

A Hyper-Heuristic of Scalarizing Functions

Raquel Hernández Gómez
Computer Science Department
CINVESTAV-IPN
Mexico City, Mexico 07360
rhernandez@computacion.cs.cinvestav.mx

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

ABSTRACT

Scalarizing functions have been successfully used by Multi-Objective Evolutionary Algorithms (MOEAs) for the fitness assignment process. Their popularity has to do with their low computational cost, their capability to generate (weakly) Pareto optimal solutions, and their effectiveness in solving many-objective optimization problems. Nevertheless, recent studies indicate that the search behavior of MOEAs strongly depends on the choice of the scalarizing function. Besides, this specification varies according to the Pareto-front geometry of the problem at hand. In this work, we present a novel hyper-heuristic for continuous search spaces, which combines the strengths and compensates for the weaknesses of different scalarizing functions. These heuristics have been proposed within the evolutionary multi-objective optimization and mathematical programming communities. Furthermore, the selection of heuristics is conducted through the s -energy, which measures the even distribution of a set of points in k -dimensional manifolds. Experimental results indicate that our proposed approach outperforms the use of a single heuristic as well as other state-of-the-art algorithms in the majority of the ZDT, DTLZ and WFG test problems.

CCS CONCEPTS

•Computing methodologies → Continuous space search;

KEYWORDS

Multi-objective optimization, Genetic algorithms, Heuristics, Selection

ACM Reference format:

Raquel Hernández Gómez and Carlos A. Coello Coello. 2017. A Hyper-Heuristic of Scalarizing Functions. In *Proceedings of GECCO '17, Berlin, Germany, July 15-19, 2017*, 8 pages.
DOI: <http://dx.doi.org/10.1145/3071178.3071220>

1 INTRODUCTION

In this work, we focus on Multi-objective Optimization Problems (MOPs), which comprise the simultaneous minimization of several, often conflicting, objective functions of the form:

$$\text{Minimize } \{f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})\} \quad \text{subject to } \mathbf{x} \in \mathcal{X}, \quad (1)$$

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

GECCO '17, Berlin, Germany

© 2017 ACM. 978-1-4503-4920-8/17/07...\$15.00

DOI: <http://dx.doi.org/10.1145/3071178.3071220>

where \mathbf{x} is the *decision vector*, $(f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T \in \mathcal{Y}$ represents the vector of $m(\geq 2)$ objective functions, $\mathcal{X} \subset \mathbb{R}^n$ is the *decision (variable) space*, and $\mathcal{Y} \subset \mathbb{R}^m$ is the *objective space*. MOPs having four or more objectives are widely known as many-objective optimization problems [19, 23], nowadays considered as a hot research topic. The solution to a MOP consists of finding a set of decision vectors that satisfy the property of Pareto optimality.¹ This is called the Pareto Optimal Set (POS), and its corresponding image in \mathcal{Y} is named the Pareto Optimal Front (POF).

Throughout the years, Multi-Objective Evolutionary Algorithms (MOEAs) have successfully shown their effectiveness in solving MOPs [2, 18, 22]. They can find discrete approximations to the POS in a single run without requiring particular assumptions, such as continuity or differentiability. In fact, MOEAs perform random search strategies that operate under Darwin's principle of natural selection, where the fittest individuals must achieve: 1) convergence to the POF and 2) uniform distribution along the objective space (diversity). One of the most preferred methods for fitness assignment is through a *Scalarizing Function (SF)* $\mathbb{R}^m \rightarrow \mathbb{R}$ [21], which reformulates the MOP into several single-objective subproblems with the aid of fixed target directions, also known as *weight vectors*. The idea behind a SF is that convergence is obtained by minimizing a certain distance metric to the *ideal point*², while diversity is kept by the set of weight vectors.

In the literature, we can find some MOEAs that incorporate one or two fixed SFs within their search engine, such as MOGLS [15] and MSOPS [13], respectively. Other algorithms operate as frameworks, where the decision maker must select the SF that best suits his (her) needs, such as MOEA/D [28]. The recent popularity of SFs has to do with their effectiveness in solving many-objective optimization problems at a low computational cost. SFs also present advantages with respect to conventional MOEAs. For instance, approaches based on Pareto optimality lose their ability to converge as the number of objectives increases, whereas algorithms based on the hypervolume indicator [29] are computationally too expensive to deal with [19, 23]. Another advantage of SFs is their nice mathematical properties. Most of them can generate at least weakly Pareto optimal solutions.³ In addition, some SFs can capture the whole POF by varying either the weights [6] or their parameter models, which control the curvature of their contour lines [20].

Nevertheless, despite the evident advantages of MOEAs relying on SFs, three crucial issues should be considered for a problem at hand: 1) the setting of the weight vectors, 2) the choice of an

¹ A decision vector $\mathbf{x}^* \in \mathcal{X}$ is Pareto optimal if $\nexists \mathbf{x} \in \mathcal{X}$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \dots, m$ and $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$ for at least one index j .

² $z_i := \min \{f_i(\mathbf{x}) \mid \mathbf{x} \in \mathcal{X}\} \forall i \in \{1, \dots, m\}$.

³ A decision vector $\mathbf{x}^* \in \mathcal{X}$ is weakly Pareto optimal if $\nexists \mathbf{x} \in \mathcal{X}$ such that $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$ for all $i = 1, \dots, m$.

appropriate SF, and, if required, 3) the setting of the model parameters. About the first issue, it has been observed that a uniform distribution of the weight vectors in $[0, 1]^m$ does not necessarily imply that approximation sets will exhibit good diversity [3]. For this reason, several efforts have focused on the adaptation of the weights (see for example [6, 8]). Regarding the second issue, some studies have exhaustively researched only the scalarizing functions WS, CHE, and PBI (see Table 1), which present some limitations. For example, WS can only capture convex Pareto fronts [20], CHE loses diversity for more than two objectives [11, 28], and, although PBI can find evenly distributed solutions from three objectives onwards [4, 28], its behavior depends on the parameter θ [27]. Moreover, recent experiments have shown that, for a given MOP, the search behavior of MOEAs strongly depends on the choice of these SFs [14, 16, 17, 21]. Finally, concerning the third issue, certain SFs require the specification of some model parameters, which may be sensitive to the Pareto-front shape. Thus, some attempts have emerged to adjust them dynamically [26, 27].

In this work, we focus on the second issue (the choice of an appropriate SF) and propose a hyper-heuristic [1], which combines the strengths and compensates for the weaknesses of seven scalarizing functions. Our aims are to provide an algorithm that is more generally applicable than current implementations and also to release the decision maker from the difficult task of selecting an appropriate SF. The proposed hyper-heuristic can be seen as a search method for selecting low-level heuristics (i.e., SFs) to solve continuous MOPs.

The remainder of this paper is organized as follows. In Section 2, we provide a brief review of the previous related work. In Section 3, we explain our proposal. In Section 4, we compare the effectiveness of our hyper-heuristic with regard to MOEA/D [28], NSGA-III [4], and MOMBI-II [11] on the Zitzler-Deb-Thiele (ZDT) [30], the Deb-Thiele-Laumanns-Zitzler (DTLZ) [5] and the Walking-Fish-Group (WFG) [12] test suites. Finally we discuss our conclusions and future work in Section 5.

2 RELATED WORK

The idea of using the simultaneous collaboration of different SFs within MOEAs dates back to 2003, when there were a few attempts to combine, at most, two SFs.

Hughes [13] scored the population using the CHE and the vector angle distance scaling (also introduced in that work) for solving continuous MOPs with two and three objectives. Although this method, called Multiple Single Objective Pareto Sampling (MSOPS), can deal with any Pareto-front geometry, it may generate dominated solutions in disconnected regions, because the second SF is not compatible with any form of Pareto optimality.

Ishibuchi et al. [16] modified MOEA/D for automatically choosing between CHE and WS. The former was applied only for concave parts of the Pareto front, where local concavity was detected as long as an individual was identical to a certain number of neighbors. However, when frequent changes occurred, this approach did not perform well in combinatorial MOPs having up to 6 objectives. The reason was probably that both SFs drove individuals to entirely different regions when considering the same weight vector [16].

Table 1: Some scalarizing functions and their features. Pareto-front shapes are abbreviated to x (convex), c (concave) or l (linear). $<$ (\leq) denotes compatibility with (weak) Pareto optimality. \parallel means that the optimal objective vector \mathbf{y}^* is nearly parallel to the weight vector \mathbf{w} .

