Exploiting the Trade-off between Convergence and Diversity Indicators

Jesús Guillermo Falcón-Cardona
Computer Science Department
CINVESTAV-IPN
Mexico City, Mexico
jfalcon@computacion.cs.cinvestav.mx

Hisao Ishibuchi
Department of Computer Science and Engineering
Southern University of Science and Technology
Shenzhen, China
hisao@sustech.edu.cn

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

***Abstract*—Recently, it has been stressed that multi-objective evolutionary algorithms (MOEAs) should produce Pareto front approximations with good diversity regardless of the Pareto front geometry. In this light, the use of selection mechanisms based on multiple quality indicators (QIs) is a promising approach due to the exploitation of their strengths. In this paper, we propose to exploit the trade-off between the IGD$^+$ and the Riesz s-energy indicators, which assess convergence and diversity of a Pareto front approximation, respectively. Since the preferences of both indicators are regularly in conflict due to their different measure scope, it is possible to design a selection mechanism that exploits such trade-off, aiming to generate Pareto front approximations with a good degree of convergence and diversity simultaneously. Our proposed density estimator is embedded in a steady-state MOEA, denoted as PFI-EMOA, which is compared with several state-of-the-art MOEAs. Our experimental results based on the WFG and WFG^{-1} test problems show that PFI-EMOA outperforms several state-of-the-art MOEAs, providing outcomes having good convergence and diversity. Additionally, the performance of PFI-EMOA does not depend on the Pareto front shape.**

***Index Terms*—Multi-objective optimization, combined quality indicator, selection mechanism**

I. Introduction

For almost 30 years, multi-objective evolutionary algorithms (MOEAs) have been successfully applied to solve nonlinear optimization problems involving two or more objective functions, i.e., the so-called multi-objective optimization problems (MOPs) [1]. Currently, there is a plethora of MOEAs that employ different mechanisms to approximate the Pareto front of an MOP. To compare the performance of MOEAs, they are tested on several benchmark problems, where each one has specific properties such as search difficulties (e.g., multi-modality, separability, bias) and Pareto front shape.[1] MOEAs have been found to be very competitive when using traditional benchmark problems [2]. However, Ishibuchi *et al.* [3] pointed out that the performance of some MOEAs[2] radically changes even if the commonly used test problems are slightly modified. In consequence, an important issue is that the performance of some MOEAs strongly depends on the Pareto front shape of the MOP being solved.

Indicator-based MOEAs (IB-MOEAs) have been recently used to overcome the performance dependence of MOEAs to the Pareto front shapes [4], [5], [6]. IB-MOEAs are based on quality indicators (QIs) which are set functions that evaluate the quality of Pareto front approximations according to specific convergence and diversity preferences [7]. Hence, a straightforward idea is to employ one or more QIs to design the selection mechanisms of an MOEA. In consequence, the Pareto front approximations generated by an IB-MOEA inherit the properties of its baseline QI(s). Since each indicator-based mechanism has specific search abilities, an IB-MOEA using multiple QIs could explore different regions of the search space, avoiding the performance dependence to the Pareto front shape [6], [8].

In 2019, Falcón-Cardona *et al.* [6] proposed the CRI-EMOA algorithm which combines the individual selection effect of two density estimators based on the Riesz s-energy (E_s) [9] and the Inverted Generational Distance plus (IGD$^+$) [10] indicators. The main idea of CRI-EMOA is to statistically analyze its convergence behavior (based on an approximation to the hypervolume indicator [11]) to decide which indicator-based density estimator (IB-DE) should be executed. If the convergence stagnates, the diversity of solutions is promoted through the E_s-DE. On the other hand, if the convergence either decreases or increases, the IGD$^+$-DE is applied. The authors empirically showed that the performance of CRI-EMOA is invariant to the Pareto front shape. However, its main drawback concerns its statistical analysis of global convergence since the approximation to the hypervolume indicator (HV) provides noisy results which could lead to not using the correct IB-DE.

Currently, there is a wide range of QIs that aim to mainly assess convergence and diversity of Pareto front approximations [7]. Since the scope of convergence and diversity QIs is

The first author acknowledges support from CONACyT and CINVESTAV-IPN to pursue graduate studies in Computer Science. The third author gratefully acknowledges support from CONACyT grant no. 2016-01-1920 (*Investigación en Fronteras de la Ciencia 2016*) and from a SEP-Cinvestav grant (application no. 4).

