functions on the selection of the most suitable model parameter value in SF.

Fig. 1: The hypervolume reached by the use of different SFs in MOEA/D.



III. Previous Related Work

This section presents a review of the related works that use either multiple Scalarizing Functions (SFs) in their search process or some bandit-based Adaptive Operator Selection (AOS) mechanism coupled to MOEA/D.

Ishibuchi et al. [20] presented an early work that combined the WS and CHE functions in MOEA/D. At each generation, a selection strategy monitors if k or more solutions in the neighborhood of an individual have the same objective vector. If that's the case, then CHE is adopted; otherwise, WS is adopted. In [21], two alternatives were proposed to combine the WS and CHE functions. Given a set of uniformly distributed weight vectors, two subpopulations focused on a particular SF. The second alternative is to assign, alternately, one SF to each vector. Recently, Hernández and Coello [37] proposed a hyper-heuristic where each individual in the population minimizes a different SF assigned by a heuristic selection mechanism based on the quality indicator called s-energy with the purpose of maintaining a uniform distribution of solutions. Here, seven different SFs were employed.

Multi-armed bandit algorithms have been a promising technique for improving the performance of MOEA/D. Li et al. [38] proposed an adaptive mechanism to select a mutation operator from Differential Evolution (DE) versions based on the multi-armed bandit. In this type of adaptive strategy, there are two processes: one is to assign a reward to a strategy based on its recent performance in the search process. The second is the choice of the best strategy based on these current reward values. For instance, in [39], an adaptive mechanism was proposed to select two components of MOEA/D: a candidate operator to generate offspring and a variation of the neighborhood size for each subproblem. A bandit-based AOS is an upper confidence model that selects a strategy according to some determinist rule based on multi-armed bandit algorithms.

In this paper, we focus on the use of the multi-armed bandit algorithms to select (instead of the evolutionary operator) the most suitable SF in different stages of the evolutionary search process of MOEA/D-DRA.

IV. Our proposed approach

In this section, we define some guidelines to combine more than one SF simultaneously, using an adaptive strategy selection and collaborative subpopulations. As indicated before, the mechanism adopted is a bandit-based AOS algorithm coupled to MOEA/D-DRA [40]. We analyze the effect that our proposed approach has on the performance of MOEA/D-DRA when solving MOPs with up to 10 objectives and complicated Pareto shapes.

A. Pool of Strategies

As mentioned in Section II, the CHE and ASF methods find optimal solutions in opposite target directions. Based on this and our prior experience, only the SFs with similar target directions will be combined simultaneously. Otherwise, the algorithm can not generate well-distributed solutions along the Pareto front. We propose the next pool of strategies for combining multiple SFs:

- $S_1 = \{ACHE, MCHE, WN\}$,
- $S_2 = \{AASF, MASF, PBI\}$,

The model parameters were suggested based on the values proposed by different authors.

- $\alpha = \{0.0, 0.0001, 0.001, 0.01, 0.5\}$,
- $p = \{0.5, 1.0, 2.0, 3.0, 5.0, 10.0, 100.0\}$,
- $\theta = \{0.1, 1.0, 2.0, 5.0, 10.0, 50.0\}$

These values consider three special cases: $\alpha = 0$ to include the ASF method, $p = 1$ for the WS function and $p = 100$ which is similar to the CHE function. Thus, $|S_1| = 17$ and $|S_2| = 16$ is the number of different configurations of SFs.

B. Collaborative populations

We pre-defined a set of uniformly distributed weight vectors using the simplex lattice design technique [41], to maintain a good diversity in MOEA/D-DRA. These weight vectors are divided into k subsets used by a subpopulation which optimizes one of the types of SF selected from a pool of options (S_1 or S_2). Each weight vector is assigned alternately to each subpopulation in the same manner as was mentioned in [21] (single grid implementation technique).

Our proposed approach uses an AOS method for deciding which SF should be employed at a time point in MOEA/D-DRA to solve different MOPs.

C. Adaptive strategy selection

The performance of a decomposition-based algorithm strongly depends on the selection of its SF. Some types of SF are more beneficial for a particular Pareto front shape or a certain number of objectives. We employ a Multi-Arm Bandit (MAB) technique to select the most appropriate SF for each subproblem in a decomposition-based MOEA. MAB considers two main aspects: to assign a credit value to each operator and to select one operator based on its historical performance. In our case, we used the term "operator" to refer to a SF and its model parameter. We adopted MOEA/D-DRA coupled to AOS based on the fitness improvement rates (FIR) [38], which is computed by each subproblem i at time t, as defined by equation (11).

$$FIR_{i,t} = \frac{g(\mathbf{x}^i|\lambda^j, \mathbf{z}^*) - g(\mathbf{y}|\lambda^j, \mathbf{z}^*)}{g(\mathbf{x}^i|\lambda^j, \mathbf{z}^*)}\,, \tag{11}$$

where $\mathbf{x}^i$ is the current solution and $\mathbf{y}$ is its generated offspring solution after applying the genetic operators. Function g is a specific SF selected from a pool of strategies. The aim of the FIR technique is to deal with the largest ranges of raw fitness values at different stages of the evolutionary search process [38]. We used a sliding window with a fixed size W and a first-in, first-out (FIFO) queue structure in order to store the FIR values of the recently used SF and its model parameter.