Acronym	Full Name	Minimize $u(\mathbf{y}; \mathbf{w}) :=$	Support	Model Parameter
WS	weighted sum	$\sum_i w_i y_i$	x <	-
EWC	exponential weighted criteria	$\sum_i (e^{p w_i} - 1) e^{p y_i}$	x, c, l <	$p = 100$
WPO	weighted power	$\sum_i \frac{(y_i)^p}{w_i}$	x, c, l < \parallel	$p = 3$
WN	weighted norm	$\left(\sum_i \frac{ y_i ^p}{w_i} \right)^{\frac{1}{p}}$	x, c, l < \parallel	$p = 0.5$
CHE	chebyshev function	$\max_i \{w_i y_i \}$	x, c, l \leq	-
ASF	achievement scalarizing function	$\max \left\{ \frac{y_i}{w_i} \right\}$	x, c, l \leq \parallel	-
AASF	augmented achievement scalarizing function	$\max \left\{ \frac{y_i}{w_i} \right\} + \alpha \sum_i \frac{y_i}{w_i}$	x, c, l < \parallel	$\alpha = 1e^{-4}$
PBI	penalty boundary intersection	$d_1 + \theta d_2$, where $d_1 := \left \mathbf{y} \cdot \frac{\mathbf{w}}{\ \mathbf{w}\ } \right $ and $d_2 := \left\ \mathbf{y} - d_1 \frac{\mathbf{w}}{\ \mathbf{w}\ } \right\ $	x, c, l \parallel	$\theta = 5$

Shortly afterwards, this issue was further analyzed in [17], exploring two alternatives. In one approach an individual optimized a previously assigned SF, whereas in the second strategy the population increased linearly at the rate of $|P|$ individuals per SF. Each different subpopulation focused on a particular SF.

On the other hand, multi-objective hyper-heuristics have received little attention, particularly in continuous search spaces, where the pool of heuristics usually focuses on variation operators.

In 2015, Gonçalves et al. [9], presented the MOEA/D Hyper-Heuristic (MOEA/D-HH), which uses an adaptive choice function to determine the Differential Evolution (DE) strategy that should be applied to generate individuals at each iteration of a MOEA/D variant. These authors proposed the following choice function to compute the score of a given heuristic h :

$$CF(h) = \alpha \phi \left(f_1(h) + f_2(g, h) \right) + \delta f_3(h), \quad (2)$$

where g is the current heuristic, f_1 and f_2 are the mean rewards of applying h alone and g followed by h , respectively. The *reward* is calculated as the difference between the CHE function value of the parent and the child. f_3 corresponds to the time elapsed since h was last selected, α is a scale factor, which is problem dependent and needs to be calibrated a priori. ϕ and δ are parameters that control the intensification and diversification of the selection of the best heuristics, respectively. Here, δ is set to $1 - \phi$ and MOEA/D-HH automatically updates ϕ through generations. The heuristic with a higher CF value is chosen to create offspring, which is later perturbed by polynomial-based mutation. The pool of heuristics consists of five DE strategies: 1) *DE/rand/1/bin* of slow convergence

speed and good exploration capability, suitable for solving multi-modal problems; 2) *DE/current-to-rand/1/bin* for enabling the algorithm to solve rotated problems more effectively; 3) *DE/nonlinear*, which includes a non-linear part of the DE mutation operator; 4) *DE/rand/2/bin* and 5) *DE/current-to-rand/2/bin*, which may provide better perturbations than the two first strategies. In [9], this on-line selection hyper-heuristic was tested on ten unconstrained instances of the CEC'09 benchmark, improving the performance of MOEA/D using a single heuristic.

Walker and Keedwell [24] introduced the Multi-Objective Sequence-based Selection Hyper-Heuristic (MOSSH), which was the first hyper-heuristic designed to solve many-objective optimization problems. MOSSH is based on a hidden Markov model to determine the mutation heuristic to be applied for generating a single child from the current parent. Thus, this approach works as a (1+1)-Evolution Strategy complemented with an external archive, which keeps all the non-dominated solutions discovered so far. The pool of seven mutation heuristics consists primarily in: 1) adding noise to the current solution using three different continuous probability distributions, and 2) replacing the parent (or only a variable) with another one, either randomly created or taken from the archive. At each iteration, the child replaces the parent if the former dominates the second. However, in another paper [25], this comparison rule was changed by strategies based on the hypervolume indicator⁴, the favor relation⁵ and the average rank⁶. Moreover, the hidden Markov model is updated if the child is added to the archive and if it was better than the parent. In [25], the three strategies were independently applied to solve continuous many-objective problems (DTLZ) having up to 6 objectives. The best results were obtained using either the hypervolume or the favor relation. Although its computational cost is low (even when using the hypervolume in high dimensionality), this hyper-heuristic is reported to face difficulties in problems with disconnected Pareto-optimal regions in the search space (like DTLZ6). Its authors indicated that this might be due to the lack of crossover. Other disadvantages are that solutions are not uniformly distributed and the external archive may grow too much.

In contrast to these works, our proposed hyper-heuristic operates in the survival selection, the population size remains constant, and the pool of heuristics is scalable. Besides, it does not introduce new parameters, and no external archive is required. In the following section, we provide its design and its implementation details.

3 OUR PROPOSED APPROACH

The proposed hyper heuristic is an extension of the elitist Genetic Algorithm MOMBI-II (Many-Objective Metaheuristic Based on the R2 Indicator II) [11], now named MOMBI-III. This extension allows the inclusion of more than one scalarizing function. The core idea is that each individual in the population minimizes a distinct scalarizing function, having its own weight vector. Our algorithm can

⁴ A solution $\mathbf{x} \in \mathcal{X}$ is better than a solution $\mathbf{y} \in \mathcal{X}$, iff $\prod_i^m z_i - f_i(\mathbf{x}) > \prod_i^m z_i - f_i(\mathbf{y})$, where $\mathbf{z} \in \mathbf{R}^m$ is a reference point.

⁵ A solution $\mathbf{x} \in \mathcal{X}$ favors a solution $\mathbf{y} \in \mathcal{X}$ ($\mathbf{x} <_f \mathbf{y}$), iff $|\{i : f_i(\mathbf{x}) < f_i(\mathbf{y}), 1 \leq i \leq m\}| > |\{j : f_j(\mathbf{x}) > f_j(\mathbf{y}), 1 \leq j \leq m\}|$.

⁶ A solution $\mathbf{x} \in \mathcal{X}$ is better than a solution $\mathbf{y} \in \mathcal{X}$, iff $\sum_i^m r_i(\mathbf{x}) < \sum_i^m r_i(\mathbf{y})$, where $r_i \in \mathbf{N}$ is the rank of the solution according to the i^{th} objective. The ranking process considers the external archive.

be considered as a method that encourages convergence through the optimization of the SFs, while diversity is accomplished by two strategies: the weights and the heuristic selection.

The pool of heuristics consists of seven scalarizing functions $H = \{\text{WS}, \text{EWC}, \text{WPO}, \text{WN}, \text{CHE}, \text{ASF}, \text{AASF}\}$. Their mathematical expressions, features and model parameters appear in Table 1, whereas their corresponding contour lines are shown in Figure 1. The model parameters were established according to the values recommended in the literature. We selected this set of SFs due to their compatibility with some form of Pareto dominance,⁷ a requirement that PBI does not meet. For a graphical counterexample see Figure 1 ($\mathbf{a} < \mathbf{b}$, but $u(\mathbf{a}; \mathbf{w}) = u(\mathbf{b}; \mathbf{w})$). We also chose these SFs because their contour lines are very different, endowing our algorithm with an increased capacity to handle different Pareto-front geometries. It is worth mentioning that in the original versions of WPO and WN, the component w_i is being multiplied by y_i [21]. We did not adopt this form since the search is driven to different regions of the objective space, as it occurred in [11, 16].

Heuristic selection is achieved through the use of s -energy [10]:

$$E_s(A) := \sum_{i \neq j} \|\mathbf{a}_i - \mathbf{a}_j\|^{-s}, \quad (3)$$

where $A = \{\mathbf{a}_1, \dots, \mathbf{a}_{|A|}\}$, $\mathbf{a}_i \in \mathbf{R}^k$, and $s > 0$ is a fixed parameter. This measure has been used to discretize k -dimensional manifolds since its minimization leads to a uniform distribution of the points in A , if $s \geq k$ [10]. Recently, the s -energy has been employed only for comparison of MOEAs, setting $s = m - 1$, since the Pareto front is at most an $(m - 1)$ -manifold [8]. In order to know the individual contribution to the s -energy, we define:

$$\Delta E_s(\mathbf{a}, A) := \frac{E_s(A) - E_s(A \setminus \{\mathbf{a}\})}{2}, \quad (4)$$

fulfilling that $E_s(A) = \sum_i \Delta E_s(\mathbf{a}_i, A)$.