[1]This is the geometry related to the Pareto front. For instance, it could be linear, convex, concave, mixed, degenerate, or disconnected.

[2]For instance, MOEAs using a set of convex weight vectors to guide the population towards the Pareto front since the Pareto front shape of some MOPs is highly correlated with the simplex formed by the weight vectors.

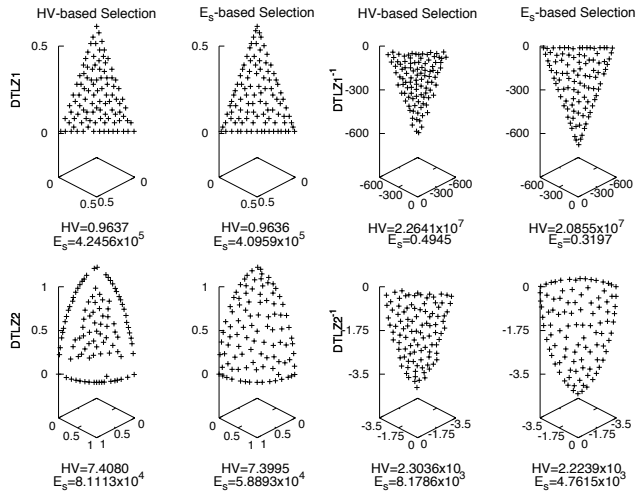


Fig. 1: Trade-off between the hypervolume indicator and the Riesz s-energy. In all cases, a well-diversified Pareto front does not have the best hypervolume value and viceversa.

different, it is clear that their preferences could be in conflict. In Figure 1, we show the Pareto fronts related to some DTLZ [12] and DTLZ^{-1} [3] benchmark problems. For each MOP, we generated two subsets by iteratively deleting the solution with worst contribution to HV and E_s and, then, we measured both indicator values. The QI values show that a high hypervolume value is not strictly related to a Pareto front approximation with high diversity. Hence, there is a trade-off between both QIs that can be exploited by an MOEA to produce Pareto front approximations, optimizing both a convergence indicator and a diversity one.

In this paper, we propose to exploit the trade-off between the IGD$^+$[3] and the Riesz s-energy to overcome the drawbacks of CRI-EMOA. To this aim, we mathematically combine IGD$^+$ and E_s in a single indicator which is then embedded into a density estimator. Our proposed approach, called Pareto front shape invariant evolutionary multi-objective algorithm (PFI-EMOA), is based on the framework of SMS-EMOA [14] but it replaces the HV-DE by our proposed density estimator using the combined indicator.

The remainder of this paper is organized as follows. Section II introduces some mathematical definitions. Section III briefly describes the previous related work. Section IV is devoted to explain the design of PFI-EMOA. Section V presents the experimental results where we compare it with state-of-the-art MOEAs. Finally, our main conclusions and future work are highligthed in Section VI.

II. Background

Throughout this paper, we consider, without loss of generality, MOPs for minimization which are defined as follows:

$$\min_{\vec{x}\in\Omega}\left\{\vec{F}(\vec{x}) = (f_1(\vec{x}),\ldots,f_m(\vec{x}))\right\}, \quad (1)$$

where $\vec{x}$ is the vector of decision variables, $\Omega \subseteq \mathbb{R}^n$ is the decision space and $\vec{F}(\vec{x})$ is the vector of $m \geq 2$ objective functions such that $f_i : \Omega \rightarrow \mathbb{R}$ for $i \in \{1,2,\ldots,m\}$. Unlike single-objective optimization problems which have a single global optimal solution, the solution of an MOP is a set of solutions that represents the best possible trade-offs among the objective functions. To identify the trade-offs among the objective functions, the Pareto dominance relation has been commonly used. Given $\vec{x},\vec{y} \in \mathbb{R}^n$, $\vec{F}(\vec{x})$ Pareto dominates $\vec{F}(\vec{y})$ (denoted as $\vec{F}(\vec{x}) \prec \vec{F}(\vec{y})$) if and only if $f_i(\vec{x}) \leq f_i(\vec{y})\ \forall i = 1,\ldots,m$ and there exists at least one index $j \in \{1,\ldots,m\}$ such that $f_j(\vec{x}) < f_j(\vec{y})$. In case that $f_i(\vec{x}) \leq f_i(\vec{y})$ for all $i = 1,\ldots,m$, we say that $\vec{x}$ weakly Pareto dominates $\vec{y}$, denoted as $\vec{F}(\vec{x}) \preceq \vec{F}(\vec{y})$. Based on the Pareto dominance relation, it is possible to identify *Pareto optimal solutions*. A solution $\vec{x}^* \in \Omega$ is Pareto optimal if there does not exist another $\vec{x} \in \Omega$ such that $\vec{F}(\vec{x}) \prec \vec{F}(\vec{x}^*)$. The set of all Pareto optimal solutions is called the Pareto set and its image in objective space is known as Pareto front.