The reward value assigned to each strategy i is given by:

$$FRR_{i,t} = \frac{Decay_i}{\sum_{j=1}^{k} Decay_j} \tag{12}$$

where $Decay_i = D^{rank_i} \times R_i$, $D \in [0, 1]$ is a decaying factor to increase the probability of selecting the best strategies, $rank_i$ is the rank assigned to each strategy (in descending order) and R_i (or reward) is the sum of all FIR values for each strategy i in the current sliding window.

We select the best SF using equation (13).

$$S_i = \text{arg max}\left\{\frac{R_{k,t}}{\sum_{i=1}^{k} R_i} + C \times \sqrt{2 \times \frac{ln(\sum_{j=1}^{m} \eta_j)}{\eta_i}}\right\} \tag{13}$$

where C is a weight factor to control the trade-off between exploration and exploration. η_i is the number of times that the strategy i was used.

D. Collaborative and adaptive strategies coupled to MOEA/D-DRA

MOEA/D with Dynamical Resource Allocation (MOEA/D-DRA) [40] is an improved version of MOEA/D [11], which was the winning algorithm in the CEC 2009 MOEA contest. MOEA/D-DRA incorporates a mechanism to compute the relative decrease of the objectives for each subproblem in order to assign computational effort according to the obtained benefits.

Next, we describe how to couple the adaptive strategy selection to MOEA/D-DRA. The first step is an initialization process. Then, we split the weight vectors into each subpopulation in order to assign a type of SF. Next, we use the adaptive strategy selection and associate one scalarizing function to each subpopulation in order to assign a different type of SF to each of them. At each generation, we monitor the FIR value for each subproblem of each subpopulation. Next, a dynamical resource allocation mechanism used in the original MOEA/D-DRA is applied. After that, a generation of new solutions via Differential Evolution and Polynomial-based mutation is employed. Finally, we update the reward subpopulation based on the FIR and Algorithm 2, while Algorithm 3 illustrates in more detail the steps described before.

Algorithm 1: Our proposed bandit-based operator selection mechanism

Input: A pool of scalarizing functions
Output: The new selected strategy
if *There are scalarizing functions that have not been selected* **then**
  S_i = one scalarizing function, which is selected randomly from the pool of strategies.
else
  $S_i =$
  $$\arg\max \left\{ \frac{Reward_{k,t}}{\sum_{i=1}^{k} Reward_i} + C \times \sqrt{2 \times \frac{ln(\sum_{j=1}^{m} \eta_j)}{\eta_i}} \right\}$$

Algorithm 2: Credit assignment algorithm.

Input: D: decay factor
Output: The reward values for each strategy
Initialize each $Reward_i = 0$
Initialize $n_i = 0$
for $i \leftarrow 1$ *to* $slidingWindow.length$ **do**
  $S = slidingWindow.GetIndexOp(i)$
  $FIR = slidingWindow.GetFIR(i)$
  $Reward_s = Reward_s + FIR$
  n_s++
Rank $Reward_i$ in descending order and set $rank_i$ to be the rank value of strategy S_i
for $i \leftarrow 1$ *to* $|S|$ **do**
  $FRR_{i,t} = \frac{Decay_i}{\sum_{j=1}^{k} Decay_j}$

E. Experimental settings

We divide our experiments in two parts. The first is focused on variations of the Pareto front geometry and the number of objectives using some MOPs defined in the *Deb-Thiele-Laumanns-Zitzler (DTLZ)* test suite [42]: DTLZ1 for linear, DTLZ3 for non-convex, DTLZ5 for degenerate and DTLZ7 for mixed Pareto front geometries. Additionally, we transformed DTLZ3 to have a convex shape. To test the scalability of our proposal, each problem was tested with $\{2, 3, 5, 7$ and $10\}$ objectives. For the second set of experiments, we adopted complicated MOPs presented in the CEC 2009 contest. For a fair comparison, we used the same MOEA/D-DRA parameters in all the MOP instances. The neighborhood size T was set to 20% of the population size (p_size). The crossover and mutation parameters were set as: $F = 0.5$, $Cr = 1$. For the DTLZ test problems, $H = \{99, 14, 6, 4, 3\}$ used by the simplex lattice design method. The population size was set to: $pop_s = \{100, 120, 210, 210, 220\}$ for $m = \{2, 3, 5, 7, 10\}$. The number of objective function evaluations was set to 40,000 for $m = 2$ and $m = 3$ and it was 50,000 for all the other cases. In the case of the UF test functions (i.e., those from the CEC 2009 contest), we adopted a population size of 600 for $m = 2$ and of 1000 for $m = 3$. The number of function evaluations was set to 300,000.