The main loop of MOMBI-III is presented in Algorithm 1. First, the population is initialized uniformly at random. Thereafter, it is evaluated according to the MOP definition. At each iteration, new individuals are created from parents selected uniformly at random (lines 4 and 5). These individuals are evaluated and added to the overall population in lines 6 and 7. Next, the reference points are updated. Here, we used the ideal point for \mathbf{z}^{min} and the nadir point for \mathbf{z}^{max} . The latter was calculated in the following way. First, we looked for those individuals that minimized WPO and AASF, using as weights the unitary vectors parallel to the axes. The maximum values of each of the objective components corresponding to these individuals constitute the point \mathbf{z}^{max} . We ensured that these individuals are different to each other and that \mathbf{z}^{max} enclosed at least $|P|$ members of the population. If these requirements were not satisfied we updated \mathbf{z}^{max} with the worst objective values of the whole population. In line 9, the objective function values are normalized using the reference points. In the following steps, the population is ranked and reduced to the desired size.

Algorithm 2 scores the population according to the pool of heuristics H . In line 1, the ranks are initialized and in lines 2 to 10 those individuals having the best value for each scalarizing function and weight vector obtain the first rank. In Algorithm 3, the population

⁷ A solution $\mathbf{x} \in \mathcal{X}$ dominates a solution $\mathbf{y} \in \mathcal{X}$, denoted by $\mathbf{x} < \mathbf{y}$ or $f(\mathbf{x}) < f(\mathbf{y})$, iff $f_i(\mathbf{x}) \leq f_i(\mathbf{y})$ for all $i = 1, \dots, m$ and $f_j(\mathbf{x}) < f_j(\mathbf{y})$ for at least one index j .

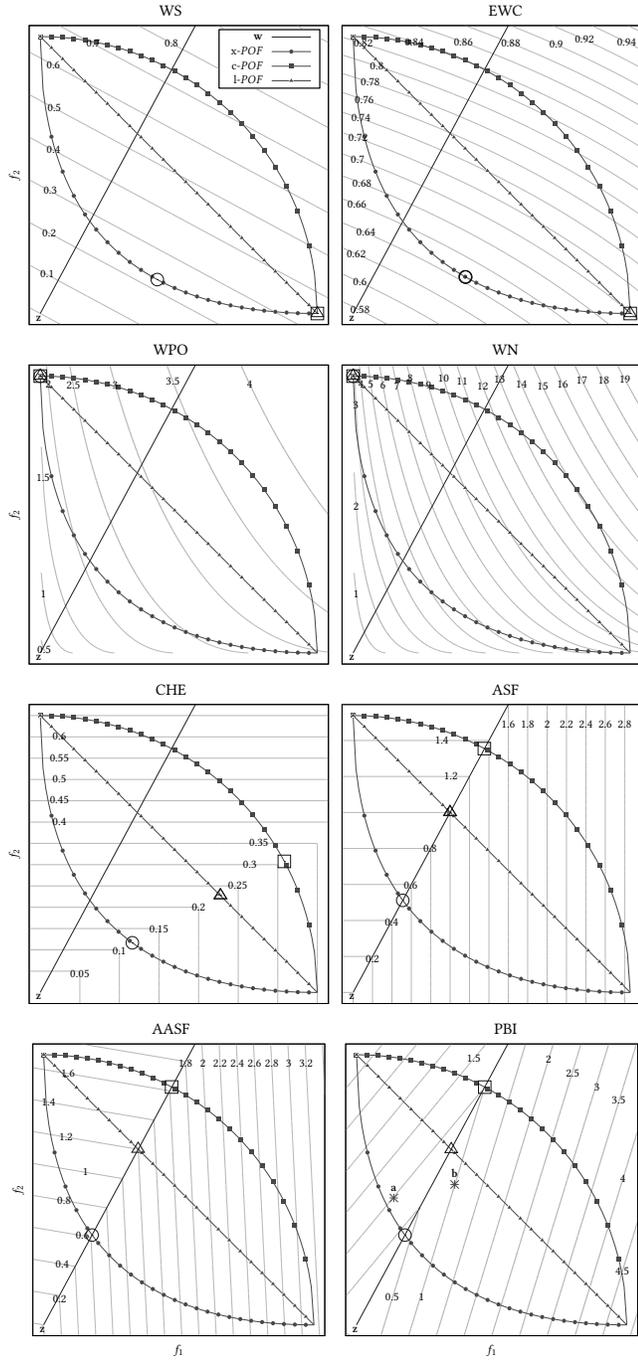


Figure 1: Contour lines of some scalarizing functions for the weight vector $w = (0.35, 0.65)$. The non-filled shapes denote optimal solutions for each of the Pareto-front shapes.

is first sorted and partitioned in layers using the ranks. A layer contains individuals with the same rank. We notice that ties are broken using Pareto dominance. With this relation we can avoid weakly-Pareto solutions in an effective manner. In lines 2 to 10, the individuals belonging to the worst ranks are removed from the

Algorithm 1 Main Loop of MOMBI-III

Input: MOP, stopping criterion, weight vectors W , heuristics H

Output: Final population P

- 1: Initialize population P at random
 - 2: Evaluate MOP for each $p \in P$
 - 3: **while** the stopping criterion is not satisfied **do**
 - 4: Select random parents from P
 - 5: $P' \leftarrow$ Generate offspring using variation operators
 - 6: Evaluate MOP for each $p \in P'$
 - 7: $Q \leftarrow P \cup P'$
 - 8: Update the reference points z^{min} and z^{max}
 - 9: Normalize objective functions by setting

$$p.y \leftarrow \frac{p.y - z^{min}}{z^{max} - z^{min}}, \forall p \in Q, \text{ where } p.y \in \mathbb{R}^m$$
 - 10: $R \leftarrow$ R2 Ranking (Q, W, H)
 - 11: $P \leftarrow$ Reduce ($Q, R, |P|$)
 - 12: **return** P
-

Algorithm 2 R2 Ranking

Input: Population Q , weight vectors W , pool of heuristics H

Output: Ranks R

- 1: $R[p] \leftarrow \infty \quad \forall p \in Q$
 - 2: **for all** $h \in H$ **do**
 - 3: **for all** $w \in W$ **do**
 - 4: **for all** $p \in Q$ **do**
 - 5: $p.\mu \leftarrow h(p.y; w)$
 - 6: Sort Q w.r.t. the field μ in increasing order
 - 7: $rank \leftarrow 1$
 - 8: **for all** $p \in Q$ **do**
 - 9: $R[p] \leftarrow \min\{R[p], rank\}$
 - 10: $rank \leftarrow rank + 1$
-

population, either by discarding an entire layer or one individual at a time. In case that more than two individuals have the same rank, we remove the one with the highest contribution to the s -energy. It is important to mention that for one set of high-fidelity objective-function evaluations, MOMBI-III can derive seven pieces of information with very low computation effort. With a careful implementation of the s -energy, the complexity of the algorithm is $O(|P|^2 m + |H| |W| |P| (\log |P| + m))$. It is assumed that $|H| \ll |P|$ and $|W| = |P|$. Thus the complexity is reduced to $O(|P|^2 (\log |P| + m))$.

4 RESULTS

In this section, we investigate the effectiveness of our proposed hyper-heuristic, considering five instances of the ZDT test suite [30], seven instances of the DTLZ benchmark [5], and all nine problems of the WFG test suite [12]. The experiments were divided in three parts: comparison with single heuristic versions (Subsection 4.2), contrast with some state-of-the-art MOEAs (Subsection 4.3), and a case study for many-objective optimization (Subsection 4.4). The first two experiments consider 2 and 3 objectives for the ZDT and DTLZ/WFG test problems, respectively, whereas the last experiment covers 4 to 10 objectives. In the following, we provide further details about these experiments.

Algorithm 3 Reduce**Input:** Population Q , ranks R , desired size n **Output:** Reduced population Q

```

1:  $\{L_1, \dots, L_k\} \leftarrow$  Sort the population in layers with respect to  $R$ .
2: while  $|Q| > n$  do
3:   if  $|L_k| \leq |Q| - n$  then {Remove members of the  $k^{th}$  layer}
4:      $r \leftarrow L_k$ 
5:      $k \leftarrow k - 1$ 
6:   else
7:     Compute the contribution to the  $s$ -energy
        $E[p.y] \leftarrow \Delta E_{m-1}(p.y, Q)$  for all  $p \in \{L_1 \cup \dots \cup L_k\}$ 
8:      $r \leftarrow \arg \max_{p \in L_k} E[p.y]$ 
9:      $L_k \leftarrow L_k \setminus \{r\}$ 
10:     $Q \leftarrow Q \setminus \{r\}$ 
11: return  $Q$ 

```

4.1 Experimental Settings

The number of decision variables for ZDT1-3 was set to 30 and for ZDT4,6 was set to 10. In the case of the DTLZ benchmark, the number of decision variables was set to $m + l - 1$, where l was 5 for DTLZ1, 10 for DTLZ2-6, and 20 for DTLZ7. The parameters for the WFG benchmark are provided in Table 2.