At each execution, an MOEA produces a Pareto front approximation (also known as *approximation set*). Given Ψ which is the set of all approximation sets, an approximation set $\mathcal{A} \in \Psi$ is a finite set of m-dimensional objective vectors. $\mathcal{A}$ is called a Pareto front approximation if none of its elements weakly dominates another element. Based on the approximation sets, we can easily extend the Pareto dominance relation to sets [15]. Given two approximation sets $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{A}$ dominates $\mathcal{B}$ (denoted as $\mathcal{A} \lhd \mathcal{B}$) if and only if $\forall \vec{b} \in \mathcal{B}, \exists \vec{a} \in \mathcal{A} : \vec{a} \preceq \vec{b}$ and $\mathcal{A} \neq \mathcal{B}$.

Quality indicators have been widely employed to quantitatively compare MOEAs' outcomes. Mathematically, an indicator is defined as follows:

Definition 1 (Quality indicator): A unary quality indicator $I : \Psi \rightarrow \mathbb{R}$ is a function that assigns a real value to every element $\mathcal{A} \in \Psi$.

Regarding QIs that measure convergence towards the Pareto front, an important property is Pareto compliance. This property determines if the total order imposed by an indicator is compliant with the partial order defined by the Pareto dominance relation. In the following, we provide the formal definition of Pareto compliance, as well as that of weak Pareto compliance.

Definition 2 (Pareto compliance): A quality indicator is Pareto-compliant if and only if $\mathcal{A} \lhd \mathcal{B} \Rightarrow I(\mathcal{A}) < I(\mathcal{B})$, assuming that a lower indicator value is better.

Definition 3 (Weak Pareto compliance): A quality indicator is weakly Pareto-compliant if and only if $\mathcal{A} \lhd \mathcal{B} \Rightarrow I(\mathcal{A}) \leq I(\mathcal{B})$, assuming that a lower indicator value is better.

III. Previous Related Work

In 2015, Ishibuchi *et al.* [10] proposed the IGD$^+$ indicator which measures convergence and diversity simultaneously. This QI measures the average distance between a reference set

[3]The decision to use IGD$^+$ instead of HV is because the latter is computationally expensive when dealing with MOPs having more than three objective functions. Additionally, the preferences of IGD$^+$ and HV are highly correlated [13].

$\mathcal{Z}$ and a Pareto front approximation $\mathcal{A}$. For this purpose, IGD$^+$ uses a modified Euclidean distance that works as follows for given $\vec{z} \in \mathcal{Z}$ and $\vec{a} \in \mathcal{A}$. If $\vec{z}$ Pareto dominates $\vec{a}$, the usual Euclidean distance is calculated. In case that both vectors are mutually nondominated, the distance from $\vec{z}$ to the dominated region of $\vec{a}$ is computed. Due to the modified Euclidean distance, IGD$^+$ is a weakly Pareto-compliant indicator. IGD$^+$ is formulated in the following equation.

$$\text{IGD}^+(\mathcal{A}, \mathcal{Z}) = \frac{1}{|\mathcal{Z}|} \sum_{\vec{z} \in \mathcal{Z}} \min_{\vec{a} \in \mathcal{A}} d^+(\vec{a}, \vec{z}) \tag{2}$$

where $d^+(\vec{a}, \vec{z}) = \sqrt{\sum_{i=1}^{m} \max(a_i - z_i, 0)^2}$ is a modified Euclidean distance and m is the dimension of the objective space. The closer IGD$^+(\mathcal{A}, \mathcal{Z})$ is to zero, the more similar $\mathcal{A}$ and $\mathcal{Z}$ are. Hence, the aim is to minimize the IGD$^+$ value. One possibility to employ IGD$^+$ in a selection mechanism is to compute the individual contributions of the solutions in $\mathcal{A}$ and delete the one having the minimum contribution. Mathematically, this IGD$^+$-DE is given as follows: $\vec{a}_{\text{worst}} = \arg\min_{\vec{a} \in \mathcal{A}} C_{\text{IGD}^+}(\vec{a}, \mathcal{A})$, where $C_{\text{IGD}^+}(\vec{a}, \mathcal{A}) = |\text{IGD}^+(\mathcal{A}, \mathcal{Z}) - \text{IGD}^+(\mathcal{A} \setminus \{\vec{a}\}, \mathcal{Z})|$ is the IGD$^+$ contribution of a solution $\vec{a} \in \mathcal{A}$.