Algorithm 3: MOEA/D-DRA-MSF

Input: A stopping criterion
ns: number of subpopulations
pop_size: population size
$\{\lambda_1, \ldots, \lambda_N\}$: A well-distributed set of weight vectors
$pool$: A pre-defined pool of strategies
Output: Pareto front estimation
Step 1. Initialization
$E_p \leftarrow \emptyset$
Compute the Euclidean distance between any two weighted vectors and then work out the T closest weighted vectors to each weighted vector.
Generate an initial set of subpopulations $P_i = \{x_1, \ldots x_{N/ns}\}$.
Evaluate indidividuals in the initial subpopulations.
Set the ideal point $z_i^* = \arg\min\{f_i(x)\}$.
while *the stopping criterion is not satisfied* **do**
  Step 2. Adaptive Strategy selection.
  Select ns scalarizing functions g and model parameters p from the pool of strategies, using Algorithm 1
  Associate one scalarizing function to each subpopulation, according to the partition strategy.
  for *each subpopulation* **do**
    Set $gen = 0$ and $\pi^i = 1$ for all $i = 1, \ldots, N$.
    Selection of subproblems for searching: the indexes of the subproblems whose objectives correspond to the MOP's individual objectives f_i are selected to form the initial I. By using a 10-tournament selection strategy based on π^i, select other $[\frac{N}{5}]m$ indexes and add them to I.
    for *each* $i \in I$, *do* **do**
      Step 3. Selection of Mating/Update Range:
      Uniformly randomly generate a number *rand* in the range $(0, 1)$.
      Then set
      $$P = \begin{cases} B(i) & \text{if } rand < \delta, \\ \{1, \ldots N\} & \text{otherwise} \end{cases}$$
      Step 4 Reproduction: Set $r_1 = i$ and randomly select two indexes r_2 and r_3 from P, and then generate a solution $\mathbf{y}$ from xr_1, xr_2 and xr_3 by a Differential Evolution (DE) operator, and then perform mutation on $\mathbf{y}$.
      Step 5. Update.
      Update $z*$:
      Update subproblem i, $u(\mathbf{f} : \mathbf{w}, \mathbf{z}^*, p)$
      Update Neighboring Solutions
      for $j \in B(i)$ **do**
        if $g(y'|w_j, z) \leq g(x|w_j, z)$ **then**
          compute FIR based on equation (11)
          set $x_j = y'$
      Step 6. Update Reward registry.
      Compute reward using Algorithm 2
      $gen = gen + 1$.
      If gen is a multiple of 50, then compute Δ_i, the relative decrease of the objective for each subproblem i during the last 50 generations, update
      $$\pi^i = \begin{cases} 1 & \text{if } \Delta^i > 0.001, \\ (0.95 + 0.05) * \frac{\Delta_i}{0.001}\pi^i & \text{otherwise} \end{cases}$$
$P \leftarrow$ non-dominated solutions from the population
return P

The parameters for the adaptive strategy selection were set as suggested in [38]. We used the decay factor $D = 1.0$ and a factor to control the exploitation and the exploration of $C = 5.0$ and $W = 0.5 \times p\_size$.

We performed 30 independent runs for each MOEA and problem instance. For comparing our results, we adopted the hypervolume indicator (HV) [2] to assess both convergence and maximum spread. We established the following reference points: $(1, 1, \ldots 1)^T$ for DTLZ1, $(7, 7, \ldots 7)^T$ for DTLZ3 and DTLZ3_$convex$, $(4, 4, \ldots 4)^T$ for DTLZ5, $(21, 21, \ldots 21)^T$ for DTLZ7.

$(2.0, \ldots, 2.0)^T$ for all the CEC 2009 test problems (UF1-10).

F. Discussion of results

In this subsection, we compare our proposed approach using multiple SFs (considering the two pools of strategies previously discussed) with respect to the original MOEA/D-DRA adopting only one SF (in our experiments, we used CHE, ASF, WN ($p = 2$) and PBI ($\theta = 5$)). Table I presents the hypervolume indicator for each DTLZ test problem. The best values are highlighted with a darker gray tone and the second best with a lighter tone. We applied the nonparametric statistic, Wilcoxon rank sum test with a 95% of confidence level to corroborate that the best result found is statistically significant respect to the others. The symbol ($\uparrow$) means that the best case (algorithmic configuration per problem) outperformed another algorithm in a significantly better way. The symbol ($\downarrow$) indicates that the difference between the best option and another algorithm is not significant.

We analyzed the results according to each proposed strategy (i.e., S_1 and S_2). In the same way, AASF, MAASF and PBI outperformed MOEA/D-DRA with ASF. One interesting observation is that the main improvements were obtained in the many-objective problems. For two objectives, the results were very similar among themselves. However, with 3 or more objectives there were some significant differences, especially in the seven- and ten-objective MOPs. We can notice that the standard deviation values increase when the MOP has multimodality such as in dtlz1, dtlz3 and dtlz3_convex. In general, the strategy S_1 is better than S_2 in the DTLZ test problems adopted.