For comparison purposes, we selected the algorithms MOEA/D [28] (based on decomposition), NSGA-III [4] (based on reference points) and MOMBI-II [11] (based on the $R2$ indicator). All adopted the same parameter values of Table 2. We select these optimizers since they have been reported to perform well in a wide range of problems. In addition, they share common features, such as the requirement of the set of weight vectors, which were generated using the method described in [28]. Here, each weight vector takes a value from $\{10^{-2}, 1/H, 2/H, \dots, H/H\}$ where $H \in \mathbf{N}$. The total number of vectors is represented by the combinatorial number C_{m-1}^{H+m-1} . Thus, to keep this number low, we adopt a cardinality similar to the population size (for 5 and 6 objectives the set was pruned using a clustering technique, whereas for 8 and 9 objectives two layers were employed, discarding duplicated points).

In the case of MOEA/D its scalarizing function was CHE for two objectives and PBI with $\theta = 5$ for the remaining objectives [4, 28]. The population of MOEA/D was normalized during its evolution in problems with different scale, such as the WFG test suite and DTLZ7, as suggested by its authors in the original paper. For MOMBI-II the scalarizing function was ASF, and its parameters were: record 5, tolerance threshold 1×10^{-3} and 0.5 for the variance threshold. The parameter models adopted for MOMBI-III and all its variants with single heuristics are established in Table 1.

The variation operators were: Polynomial-based mutation and Simulated Binary Crossover (SBX). For the mutation operator, its probability and distribution index were set to $1/n$ and 20, respectively. For the crossover operator, these parameters varied according to the number of objectives, for two were 0.9 and 20, whereas for higher dimensionality were 1.0 and 30. The stopping criterion consisted of reaching a maximum number of evaluations of the MOP (see Table 2).

Table 2: Parameters adopted in our experiments

Objectives (m)	2	3	4	5	6	7	8	9	10
Population size	100	136	166	180	200	210	230	250	266
Objective function evaluations ($\times 10^3$)	40	60	70	80	80	90	100	100	110
WFG variables (n)	24	26	28	30	32	34	36	38	40
WFG position-related parameters	2	2	3	4	5	6	7	8	9
Weight-vector partitions (H)	99	15	8	6	5	4	3,4	3,4	2,3
MOEA/D niche	20	27	33	36	40	42	46	50	53

Finally, for the performance assessment of the algorithms, we relied on the hypervolume indicator [29], which measures convergence and diversity at the same time. The reference points for the hypervolume indicator were $(4, 4, \dots)$ for DTLZ3, $(2, 2, \dots, 2, 8)$ for DTLZ7, $(3, 5, 7, \dots)$ for WFG, and $(2, 2, \dots)$ for ZDT, DTLZ1,2, and DTLZ4-6. These points are slightly worse in all objectives than the nadir point. The hypervolume indicator has some bias for favoring non-linear Pareto fronts with clusters near the middle point (knee region). To compensate this situation, we also adopted the $(m-1)$ -energy (see expression 3), which rewards even distributions; and the Solow-Polasky indicator [7], defined by:

$$SPI(A) := (1, \dots, 1)_{1 \times |A|} C^{-1} (1, \dots, 1)_{|A| \times 1}^T \quad (5)$$

where C is a matrix with the entries $c_{i,j} := e^{-\theta \|a_i - a_j\|}$, for all $i = 1, \dots, |A|$ and $j = 1, \dots, |A|$. θ is a normalizing parameter, which was set to 10 in all experiments [7]. This indicator is to be maximized and it has been used to measure biological diversity (its value can be interpreted as the number of species). Due to the space limitation, the complete study is available for download at <http://delta.cs.cinvestav.mx/~ccoello/2017.html>.

We performed 30 independent runs for each MOEA and test problem. We applied the Wilcoxon rank sum test (one-tailed) to the mean of these indicators, in order to determine whether if MOMBI-III performed better (\uparrow) or not (\downarrow) than the other approaches at the confidence interval of 99%. Moreover, due to multiple comparisons, this value was adjusted by the Bonferroni correction.

4.2 Single Heuristics

Hyper-heuristics are meant to perform better than their constitutive heuristics in a wide range of problems [1]. Thus, motivated by this requirement, MOMBI-III was compared with the pool of heuristics. In this experiment, each SF was coupled independently to MOMBI-III. The results are presented in Tables 3, 4 and 5. Regarding the hypervolume indicator, there is no doubt that MOMBI-III showed a clear advantage over these single heuristics, except for the problems ZDT1,2,6, DTLZ2,4, WFG4,6,9, where our proposed method was outperformed by EWC, AASF and WPO. In spite of this, our MOMBI-III achieved the first and second places 16 times, out of 21. On the other hand, when evaluating with respect to the s -energy, our method outperformed the other variants in 20 cases. Only for the concave ZDT2, WN ranked first, and in this case only the extreme points were found. This suggests that MOMBI-III effectively minimized this measure. Similarly, when using the Solow-Polasky indicator

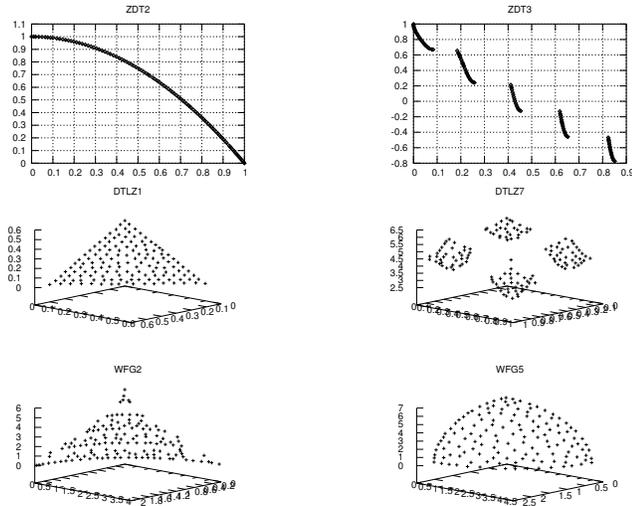


Figure 2: Pareto fronts produced by MOMBI-III

MOMBI-III outperformed the single heuristics, only in DTLZ1, CHE found a better value without significantly surpassing our algorithm. It is worth noticing that a zero value of the Solow-Polasky indicator means that it was not possible to calculate the inverse of the C matrix (see expression 5), since all the population concentrates on extreme points. In Figure 2, we show some approximations to the Pareto front produced by MOMBI-III, corresponding to the median hypervolume indicator.

4.3 State-of-the-art Algorithms

In this section, we compared MOMBI-III with respect to the well-known MOEA/D and NSGA-III. Additionally, we considered MOMBI-II, since our proposed method improves it. Experimental results are shown in Tables 6 and 7. In the case of the hypervolume indicator, we can observe again an overwhelming outperformance of our proposed method, whose scores were in the top places on 18 instances out of 21. It was only outperformed by MOEA/D and NSGA-III on the concave problems ZDT6 and DTLZ2. Moreover, MOMBI-III won over its predecessor MOMBI-II in 15 instances, providing solid evidence of its superiority. Concerning the s -energy, MOMBI-III performed better in 13 instances, being outperformed on ZDT2 by the other MOEAs; DTLZ1 by MOEA/D and NSGA-III; and WFG4,6-8 by MOEA/D. Finally, our proposed approach was better than MOMBI-II in almost all the test problems adopted.

4.4 Many-Objective Problems

In this experiment, we investigated the behavior of MOMBI-III in many-objective instances of DTLZ1. For this purpose, we performed a comparative study with the same algorithms of the previous experiment. The results of the hypervolume indicator are presented in Table 8. In this case, MOMBI-III outperformed MOEA/D and MOMBI-II in all instances and performed slightly better than NSGA-III. With this example, we intend to show the promising behavior of our proposal in many-objective problems. Nevertheless, further studies in this direction are still required.

5 CONCLUSIONS AND FUTURE WORK

One of the recent trends in multi-objective optimization is the development of generic methods, which can produce solutions of acceptable quality using a set of easy-to-implement low-level heuristics. These methods are known as hyper-heuristics and can be seen as high-level methodologies, which automatically produce an adequate combination of single heuristics for solving a broad set of problems. In this paper, we presented for the first time, a Hyper-Heuristic of Scalarizing Functions (MOMBI-III) for solving continuous multi-objective optimization problems, which are transformed into single-objective ones. The adopted set of scalarizing functions has several advantages, from which the most relevant are its low computational cost and compatibility with Pareto dominance. Although MOMBI-III can incorporate scalarizing functions that are incompatible with any form of Pareto optimality (e.g., PBI), an extra effort is required at each iteration to filter dominated solutions. Furthermore, MOMBI-III incorporates the s -energy for generating even distributions in objective space. Our experimental results showed that MOMBI-III significantly outperformed single heuristics as well as advanced algorithms, such as NSGA-III and MOEA/D in the majority of the instances of the ZDT, DTLZ and WFG test suites. Furthermore, our proposal showed promise in solving many-objective problems. As part of our future work, we would like to expand the set of heuristics, incorporating more values for the model parameters of the current scalarizing functions. Such values might be automatically adapted. We would also like to explore a probabilistic approach for the heuristic selection process, since this is currently performed in an exhaustive way.