A remarkable diversity indicator, which has not been extensively employed in the evolutionary multi-objective optimization community, is the Riesz s-energy (E_s) proposed by Hardin and Saff [9]. This indicator measures the uniformity of a set of points in m-dimensional manifolds. It is mathematically defined as follows:

$$E_s(\mathcal{A}) = \sum_{i \neq j} \|\vec{a}_i - \vec{a}_j\|^{-s} \tag{3}$$

where $\|\cdot\|$ represents the Euclidean distance and $s > 0$ is a parameter that controls the degree of uniformity of the solutions in $\mathcal{A}$. It is worth noting that as $s \to \infty$, with $|\mathcal{A}|$ fixed, E_s leads to the best packing problem [9]. The aim is to minimize the Riesz s-energy, which implies a Pareto front approximation with high diversity. Regarding the individual contribution of a solution $\vec{a} \in \mathcal{A}$, it is given as $C_{E_s}(\vec{a}, \mathcal{A}) = \frac{1}{2}[E_s(\mathcal{A}) - E_s(\mathcal{A} \setminus \{\vec{a}\})]$. On the basis of C_{E_s}, the E_s-DE is defined as follows: $\vec{a}_{\text{worst}} = \arg\max_{\vec{a} \in \mathcal{A}} C_{E_s}(\vec{a}, \mathcal{A})$.

Due to the nice properties of both IGD$^+$ and E_s, Falcón-Cardona *et al.* [6] proposed the CRI-EMOA algorithm that takes advantage of their selection properties. CRI-EMOA is a steady-state MOEA that switches between IGD$^+$-DE and E_s-DE depending on a statistical analysis of convergence. At each iteration, CRI-EMOA measures the convergence of its population, based on an approximation of HV that adds up all the Euclidean distances between the solutions in the population and a reference point. The convergence values are stored in a circular array of size T_w. Once the circular array is full, CRI-EMOA generates at each iteration a linear model of the convergence behavior from which it is obtained the angle θ related to the slope of the model and the coefficient of variation β of the samples. Based on these two values and two given thresholds $\bar{\beta}$ and $\bar{\theta}$, the IB-DE to be executed is selected. If the number of dominance layers produced by the nondominated sorting algorithm [16] is equal to one and $\beta \leq \bar{\beta}$ and $\theta \in [-\bar{\theta}, \bar{\theta}]$, it means that the convergence is stagnated because of a low variation of the convergence values and, hence, diversity should be promoted via E_s-DE. Otherwise, IGD$^+$-DE is performed. Its main drawbacks are related to the HV approximation and the user-supplied thresholds, which impact the performance of the switching mechanism between the IB-DEs.

IV. Our Proposed Approach

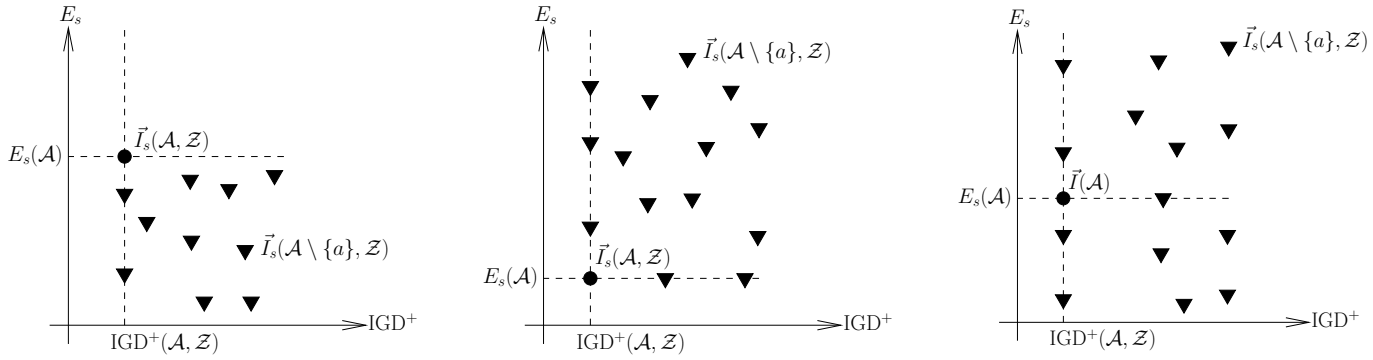
In this section, we first introduce the combination of IGD$^+$ and E_s in a single indicator from which a density estimator is designed. Then, we present the general description of PFI-EMOA.