In a second experiment, we compare our proposal with respect to state-of-the-art MOEAs such as the original version of MOEA/D-DRA [40], and with respect to MOEA/D-DRA-MAB [38] which used the same multi-armed bandit algorithm but applied to a pool of Differential Evolution operators. Moreover, we also compared results with respect to ADEMO/D [31] which includes a learning period strategy to select from a pool of DE operators. The source code for these algorithms was obtained from `https://coda-group.github.io/publications.html`.

For assessing performance, we computed the IGD+ [43] and HV indicators (see Table II), then we apply Wilcoxon test in the same manner as previous experiments. In almost all cases, MOEA/D-DRA with multiple SFs outperformed the other approaches with respect to which it was compared, but the most evident improvement was observed in UF5, UF6 and UF10 in both indicators (HV and IGD+).

V. Conclusions and Future Work

In this paper, we have presented a comparative study to determine some guidelines to combine, in a simultaneous way, several scalarizing functions. One of our most important contributions is that we have identified that the pool of scalarizing functions should establish the same target directions in order to generate well-distributed (i.e., uniform) solutions along the Pareto front.

Two interesting observations were the following:

1) The assignment of the weight vectors to each subpopulation produces a good distributions of solutions, and
2) properly setting the parameters of the SF per subpopulation, based on the AOS rule, provides a good convergence to the Pareto optimal front.

Also, we noticed that the appropriate parameters settings depend directly on the number of objectives and on the Pareto front geometry. We claim that the bandit-based AOS adopted in our proposed approach is a good option to detect appropriate parameters settings in SFs, while requiring a lower computational effort than the use of static parameter tuning strategies (offline tuning methods).

As part of our future work, we plan to explore other adaptive strategies and to compare them. We plan to study the effect of other AOS parameters values and types of SFs in our proposal. We are also interested in studying the relation between the weight vectors and the corresponding scalarizing functions.

Acknowledgments

The first author acknowledges support from CONACyT and CINVESTAV-IPN to pursue graduate studies in Computer Science. The second author gratefully acknowledges support from CONACyT project no. 221551.

References

[1] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA–II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, April 2002.

[2] E. Zitzler and L. Thiele, "Multiobjective Optimization Using Evolutionary Algorithms—A Comparative Study," in *Parallel Problem Solving from Nature V*, A. E. Eiben, Ed. Amsterdam: Springer-Verlag, September 1998, pp. 292–301.

[3] R. Hernández Gómez and 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*, ser. GECCO '15. New York, NY, USA: ACM, 2015, pp. 679–686.

[4] H. L. Liu, L. Chen, Q. Zhang, and K. Deb, "An evolutionary many-objective optimisation algorithm with adaptive region decomposition," in *2016 IEEE Congress on Evolutionary Computation (CEC)*, July 2016, pp. 4763–4769.

[5] H. Sato, "Inverted pbi in moea/d and its impact on the search performance on multi and many-objective optimization," in *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '14. New York, NY, USA: ACM, 2014, pp. 645–652. [Online]. Available: http://doi.acm.org/10.1145/2576768.2598297

[6] H. Ishibuchi, K. Doi, H. Masuda, and Y. Nojima, "Relation between weight vectors and solutions in moea/d," in *2015 IEEE Symposium Series on Computational Intelligence*, Dec 2015, pp. 861–868.

[7] H. Ishibuchi, N. Akedo, and Y. Nojima, "Behavior of Multiobjective Evolutionary Algorithms on Many-Objective Knapsack Problems," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 2, pp. 264–283, April 2015.

[8] A. A. M. no, C. A. C. Coello, and E. Mezura-Montes, "MODE-LD+SS: A Novel Differential Evolution Algorithm Incorporating Local Dominance and Scalar Selection Mechanisms for Multi-Objective Optimization," in *2010 IEEE Congress on Evolutionary Computation (CEC'2010)*. Barcelona, Spain: IEEE Press, July 18–23 2010, pp. 3284–3291.

[9] S. Zapotecas Martínez and C. A. Coello Coello, "A Direct Local Search Mechanism for Decomposition-based Multi-Objective Evolutionary Algorithms," in *2012 IEEE Congress on Evolutionary Computation (CEC'2012)*. Brisbane, Australia: IEEE Press, June 10-15 2012, pp. 3431–3438.