ACKNOWLEDGMENTS

The first author acknowledges support from CINVSTAV-IPN and CONACyT to pursue graduate studies in Computer Science. The second author gratefully acknowledges support from CONACyT project no. 221551 and from a Cátedra Marcos Moshinsky in Mathematics.

REFERENCES

- [1] Edmund K Burke, Michel Gendreau, Matthew Hyde, Graham Kendall, Gabriela Ochoa, Ender Özcan, and Rong Qu. Hyper-heuristics: A Survey of the State of the Art. *Journal of the Operational Research Society* 64, 12 (2013), 1695–1724.
- [2] Carlos A. Coello Coello, Gary B. Lamont, and David A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems* (second ed.). Springer, New York. ISBN 978-0-387-33254-3.
- [3] I. Das and J. E. Dennis. A Closer Look at Drawbacks of Minimizing Weighted Sums of Objectives for Pareto set Generation in Multicriteria Optimization Problems. *Structural Optimization* 14, 1 (1997), 63–69.
- [4] Kalyanmoy Deb and Himanshu 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, 4 (Aug 2014), 577–601.
- [5] Kalyanmoy Deb, Lothar Thiele, Marco Laumanns, and Eckart Zitzler. Scalable Test Problems for Evolutionary Multiobjective Optimization. In *Evolutionary Multiobjective Optimization*, Ajith Abraham, Lakhmi Jain, and Robert Goldberg (Eds.). Springer London (2005), 105–145.
- [6] G. Eichfelder. An Adaptive Scalarization Method in Multiobjective Optimization. *SIAM Journal on Optimization* 19, 4 (jan 2009), 1694–1718.
- [7] Michael T.M. Emmerich, André H. Deutz, and Johannes W. Krusselbrink. On Quality Indicators for Black-Box Level Set Approximation. In *EVOLVE - A Bridge Between Probability, Set Oriented Numerics and Evolutionary Computation*, E. Tantar, A. A. Tantar, P. Bouvry, P. Moral, P. Legrand, C. A. Coello Coello, and O. Schütze (Eds.). Springer-Verlag. Studies in Computational Intelligence Vol. 447, Heidelberg, Germany, Chapter 4 (2013), 157–185. 978-3-642-32725-4.

Table 3: Median and standard deviation of the hypervolume indicator for single heuristics and MOMBI-III. The two best values are shown in gray scale, where a darker tone corresponds to the best value.

Problem	WS	EWC	WPO	WN	CHE	ASF	AASF	MOMBI-III
ZDT1	3.6539e+00 7.26e-4	3.6620e+00 8.08e-5	3.6333e+00 1.60e-2	3.0000e+00 1.78e-4	3.6614e+00 4.55e-5	3.6614e+00 4.55e-5	3.6614e+00 9.04e-5	3.6616e+00 7.40e-5
ZDT2	3.0000e+00 1.71e-7	3.3286e+00 1.02e-4	3.3276e+00 1.20e-3	3.0000e+00 4.02e-7	3.3281e+00 1.60e-4	3.3281e+00 1.60e-4	3.3281e+00 1.65e-4	3.3283e+00 1.51e-4
ZDT3	4.7070e+00 8.32e-2	4.8148e+00 5.20e-5	4.7813e+00 2.41e-1	4.0361e+00 2.52e-4	4.8140e+00 1.67e-4	4.8140e+00 1.67e-4	4.8141e+00 1.40e-4	4.8152e+00 4.91e-5
ZDT4	3.6521e+00 1.72e-3	3.6590e+00 1.26e-2	3.5424e+00 1.10e-1	2.9967e+00 2.40e-3	3.6584e+00 1.93e-3	3.6584e+00 1.93e-3	3.6580e+00 3.42e-3	3.6573e+00 2.47e-3
ZDT6	2.7741e+00 4.13e-4	3.0348e+00 1.25e-3	3.0347e+00 1.18e-3	2.7728e+00 1.14e-3	3.0312e+00 2.59e-3	3.0312e+00 2.59e-3	3.0309e+00 2.56e-3	3.0316e+00 2.20e-3
DTLZ1	7.8809e+00 1.42e-3	7.9637e+00 2.81e-4	7.8612e+00 7.14e-2	7.8807e+00 2.17e-4	7.9919e+00 2.21e-2	7.9917e+00 2.50e-2	7.9364e+00 2.32e-2	7.9745e+00 1.14e-4
DTLZ2	7.0000e+00 3.43e-7	7.3928e+00 7.36e-4	7.4174e+00 1.37e-3	7.0000e+00 2.97e-7	7.3919e+00 4.29e-3	7.4190e+00 1.73e-3	7.4245e+00 3.68e-4	7.4236e+00 9.48e-4
DTLZ3	6.2978e+01 1.30e+0	6.3350e+01 1.44e-2	6.2969e+01 1.78e+0	6.2979e+01 2.30e-2	6.3384e+01 1.05e-1	6.3385e+01 2.47e-2	6.3400e+01 1.52e-2	6.3415e+01 1.36e-2
DTLZ4	7.0000e+00 2.10e-1	7.3931e+00 3.01e-1	7.4186e+00 2.55e-1	7.0000e+00 3.28e-1	7.3983e+00 3.00e-1	7.4234e+00 3.49e-1	7.4252e+00 3.08e-1	7.4247e+00 1.84e-1
DTLZ5	5.6716e+00 9.91e-3	5.8123e+00 4.24e-2	6.0001e+00 2.11e-2	5.6716e+00 1.77e-2	5.9656e+00 3.00e-2	6.0410e+00 5.87e-3	6.0412e+00 1.03e-2	6.1021e+00 1.19e-3
DTLZ6	5.4390e+00 9.07e-2	5.7566e+00 4.66e-2	5.7683e+00 8.14e-2	5.3947e+00 9.99e-2	5.7426e+00 7.22e-2	5.8099e+00 7.57e-2	5.8163e+00 8.00e-2	5.8509e+00 9.41e-2
DTLZ7	1.5994e+01 2.18e-3	1.7252e+01 1.08e-1	1.6218e+01 6.09e-2	1.5868e+01 3.30e-3	1.7230e+01 9.60e-2	1.7475e+01 2.81e-2	1.7482e+01 3.77e-2	1.7545e+01 1.02e-2
WFG1	5.4395e+01 1.32e+0	4.8913e+01 1.58e+0	4.4832e+01 1.80e+0	4.5190e+01 5.64e+0	5.2472e+01 1.63e+0	5.2447e+01 1.63e+0	5.2128e+01 1.75e+0	5.4921e+01 1.67e+0
WFG2	9.9789e+01 3.57e-1	1.0035e+02 1.40e-1	9.7306e+01 1.64e-1	7.1243e+01 1.21e+0	9.9489e+01 2.85e-1	1.0031e+02 1.22e-1	1.0034e+02 1.36e-1	1.0082e+02 1.04e-1
WFG3	5.4610e+01 4.01e-1	7.3941e+01 5.60e-1	7.2792e+01 8.79e-2	5.6352e+01 3.26e-1	7.4927e+01 2.44e-1	7.5219e+01 1.92e-1	7.5138e+01 2.00e-1	7.5220e+01 1.54e-1
WFG4	5.6991e+01 1.34e-2	7.5551e+01 9.92e-2	7.6980e+01 9.20e-2	5.6995e+01 9.27e-3	7.5590e+01 1.70e-1	7.6737e+01 8.43e-2	7.6743e+01 7.67e-2	7.6613e+01 8.90e-2
WFG5	5.3487e+01 5.78e-6	7.2536e+01 9.41e-2	7.3541e+01 4.35e-2	5.3487e+01 4.34e-6	7.2351e+01 1.53e-1	7.3688e+01 4.58e-2	7.3728e+01 6.67e-2	7.3823e+01 5.24e-2
WFG6	5.4326e+01 3.83e-1	7.2957e+01 2.84e-1	7.4509e+01 3.06e-1	5.4374e+01 3.99e-1	7.2872e+01 3.65e-1	7.4153e+01 3.93e-1	7.4305e+01 3.60e-1	7.4245e+01 2.87e-1
WFG7	5.7011e+01 4.57e-3	7.5688e+01 4.23e-1	7.6816e+01 5.35e-2	5.7012e+01 3.36e-3	7.5529e+01 2.59e-1	7.6852e+01 7.72e-2	7.6900e+01 6.40e-2	7.7048e+01 4.94e-2
WFG8	5.2754e+01 7.05e-1	7.1874e+01 3.21e-1	7.2693e+01 3.06e-1	5.3782e+01 6.93e-1	7.1874e+01 1.88e-1	7.3041e+01 1.80e-1	7.2915e+01 1.71e-1	7.2842e+01 2.28e-1
WFG9	5.4922e+01 1.76e+0	7.3757e+01 2.56e-1	7.6333e+01 1.34e+0	5.5106e+01 3.39e+0	7.3891e+01 1.28e+0	7.5062e+01 1.03e+0	7.5058e+01 1.12e+0	7.5127e+01 2.18e+0

Table 4: Median and standard deviation of the s-energy measure for single heuristics and MOMBI-III.