TABLE I: Statistical results for strategies S_1 and S_2. We show the mean and the standard deviations (in parentheses)

m	MOP	CHE	WN($p = 2$)	MOEA/D-DRA-MSF1	ASF	PBI($\theta = 5$)	MOEA/D-DRA-MSF2
	dtlz1	↑8.7122e-01(7.1742e-04)	8.7296e-01(3.4337e-01)	↑8.7034e-01(1.3440e-01)	↓8.7140e-01(5.2209e-04)	↑8.7189e-01(4.6566e-04)	↓8.7114e-01(5.4564e-02)
	dtlz3	4.8138e+01(1.3129e-02)	↑4.2977e+01(1.2991e+00)	↑4.8125e+01(1.1017e+02)	↑4.8134e+01(1.4122e-02)	↑4.3496e+01(1.4836e+00)	↓4.8131e+01(1.4154e-02)
2	dtlz3*	↑4.8585e+01(1.4334e-02)	↑4.8008e+01(2.3724e+02)	↑3.9507e+01(1.6363e+00)	4.8622e+01(9.7710e-03)	↑4.8076e+01(1.1369e-01)	↓4.8570e+01(1.9431e-02)
	dtlz5	↑1.5209e+01(5.3291e-15)	↑1.5000e+01(0.0000e+00)	1.5210e+01(3.7268e-04)	↑1.4107e+01(2.7241e-05)	↑1.3207e+01(1.3291e-11)	↑1.5210e+01(2.0331e-03)
	dtlz7	↑3.6431e+02(1.2806e+01)	↑3.6581e+02(1.2689e+01)	3.6591e+02(1.2713e+00)	↑3.6591e+02(8.7256e+00)	↑3.6587e+02(9.5238e+00)	↑3.6589e+02(8.7220e+00)
	dtlz1	4.9629e-01(1.1656e-02)	↑3.3465e-01(1.3951e+01)	↑3.3000e-03(9.3354e-02)	↓4.9626e-01(7.1114e-02)	↑3.2531e-03(1.2794e-01)	↓4.3143e-03(1.0891e-01)
	dtlz3	↑5.1501e+01(5.5516e+01)	↑1.9355e-01(1.7713e+03)	1.4713e+02(9.5746e+02)	↑1.4584e+02(4.2952e+01)	↑9.4666e+00(1.9882e+02)	↑9.5988e+01(5.4480e+01)
3	3*	↑1.7266e+00(3.7764e+02)	9.0919e+01(1.4094e+03)	↓9.0237e+01(3.2449e+02)	↑8.5693e+01(3.6189e+02)	↑1.3596e+01(1.1116e+02)	↑3.3232e+01(8.4332e+01)
	dtlz5	4.3524e+01(3.3929e-02)	↑4.3373e+01(8.3057e-01)	↑4.3517e+01(1.9373e-02)	↑4.3224e+01(1.3812e-02)	↑4.2307e+01(4.4186e+00)	↑4.0095e+01(1.7415e+00)
	dtlz7	↑6.7217e+03(1.9440e+02)	↑6.4022e+03(2.9728e+02)	↑6.7218e+03(2.4975e+02)	↑6.7977e+03(1.9758e+02)	↑6.8171e+03(2.2749e+02)	6.8291e+03(1.6706e+02)
	dtlz1	↑4.9929e-01(1.5924e-01)	↑2.5117e-01(6.1529e+02)	↑4.9972e-01(1.7369e-01)	8.5079e-01(7.3003e-02)	↑2.9759e-01(2.3140e-01)	↑8.1317e-01(9.2416e-02)
	dtlz3	↑5.8410e+03(3.1219e+03)	↑5.6629e+03(1.9979e+05)	↑4.4650e+03(3.2435e+03)	1.3322e+04(1.2718e+03)	↑2.3804e+03(4.8386e+04)	↑1.1450e+04(2.0792e+03)
5	3*	↑1.2414e+04(5.1761e+04)	1.6408e+04(4.0656e+10)	↑8.0325e+03(2.0862e+03)	↑6.8641e+03(2.5797e+03)	↓1.1697e+04(1.7868e+03)	↑1.2897e+04(1.2760e+03)
	dtlz5	↑4.0830e+02(1.4979e+02)	↓6.0898e+02(2.3057e+01)	↑6.1298e+02(5.5498e+01)	↑2.8630e+02(8.1132e+01)	6.1298e+02(1.8570e+01)	↑6.2335e+02(1.1285e+02)
	dtlz7	↑2.3765e+06(1.6408e+05)	↑2.4342e+06(2.1796e+05)	2.5201e+06(1.3996e+05)	↑2.3798e+06(1.6727e+05)	↑2.3684e+06(1.5960e+05)	↑2.3782e+06(1.3541e+05)
	dtlz1	↑3.8993e-01(8.9484e-02)	1.3453e+00(2.8435e+01)	↑4.9979e-01(1.3579e-01)	↑6.9847e-01(5.0892e-02)	↑7.3862e-01(4.2553e-02)	↓1.2867e-00(8.1133e-02)
	dtlz3	↑1.2494e+05(1.6655e+05)	↑1.1803e+05(7.6051e+06)	5.8770e+05(6.9250e+04)	↑4.7643e+05(7.2851e+04)	↑7.7523e+04(1.8189e+06)	↑5.3980e+05(6.6100e+04)
7	3*	↑4.2864e+05(9.2783e+05)	↑6.6023e+05(3.6630e+04)	1.1480e+06(2.3005e+09)	↑4.5671e+05(9.6351e+04)	↑2.4514e+05(1.1749e+05)	↑4.5453e+05(9.9030e+04)
	dtlz5	↑2.3702e+03(4.3165e+03)	↑3.6759e+03(2.3452e+03)	↑9.2144e+03(9.7608e+02)	↑4.2746e+03(1.0812e+03)	↑4.2830e+03(1.1001e+03)	9.5339e+03(1.4573e+03)
	dtlz7	↑7.6908e+08(7.9234e+07)	↑7.6876e+08(1.3667e+08)	↑9.0892e+08(5.8621e+07)	↑7.7546e+08(9.7418e+07)	9.1073e+08(6.6888e+07)	↓8.4123e+08(7.5738e+07)
	dtlz1	↑3.7893e-01(9.4529e-02)	↓5.0146e-01(8.2056e-02)	3.0936e+00(9.6209e+01)	↑3.5268e-01(9.9343e-02)	↑7.2013e-01(3.1656e-02)	↑6.5351e-01(5.9505e-02)
	dtlz3	↑5.7924e+07(5.0050e+07)	↑8.1510e+04(2.3712e+09)	1.6133e+08(2.8166e+07)	↑2.7653e+07(5.1446e+07)	↑8.1303e+07(1.8458e+08)	↑8.2516e+07(4.4397e+07)
10	3*	↑4.7820e+07(1.6563e+08)	1.0000e+09(2.3990e+12)	↑2.1945e+08(1.5715e+07)	↑1.6077e+08(8.0149e+07)	↑1.0528e+08(4.3691e+07)	↑2.3384e+08(4.9907e+07)
	dtlz5	↑1.3934e+05(2.8259e+05)	↓2.3465e+05(1.4312e+05)	↓3.7277e+05(1.0007e+05)	↑3.9270e+05(4.3071e+04)	5.6638e+05(6.6277e+04)	↑5.5157e+05(2.7411e+04)
	dtlz7	1.0000e+09(6.9947e+11)	1.0000e+09(6.0816e+12)	1.0000e+09(1.1297e+12)	1.0000e+09(9.3084e+11)	1.0000e+09(7.3536e+11)	1.0000e+09(6.8013e+11)

TABLE II: Statistical results for strategies S_1 and S_2. We show the mean and the standard deviations (in parentheses)

MOP	MOEA/D-DRA	MOEA/D-DRA-MAB	ADEMO/D	MOEA/D-DRA-MSF1	MOEA/D-DRA-MSF2
Hypervolume indicator					
UF1	↑3.5847(0.0587)	↓3.4397(0.0895)	↑3.5791(0.0373)	3.6611(0.0015)	↑3.6335(0.0204)
UF2	↑3.6026(0.0344)	↑3.5965(0.0193)	↑3.6325(0.0137)	3.6532(0.0112)	↓3.6396(0.0346)
UF3	↑3.4353(0.1568)	↑3.2093(0.1340)	3.4164(0.1265)	3.6551(0.0259)	↑3.3769(0.0696)
UF4	↑3.1783(0.0135)	↑3.1978(0.0107)	↓3.2337(0.0126)	↑3.2566(0.0101)	3.2878(0.0030)
UF5	↑0.7428(0.8705)	↑1.7446(0.2659)	↑1.8762(0.2609)	↑2.7135(0.3376)	3.0424(0.1003)
UF6	↑2.5232(0.2259)	↑2.6236(0.2125)	↑2.7444(0.1814)	↑2.8315(0.4389)	3.1201(0.1323)
UF7	↑3.4408(0.0391)	↑3.2580(0.3566)	↑3.3883(0.2606)	3.4870(0.0316)	↑3.4850(0.0194)
UF8	↑6.9568(0.3853)	↑6.9779(0.3591)	↓7.3229(0.0242)	7.4003(0.0226)	↑7.3145(0.0017)
UF9	↑6.9542(0.3346)	↑7.2106(0.2974)	↓7.2993(0.2163)	↓7.3362(0.2213)	7.6577(0.0348)
UF10	↑0.7238(1.0449)	↑4.4226(0.7018)	↑4.7172(0.9373)	↑4.0287(0.6905)	6.1128(0.1576)
IGD+ indicator					
UF1	↑0.0389(0.0291)	↑0.0392(0.0126)	↑0.0142(0.0033)	0.0013(0.0001)	↑0.0063(0.0023)
UF2	↑0.0271(0.0131)	↑0.0077(0.0011)	↑0.0035(0.0006)	0.0034(0.0012)	↑0.0040(0.0028)
UF3	↑0.1476(0.0986)	↑0.0712(0.0240)	↑0.0531(0.0256)	0.0037(0.0051)	↑0.0500(0.0092)
UF4	↑0.0559(0.0045)	↑0.0449(0.0015)	↑0.0333(0.0014)	↑0.0271(0.0019)	0.0172(0.0003)
UF5	↑1.2284(0.4610)	↑0.5648(0.1253)	↑0.5132(0.1058)	↓0.2564(0.1017)	0.2167(0.0385)
UF6	↑0.4302(0.1407)	↑0.1466(0.0741)	0.1211(0.0413)	↑0.1686(0.1105)	↑0.1428(0.0717)
UF7	↑0.0217(0.0152)	↑0.0603(0.0836)	↑0.0293(0.0638)	0.0019(0.0018)	↑0.0044(0.0011)
UF8	↑0.0693(0.0258)	↓0.0690(0.0213)	↑0.0348(0.0098)	0.0115(0.0108)	↑0.0526(0.0004)
UF9	↑0.2455(0.0462)	↑0.2217(0.0637)	↑0.2098(0.0494)	↑0.2123(0.0487)	0.1372(0.0048)
UF10	↑1.7270(0.7275)	↑0.2542(0.0528)	↓0.2394(0.0724)	↑0.2571(0.0757)	0.1544(0.0288)