Problem	WS	EWC	WPO	WN	CHE	ASF	AASF	MOMBI-III
ZDT1	1.366e+05 3.34e+04	5.741e+04 1.29e+02	1.275e+05 4.58e+03	4.472e+07 1.34e+08	6.486e+04 1.03e+01	6.486e+04 1.03e+01	6.486e+04 5.02e+01	5.639e+04 1.34e+02
ZDT2	1.684e+08 9.93e+08	5.702e+04 1.13e+04	6.501e+04 1.05e+02	1.414e+00 3.24e+07	5.606e+04 2.22e+02	5.606e+04 2.22e+02	5.605e+04 3.14e+03	5.609e+04 1.10e+02
ZDT3	1.931e+06 8.62e+05	6.505e+04 1.04e+04	1.203e+06 1.61e+05	5.737e+04 1.08e+08	4.938e+04 4.73e+03	4.938e+04 4.73e+03	4.884e+04 5.74e+03	4.404e+04 2.80e+02
ZDT4	1.176e+05 2.11e+04	5.764e+04 2.25e+03	1.415e+05 6.14e+04	8.762e+06 3.16e+08	6.486e+04 7.17e+02	6.486e+04 7.17e+02	6.485e+04 5.80e+02	5.629e+04 2.69e+02
ZDT6	2.930e+09 1.50e+16	7.231e+04 9.64e+03	8.987e+04 1.03e+04	9.295e+05 9.01e+10	7.012e+04 5.11e+02	7.012e+04 5.11e+02	7.007e+04 3.52e+03	6.936e+04 6.07e+02
DTLZ1	1.103e+41 4.69e+60	6.064e+06 8.42e+10	3.319e+13 2.92e+17	3.302e+44 6.39e+61	1.526e+06 4.75e+10	8.164e+05 3.23e+14	8.153e+05 1.11e+12	7.981e+05 8.30e+03
DTLZ2	2.792e+23 2.48e+35	1.212e+06 2.57e+06	1.248e+05 1.04e+02	9.271e+22 5.04e+43	1.771e+07 2.05e+11	1.238e+05 2.83e+02	1.206e+05 7.42e+02	1.182e+05 6.26e+02
DTLZ3	1.526e+12 5.09e+34	6.366e+05 7.94e+08	1.873e+14 3.76e+23	2.113e+13 1.77e+33	3.949e+05 6.15e+09	1.364e+05 2.88e+04	1.429e+05 7.87e+04	1.167e+05 2.13e+03
DTLZ4	3.255e+23 8.42e+35	1.077e+06 2.91e+24	1.248e+05 3.72e+15	4.021e+25 3.53e+37	3.067e+05 1.49e+16	1.222e+05 3.07e+16	1.203e+05 1.36e+13	1.176e+05 1.66e+06
DTLZ5	5.059e+26 3.88e+62	1.433e+20 3.94e+23	1.037e+18 1.14e+20	1.086e+27 6.09e+57	5.090e+16 8.80e+19	2.547e+18 1.11e+22	6.107e+19 1.58e+21	7.102e+06 1.55e+14
DTLZ6	2.721e+39 1.45e+44	3.706e+07 1.80e+08	2.703e+07 1.19e+09	3.910e+27 4.31e+33	3.406e+06 9.04e+14	3.433e+07 1.52e+13	2.982e+09 1.01e+22	8.292e+05 1.09e+14
DTLZ7	2.209e+14 3.57e+18	8.873e+05 3.36e+08	7.196e+06 1.02e+08	1.148e+16 7.92e+17	6.833e+05 1.45e+10	1.111e+06 3.45e+06	1.257e+06 3.84e+06	9.824e+04 4.83e+03
WFG1	1.451e+09 1.35e+11	1.355e+05 4.40e+06	1.582e+07 8.35e+10	8.852e+09 6.25e+16	1.945e+05 3.17e+05	8.232e+04 1.99e+05	7.936e+04 1.58e+06	4.654e+04 6.16e+03
WFG2	4.249e+07 8.85e+09	6.491e+04 3.35e+05	2.765e+06 1.25e+07	2.854e+11 9.46e+12	1.191e+05 1.24e+07	2.845e+04 8.39e+05	2.934e+04 8.69e+04	2.064e+04 1.38e+03
WFG3	3.107e+13 2.45e+18	1.800e+06 3.20e+09	2.885e+06 5.79e+09	3.518e+13 8.16e+15	1.830e+05 5.31e+06	9.107e+05 6.35e+06	7.468e+05 8.65e+06	3.911e+04 2.52e+03
WFG4	1.312e+10 3.10e+11	1.500e+04 7.94e+04	9.368e+03 1.18e+02	4.125e+09 2.20e+13	3.636e+05 2.38e+06	9.360e+03 9.61e+02	9.438e+03 2.51e+03	8.355e+03 1.14e+02
WFG5	9.993e+16 2.50e+19	1.926e+04 3.78e+04	9.202e+03 3.00e+01	2.037e+16 1.47e+21	9.692e+06 4.80e+07	9.111e+03 1.13e+03	9.047e+03 1.32e+02	8.419e+03 9.45e+01
WFG6	2.161e+10 1.29e+12	1.562e+04 3.91e+05	9.186e+03 6.60e+01	2.731e+10 1.02e+13	2.838e+05 3.40e+07	9.216e+03 9.11e+02	9.245e+03 9.29e+02	8.354e+03 1.01e+02
WFG7	1.235e+11 2.32e+12	2.383e+04 1.06e+06	9.439e+03 4.35e+01	3.176e+11 8.53e+13	5.113e+05 9.90e+06	9.324e+03 6.27e+02	9.301e+03 2.57e+02	8.415e+03 1.17e+02
WFG8	3.879e+09 6.28e+12	1.928e+04 1.58e+05	2.482e+06 8.80e+06	9.632e+08 1.51e+11	3.408e+05 9.34e+05	3.435e+04 2.19e+05	5.806e+04 2.13e+05	8.314e+03 1.23e+02
WFG9	6.356e+08 5.32e+13	2.246e+04 4.22e+04	9.663e+03 2.72e+02	3.228e+09 8.34e+12	1.047e+05 1.43e+07	1.371e+04 5.40e+04	1.196e+04 6.04e+04	8.565e+03 9.60e+01

Table 5: Median and standard deviation of the Solow-Polanyi indicator for single heuristics and MOMBI-III.