[10] A. Lara, S. Alvarado, S. Salomon, G. Avigad, C. A. Coello Coello, and O. Schütze, "The Gradient Free Directed Search Method as Local Search within Multi-Objective Evolutionary Algorithms," in *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation II*, O. Schütze, C. A. Coello Coello, A.-A. Tantar, E. Tantar, P. Bouvry, P. Del Moral, and P. Legrand, Eds. Berlin, Germany: Springer, Advances in Intelligent Systems and Computing Vol. 175, 2012, pp. 153–168, iSBN 978-3-642-31519-0.

[11] Q. Zhang and H. Li, "MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, December 2007.

[12] H. Li and Q. Zhang, "Multiobjective Optimization Problems with Complicated Pareto Sets, MOEA/D and NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 284–302, April 2009.

[13] H. L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 450–455, June 2014.

[14] K. Li, K. Deb, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 5, pp. 694–716, Oct 2015.

[15] R. Wang, T. Zhang, and B. Guo, "An Enhanced MOEA/D Using Uniform Directions and a Pre-organization Procedure," in *2013 IEEE Congress on Evolutionary Computation (CEC'2013)*. Cancún, México: IEEE Press, 20-23 June 2013, pp. 2390–2397, iSBN 978-1-4799-0454-9.

[16] C.-M. Chen, Y.-p. Chen, and Q. Zhang, "Enhancing moea/d with guided mutation and priority update for multi-objective optimization," in *Proceedings of the Eleventh Conference on Congress on Evolutionary Computation*, ser. CEC'09. Piscataway, NJ, USA: IEEE Press, 2009, pp. 209–216.

[17] H. Sato, "Analysis of inverted pbi and comparison with other scalarizing functions in decomposition based moeas," *Journal of Heuristics*, vol. 21, no. 6, pp. 819–849, 2015.

[18] I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "Generalized Decomposition," in *Evolutionary Multi-Criterion Optimization, 7th International Conference, EMO 2013*, R. C. Purshouse, P. J. Fleming, C. M. Fonseca, S. Greco, and J. Shaw, Eds. Sheffield, UK: Springer. Lecture Notes in Computer Science Vol. 7811, March 19-22 2013, pp. 428–442.

[19] B. Derbel, D. Brockhoff, A. Liefooghe, and S. Verel, *On the Impact of Multiobjective Scalarizing Functions*. Cham: Springer International Publishing, 2014, pp. 548–558.

[20] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Adaptation of Scalarizing Functions in MOEA/D: An Adaptive Scalarizing Function-Based Multiobjective Evolutionary Algorithm," in *Evolutionary Multi-Criterion Optimization: 5th International Conference, EMO 2009, Nantes, France, April 7-10, 2009. Proceedings*, M. Ehrgott, C. M. Fonseca, X. Gandibleux, J.-K. Hao, and M. Sevaux, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 438–452.

[21] ——, "Simultaneous use of different scalarizing functions in moea/d," in *Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '10. New York, NY, USA: ACM, 2010, pp. 519–526. [Online]. Available: http://doi.acm.org/10.1145/1830483.1830577

[22] M. Zeleny, "Compromise programming," in *Multiple Criteria Decision Making*, L. Cochrane and M. Zeleny, Eds. Columbia, USA: University of South Carolina Press, 1973, pp. 262–301.

[23] L. Zadeh, "Optimality and Non-scalar-valued Performance Criteria," *IEEE*, vol. 8, no. 1, pp. 59–60, Jan 1963.

[24] G. Dellino, M. Fedele, and C. Meloni, *Dynamic Objectives Aggregation Methods in Multi-objective Evolutionary Optimization*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 85–103.