Problem	WS	EWC	WPO	WN	CHE	ASF	AASF	MOMBI-III
ZDT1	8.24e+00 1.55e-2	8.36e+00 2.47e-2	7.35e+00 1.30e-1	2.00e+00 1.47e-2	8.33e+00 3.78e-3	8.33e+00 3.78e-3	8.33e+00 5.34e-3	8.37e+00 1.09e-3
ZDT2	2.00e+00 2.85e-3	8.37e+00 7.66e-3	8.35e+00 6.87e-3	2.00e+00 5.22e-3	8.37e+00 8.79e-4	8.37e+00 8.79e-4	8.37e+00 3.16e-3	8.37e+00 7.20e-4
ZDT3	5.53e+00 6.97e-1	1.12e+01 2.47e-2	6.82e+00 4.89e-1	2.00e+00 1.49e-3	1.15e+01 1.26e-2	1.15e+01 1.26e-2	1.15e+01 1.11e-2	1.15e+01 4.33e-3
ZDT4	8.24e+00 2.36e-2	8.37e+00 1.36e-1	6.69e+00 7.26e-1	2.00e+00 9.88e-2	8.33e+00 2.56e-2	8.33e+00 2.56e-2	8.33e+00 5.72e-3	8.37e+00 9.37e-3
ZDT6	2.00e+00 8.45e-3	7.16e+00 9.58e-2	7.08e+00 8.49e-2	2.01e+00 1.33e-2	7.17e+00 9.85e-2	7.17e+00 9.85e-2	7.14e+00 1.65e-1	7.18e+00 1.34e-1
DTLZ1	0.00e+00 3.83e+0	8.34e+00 2.79e-2	3.05e+00 2.35e+0	0.00e+00 9.61e-1	9.44e+00 1.86e+0	8.98e+00 2.33e+0	8.99e+00 1.97e+0	9.24e+00 4.25e-1
DTLZ2	0.00e+00 0.00e+0	2.45e+01 1.15e-1	3.35e+01 1.19e-2	0.00e+00 0.00e+0	2.99e+01 4.05e-1	3.37e+01 3.18e-2	3.41e+01 4.38e-2	3.44e+01 9.22e-2
DTLZ3	0.00e+00 1.54e+0	2.56e+01 3.39e-1	3.00e+00 1.28e+1	0.00e+00 1.07e+0	3.24e+01 6.53e+0	3.39e+01 2.83e-1	3.40e+01 4.05e-1	3.47e+01 6.42e-1
DTLZ4	0.00e+00 0.00e+0	2.45e+01 7.48e+0	3.35e+01 6.28e+0	0.00e+00 0.00e+0	3.05e+01 7.70e+0	3.40e+01 9.57e+0	3.42e+01 7.77e+0	3.45e+01 4.70e+0
DTLZ5	0.00e+00 1.37e+0	7.88e+00 1.41e-1	8.29e+00 3.06e-1	2.23e+00 1.42e+0	8.35e+00 2.04e-1	8.48e+00 1.02e-1	8.49e+00 1.32e-1	8.98e+00 4.33e-1
DTLZ6	0.00e+00 0.00e+0	1.17e+01 1.03e+0	1.17e+01 1.29e+0	4.33e+00 1.15e+0	1.38e+01 1.35e+0	1.21e+01 1.23e+0	1.20e+01 1.30e+0	1.48e+01 2.00e+0
DTLZ7	5.97e+00 4.23e-1	3.98e+01 1.51e+0	1.87e+01 8.12e-1	4.84e+00 9.68e-1	3.29e+01 8.95e-1	3.77e+01 8.64e-1	3.77e+01 9.35e-1	4.53e+01 1.16e+0
WFG1	2.15e+01 1.93e+0	3.74e+01 3.85e+0	1.99e+01 2.71e+0	1.27e+01 3.13e+0	5.06e+01 1.87e+0	6.04e+01 3.30e+0	6.07e+01 2.99e+0	6.80e+01 4.92e+0
WFG2	3.55e+01 1.33e+0	7.81e+01 2.59e+0	2.70e+01 3.63e-1	7.19e+00 5.35e-1	6.95e+01 2.59e+0	9.41e+01 1.83e+0	9.46e+01 1.62e+0	1.00e+02 2.50e+0
WFG3	4.81e+00 2.70e-1	5.53e+01 2.24e+0	3.09e+01 4.90e-1	4.77e+00 2.83e-2	5.20e+01 1.56e+0	5.18e+01 1.45e+0	5.11e+01 1.53e+0	7.51e+01 1.17e+0
WFG4	3.47e+00 2.79e-2	1.15e+02 8.66e-1	1.22e+0					

Table 6: Median and standard deviation of the hypervolume indicator for the compared MOEAs and MOMBI-III.

Problem	MOEA/D	NSGA-III	MOMBI-II	MOMBI-III
ZDT1	3.660e+00 1.58e-3 ↑	3.661e+00 1.78e-4 ↑	3.661e+00 8.04e-5 ↑	3.662e+00 7.59e-5
ZDT2	3.326e+00 1.28e-3 ↑	3.328e+00 2.27e-4 ↑	3.328e+00 1.22e-4 ↑	3.328e+00 1.14e-4
ZDT3	4.811e+00 2.95e-3 ↑	4.813e+00 3.36e-4 ↑	4.814e+00 8.81e-5 ↑	4.815e+00 6.16e-2
ZDT4	3.649e+00 5.11e-3 ↑	3.658e+00 7.45e-3	3.658e+00 4.11e-3	3.659e+00 1.64e-3
ZDT6	3.036e+00 1.26e-3 ↓	3.024e+00 4.58e-3 ↑	3.031e+00 2.28e-3	3.030e+00 2.55e-3
DTLZ1	7.975e+00 1.63e-4	7.975e+00 5.56e-4	7.937e+00 4.78e-3 ↑	7.975e+00 5.54e-5
DTLZ2	7.426e+00 2.34e-5 ↓	7.425e+00 4.06e-4	7.376e+00 7.14e-3 ↑	7.423e+00 9.85e-4
DTLZ3	6.339e+01 1.87e-2 ↑	6.340e+01 2.83e-2 ↑	6.336e+01 1.81e-2 ↑	6.341e+01 6.25e-3
DTLZ4	7.426e+00 1.05e+0	7.425e+00 4.37e-1	7.407e+00 4.91e-3 ↑	7.424e+00 1.84e-1
DTLZ5	6.050e+00 2.15e-4 ↑	5.954e+00 2.18e-1 ↑	6.015e+00 3.26e-3 ↑	6.103e+00 1.09e-4
DTLZ6	5.821e+00 7.95e-2	5.444e+00 1.15e-1 ↑	5.748e+00 6.70e-2 ↑	5.877e+00 7.63e-2
DTLZ7	9.729e+00 2.61e-2 ↑	1.739e+01 2.88e-2 ↑	1.736e+01 1.15e-2 ↑	1.754e+01 1.24e-2
WFG1	5.305e+01 1.45e+0	4.907e+01 1.59e+0 ↑	5.443e+01 1.79e+0	5.492e+01 1.67e+0
WFG2	9.666e+01 1.16e+0 ↑	1.003e+02 1.80e-1 ↑	1.001e+02 1.61e-1 ↑	1.008e+02 1.04e-1
WFG3	7.283e+01 6.74e-1 ↑	7.408e+01 1.53e-1 ↑	7.505e+01 1.47e-1 ↑	7.522e+01 1.54e-1
WFG4	7.382e+01 4.23e-1 ↑	7.656e+01 1.04e-1	7.668e+01 9.41e-2	7.661e+01 8.90e-2
WFG5	7.134e+01 5.35e-1 ↑	7.373e+01 8.87e-2 ↑	7.353e+01 8.09e-2 ↑	7.383e+01 4.10e-2
WFG6	7.153e+01 6.24e-1 ↑	7.412e+01 2.69e-1	7.401e+01 3.76e-1	7.422e+01 3.13e-1
WFG7	7.308e+01 8.38e-1 ↑	7.685e+01 7.76e-2 ↑	7.682e+01 8.26e-2 ↑	7.700e+01 5.64e-2
WFG8	6.945e+01 9.24e-1 ↑	7.285e+01 2.67e-1	7.266e+01 2.02e-1 ↑	7.293e+01 2.29e-1
WFG9	6.821e+01 1.79e+0 ↑	7.392e+01 9.19e-1 ↑	7.489e+01 1.10e+0	7.513e+01 2.70e-1

Table 7: Median and standard deviation of the s-energy measure for the compared MOEAs and MOMBI-III.

Problem	MOEA/D	NSGA-III	MOMBI-II	MOMBI-III
ZDT1	6.48e+04 5.26e+01 ↑	6.49e+04 1.45e+02 ↑	6.49e+04 1.43e+02 ↑	5.69e+04 1.88e+02
ZDT2	5.61e+04 1.85e+02	5.61e+04 3.70e+03 ↓	5.61e+04 3.45e+02 ↓	5.61e+04 1.02e+02
ZDT3	1.69e+06 5.51e+07 ↑	1.32e+05 5.69e+04 ↑	2.08e+07 1.28e+07 ↑	4.40e+04 2.14e+03
ZDT4	6.45e+04 2.29e+02 ↑	6.48e+04 5.55e+03 ↑	6.49e+04 4.81e+03 ↑	5.65e+04 3.04e+02
ZDT6	7.07e+04 2.30e+02 ↑	7.00e+04 1.13e+04 ↑	7.07e+04 7.50e+05 ↑	6.94e+04 4.38e+02
DTLZ1	7.82e+05 2.25e+03 ↓	7.84e+05 1.12e+08 ↓	8.15e+05 1.43e+03 ↑	8.01e+05 4.57e+03
DTLZ2	1.19e+05 5.54e+00 ↑	1.19e+05 2.14e+02 ↑	1.25e+05 3.36e+02 ↑	1.18e+05 5.05e+02
DTLZ3	1.16e+05 3.65e+07	1.36e+05 9.25e+06 ↑	1.39e+05 3.62e+10 ↑	1.16e+05 1.51e+03
DTLZ4	1.19e+05 - ↑	1.19e+05 9.79e+12 ↑	1.24e+05 7.82e+02 ↑	1.18e+05 1.68e+06
DTLZ5	6.38e+15 8.00e+15 ↑	8.67e+14 1.89e+37 ↑	6.45e+16 1.41e+18 ↑	3.04e+06 1.52e+05
DTLZ6	1.21e+09 5.23e+10 ↑	3.28e+11 1.64e+21 ↑	4.89e+13 4.33e+15 ↑	4.49e+05 1.78e+05
DTLZ7	3.07e+10 1.73e+18 ↑	1.73e+06 1.01e+07 ↑	5.54e+11 1.30e+12 ↑	5.08e+04 2.22e+03
WFG1	8.87e+09 7.33e+11 ↑	1.57e+05 5.71e+06 ↑	3.48e+08 6.75e+10 ↑	4.65e+04 6.16e+03
WFG2	2.50e+04 1.21e+08 ↑	2.00e+04 8.16e+04	9.64e+08 3.12e+11 ↑	2.06e+04 1.38e+03
WFG3	2.20e+09 1.26e+10 ↑	2.66e+06 2.47e+08 ↑	8.20e+08 1.89e+11 ↑	3.91e+04 2.52e+03
WFG4	8.18e+03 1.28e+02	8.93e+03 4.38e+01 ↑	9.31e+03 1.00e+05 ↑	8.35e+03 1.14e+02
WFG5	1.66e+05 8.21e+07 ↑	8.85e+03 2.70e+01 ↑	9.16e+03 3.99e+07 ↑	8.40e+03 1.25e+02
WFG6	7.66e+03 1.41e+02 ↓	8.94e+03 7.02e+04 ↑	1.01e+04 6.12e+05 ↑	8.40e+03 1.29e+02
WFG7	7.67e+03 1.40e+02 ↓	8.86e+03 1.20e+03 ↑	9.17e+03 1.64e+06 ↑	8.45e+03 1.18e+02
WFG8	8.03e+03 2.16e+02 ↓	2.18e+04 5.33e+05 ↑	3.89e+08 2.05e+10 ↑	8.29e+03 1.42e+02
WFG9	2.93e+05 3.65e+15 ↑	1.08e+04 3.91e+04 ↑	5.02e+05 1.12e+07 ↑	4.31e+03 4.83e+01

Table 8: Median and standard deviation of the hypervolume indicator on many-objective instances of DTLZ1.

m	MOEA/D (1)	NSGA-III (2)	MOMBI-II (3)	MOMBI-III (4)
4	1.5995e+01 1.23e-4 ↑	1.5995e+01 8.66e-5	1.5945e+01 8.54e-3 ↑	1.5995e+01 4.35e-5
5	3.1999e+01 9.18e-5 ↑	3.1999e+01 6.97e-5 ↑	3.1930e+01 1.68e-2 ↑	3.1999e+01 3.65e-5
6	6.3998e+01 4.27e-4 ↑	6.4000e+01 9.66e-5 ↑	6.3922e+01 2.08e-2 ↑	6.4000e+01 7.23e-5
7	1.2800e+02 8.25e-4 ↑	1.2800e+02 1.36e-2	1.2784e+02 5.71e-2 ↑	1.2800e+02 1.17e-4
8	2.5599e+02 4.47e-3 ↑	2.5600e+02 2.00e-2	2.5577e+02 8.23e-2 ↑	2.5600e+02 8.56e-4
9	5.1196e+02 2.69e-2 ↑	5.1199e+02 7.76e-2 ↑	5.1164e+02 1.78e-1 ↑	5.1200e+02 1.36e-3
10	1.0238e+03 7.78e-2 ↑	1.0240e+03 3.41e-2 ↑	1.0232e+03 2.63e-1 ↑	1.0240e+03 8.59e-4

[8] Ioannis Giagkiozis, Robin C. Purshouse, and Peter 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.), Springer, Lecture Notes in Computer Science Vol. 7811, 428–442.

[9] Richard A. Gonçalves, Josiel N. Kuk, Carolina P. Almeida, and Sandra M. Venske. MOEA/D-HH: A Hyper-Heuristic for Multi-objective Problems. In *Evolutionary Multi-Criterion Optimization, 8th International Conference, EMO 2015*, A. Gaspar-Cunha, C. H. Antunes, and C. Coello Coello (Eds.), Springer, Lecture Notes in Computer Science Vol. 9018, Guimarães, Portugal, 94–108.

[10] D.P. Hardin and E.B. Saff. Discretizing Manifolds via Minimum Energy Points. *Notices of the American Mathematical Society* 51, 10 (November 2004), 1186–1194.

[11] Raquel Hernández Gómez and Carlos A. Coello Coello. Improved Metaheuristic Based on the R2 Indicator for Many-Objective Optimization. In *2015 Genetic and Evolutionary Computation Conference (GECCO 2015)*. ACM Press, Madrid, Spain, 679–686. ISBN 978-1-4503-3472-3.

[12] 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* 10, 5 (Oct 2006), 477–506.

[13] Evan J. Hughes. Multiple Single Objective Pareto Sampling. In *Proceedings of the 2003 Congress on Evolutionary Computation (CEC'2003)*, Vol. 4. IEEE Press, Canberra, Australia, 2678–2684.

[14] Hisao Ishibuchi, Naoya Akedo, and Yusuke Nojima. Behavior of Multiobjective Evolutionary Algorithms on Many-Objective Knapsack Problems. *IEEE Transactions on Evolutionary Computation* 19, 2 (April 2015), 264–283.

[15] Hisao Ishibuchi and Tadahiko Murata. 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, 3 (August 1998), 392–403.

[16] Hisao Ishibuchi, Yuji Sakane, Noritaka Tsukamoto, and Yusuke 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*, M. Ehrgott, C. M. Fonseca, X. Gandibleux, J.-K. Hao, and M. Sevaux (Eds.), Springer, Lecture Notes in Computer Science Vol. 5467, Nantes, France, 438–452.

[17] Hisao Ishibuchi, Yuji Sakane, Noritaka Tsukamoto, and Yusuke Nojima. Simultaneous Use of Different Scalarizing Functions in MOEA/D. In *Proceedings of the 12th annual conference on Genetic and Evolutionary Computation (GECCO'2010)*. ACM Press, Portland, Oregon, USA, 519–526. ISBN 978-1-4503-0072-8.

[18] Joshua Knowles, David Corne, and Kalyanmoy Deb (Eds.). *Multiobjective Problem Solving from Nature. From Concepts to Applications*. Springer (2008), Berlin. ISBN 978-3-540-72963-1.

[19] Bingdong Li, Jinlong Li, Ke Tang, and Xin Yao. Many-Objective Evolutionary Algorithms: A Survey. *Comput. Surveys* 48, 1 (September 2015).

[20] Achille Messac, Cyrille P. Puemi-Sukam, and Emanuel Melachrinoudis. Aggregate Objective Functions and Pareto Frontiers: Required Relationships and Practical Implications. *Optimization and Engineering* 1, 2 (2000), 171–188.

[21] Miriam Pescador-Rojas, Raquel Hernández Gómez, Elizabeth Montero, Nicolás Rojas-Morales, María-Cristina Riff, and Carlos 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. Klarmoth, O. Schütze, M. Wiecek, Y. Jin, and C. Grimme (Eds.), Springer International Publishing, Cham, 499–513.

[22] K.C. Tan, E.F. Khor, and T.H. Lee. *Multiobjective Evolutionary Algorithms and Applications*. Springer-Verlag (2005), London. ISBN 1-85233-836-9.

[23] Christian von Lüken, Benjamin Baran, and Carlos Brizuela. A Survey on Multiobjective Evolutionary Algorithms for Many-objective Problems. *Computational Optimization and Applications* 58, 3 (July 2014), 707–756.

[24] David J. Walker and Ed Keedwell. Multi-objective Optimisation with a Sequence-based Selection Hyper-heuristic. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion (GECCO '16 Companion)*. ACM, New York, NY, USA, 81–82.

[25] David J. Walker and Ed Keedwell. Towards Many-Objective Optimisation with Hyper-Heuristics: Identifying Good Heuristics with Indicators. In *Parallel Problem Solving from Nature – PPSN XIV, 14th International Conference*, J. Handl, E. Hart, P. R. Lewis, M. López-Ibáñez, G. Ochoa, and B. Paechter (Eds.), Springer, Lecture Notes in Computer Science Vol. 9921 (2016), 493–502. ISBN 978-3-319-45822-9.

[26] R. Wang, Q. Zhang, and T. Zhang. Decomposition-Based Algorithms Using Pareto Adaptive Scalarizing Methods. *IEEE Transactions on Evolutionary Computation* 20, 6 (Dec 2016), 821–837.

[27] Shengxiang Yang, Shouyong Jiang, and Yong Jiang. Improving the Multiobjective Evolutionary Algorithm based on Decomposition with New Penalty Schemes. *Soft Computing* (2016), 1–15.

[28] Qingfu Zhang and Hui Li. MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition. *IEEE Transactions on Evolutionary Computation* 11, 6 (December 2007), 712–731.

[29] Eckart Zitzler. *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*. Ph.D. Dissertation. Swiss Federal Institute of Technology (ETH), Zurich, Switzerland (1999).

[30] Eckart Zitzler, Kalyanmoy Deb, and Lothar Thiele. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evol. Comput.* 8, 2 (Jun 2000), 173–195.