[25] R. Wang, Q. Zhang, and T. Zhang, "Pareto Adaptive Scalarising Functions for Decomposition Based Algorithms," in *Evolutionary Multi-Criterion Optimization: 8th International Conference, EMO 2015, Guimarães, Portugal, March 29 –April 1, 2015. Proceedings, Part I*, A. Gaspar-Cunha, C. Henggeler Antunes, and C. C. Coello, Eds. Cham: Springer International Publishing, 2015, pp. 248–262.

[26] V. B. Jr, "On the Relationship of the Tchebycheff Norm and the Efficient Frontier of Multiple-criteria Objectives," in *Lecture Notes in Economic and Mathematical Systems*, H. Thiriez and S. Zionts, Eds. Berlin, Germany: Springer, 1973, vol. 130, pp. 76–86.

[27] W. Huang and H. Li, "On the differential evolution schemes in moea/d," in *2010 Sixth International Conference on Natural Computation*, vol. 6, Aug 2010, pp. 2788–2792.

[28] C.-M. Chen, Y. ping Chen, and Q. Zhang, "Enhancing MOEA/D with Guided Mutation and Priority Update for Multi-Objective Optimization," in *2009 IEEE Congress on Evolutionary Computation (CEC'2009)*. Trondheim, Norway: IEEE Press, May 2009, pp. 209–216.

[29] Q. Zhang, W. Liu, and H. Li, "The Performance of a New Version of MOEA/D on CEC09 Unconstrained MOP Test Instances," in *2009 IEEE Congress on Evolutionary Computation (CEC'2009)*. Trondheim, Norway: IEEE Press, May 2009, pp. 203–208.

[30] T. C. Chiang and Y. P. Lai, "Moea/d-ams: Improving moea/d by an adaptive mating selection mechanism," in *2011 IEEE Congress of Evolutionary Computation (CEC)*, June 2011, pp. 1473–1480.

[31] S. M. Venske, R. A. Goncalves, and M. R. Delgado, "ADEMO/D: Multiobjective optimization by an adaptive differential evolution algorithm," *Neurocomputing*, vol. 127, pp. 65–77, March 15 2014.

[32] R. E. Steuer and E.-U. Choo, "An Interactive Weighted Tchebycheff Procedure for Multiple Objective Programming," *Mathematical Programming*, vol. 26, no. 3, pp. 326–344, 1983.

[33] I. Kaliszewski, "A modified weighted tchebycheff metric for multiple objective programming," *Comput. Oper. res.*, vol. 14, no. 4, pp. 315–323, 1987.

[34] K. Miettinen, *Nonlinear Multiobjective Optimization*. International Series in Operations Research & Management Science, 1998, vol. 12.

[35] A. P. Wierzbicki, "The Use of Reference Objectives in Multiobjective Optimization," in *Multiple Criteria Decision Making Theory and Application: Proceedings of the Third Conference Hagen/Königswinter*, G. Fandel and T. Gal, Eds. Berlin, Germany: Springer, August 20-24 1980, pp. 468–486.

[36] M. Pescador-Rojas, R. Hernández Gómez, E. Montero, N. Rojas-Morales, M.-C. Riff, and C. A. Coello Coello, "An Overview of Weighted and Unconstrained Scalarizing Functions," in *Evolutionary Multi-Criterion Optimization, 9th International Conference, EMO 2017*, H. Trautmann, G. Rudolph, K. Klamroth, O. Schütze, M. Wiecek, Y. Jin, and C. Grimme, Eds. Münster, Germany: Springer. Lecture Notes in Computer Science Vol. 10173, March 19-22 2017, pp. 499–513, iSBN 978-3-319-54156-3.

[37] R. H. Gómez and C. A. C. Coello, "A hyper-heuristic of scalarizing functions," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '17. New York, NY, USA: ACM, 2017, pp. 577–584.

[38] K. Li, . Fialho, S. Kwong, and Q. Zhang, "Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 1, pp. 114–130, Feb 2014.

[39] Y. Qi, L. Bao, X. Ma, Q. Miao, and X. Li, "Self-adaptive multi-objective evolutionary algorithm based on decomposition for large-scale problems," *Inf. Sci.*, vol. 367, no. C, pp. 529–549, Nov. 2016.

[40] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of moea/d on cec09 unconstrained mop test instances," in *Proceedings of the Eleventh Conference on Congress on Evolutionary Computation*, ser. CEC'09. Piscataway, NJ, USA: IEEE Press, 2009, pp. 203–208. [Online]. Available: http://dl.acm.org/citation.cfm?id=1689599.1689626

[41] H. Scheffé, "Experiments with Mixtures," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 20, p. 344360, 1958.

[42] S. Huband, P. Hingston, L. Barone, and L. While, "A Review of Multiobjective Test Problems and a Scalable Test Problem Toolkit," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 5, pp. 477–506, October 2006.

[43] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, *Modified Distance Calculation in Generational Distance and Inverted Generational Distance*. Cham: Springer International Publishing, 2015, pp. 110–125.