A. Design of the Combined Indicator and Density Estimator

The underlying idea of PFI-EMOA is to unify IGD$^+$ and E_s in a single QI such that we exploit the trade-off between both indicators and, then, the new QI is embedded in a density estimator. In 2019, Falcón-Cardona *et al.* [17] proposed to combine a set of QIs (given by an indicator vector $\vec{I} = (I_1, \ldots, I_k)$) from which one or more are weakly Pareto-compliant QIs with at least one being Pareto-compliant, using an order-preserving function $u : \mathbb{R}^k \to \mathbb{R}$ such that the resulting combined indicator $\mathcal{I} = u(\vec{I})$ is Pareto-compliant. In this paper, we follow this combination approach to exploit the trade-off between IGD$^+$ and E_s.

Let $\vec{I}_s(\mathcal{A}, \mathcal{Z}) = (\text{IGD}^+(\mathcal{A}, \mathcal{Z}), E_s(\mathcal{A}))$ be our indicator vector. From the set of order-preserving functions, we chose, without loss of generality, the augmented Tchebycheff function (ATCH) which is defined as follows: $\text{ATCH}_{\vec{w}}(\vec{x}) = \max_{i=1,\ldots,k}\{w_i x_i\} + \alpha \sum_{i=1}^{k} x_i$, where $\alpha > 0$ and $\vec{w} \in \mathbb{R}^k$ is a weight vector, holding $\sum_{i=1}^{k} w_i = 1$ and all of its components should be strictly positive such that all the indicators contribute to the ATCH value. Hence, the combined indicator is $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A}, \mathcal{Z}))$. Since we are combining a weakly Pareto-compliant QI with a non-Pareto-compliant one, according to the theorem proposed by Falcón-Cardona *et al.* [17], the proposed combined indicator is not Pareto-compliant.

To design a density estimator based on $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A}, \mathcal{Z}))$, it is necessary to analyze the three cases of Figure 2 that arise due to the non-Pareto compliance of E_s. This figure shows the relation between $\vec{I}_s(\mathcal{A}, \mathcal{Z})$ and $\vec{I}_s(\mathcal{A} \setminus \{\vec{a}\}, \mathcal{Z})$. In Case 1, every time a solution is removed from $\mathcal{A}$, the E_s value gets better, which implies that $\vec{I}_s(\mathcal{A}, \mathcal{Z})$ is mutually nondominated with respect to all vectors $\vec{I}_s(\mathcal{A} \setminus \{\vec{a}\}, \mathcal{Z})$. Hence, $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A} \setminus \{a\}, \mathcal{Z}))$ could be less than, greater than or equal to $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A}, \mathcal{Z}))$. On the other hand, in Case 2, all vectors $\vec{I}_s(\mathcal{A} \setminus \{\vec{a}\}, \mathcal{Z})$ are dominated by $\vec{I}_s(\mathcal{A}, \mathcal{Z})$ since they all have a greater E_s value. Due to the order-preserving property of ATCH, $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A}, \mathcal{Z})) < \text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A} \setminus \{a\}, \mathcal{Z}))$, for all $\vec{a} \in \mathcal{A}$. Finally, case 3 combines cases 1 and 2, thus, $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A} \setminus \{a\}, \mathcal{Z}))$ could be less than, greater than or equal to $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A}, \mathcal{Z}))$. In the three cases, it is evident



(a) Case 1 (b) Case 2 (c) Case 3

Fig. 2: Due to the Pareto non-compliance of the Riesz s-energy indicator, there are three cases for selection when using $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A}, \mathcal{Z}))$.

that there exist non-contributing solutions to the IGD^+ indicator that must be eliminated every time they appear because they do not contribute to the convergence of the algorithm. In consequence, we delete the worst-contributing solution to the Riesz s-energy indicator in this case. Regarding the rest of solutions, we delete the solution having the minimum $\text{ATCH}_{\vec{w}}(\vec{I}_s(\mathcal{A} \setminus \{a\}, \mathcal{Z}))$ value.

B. PFI-EMOA: General Description

Algorithm 1 PFI-EMOA general framework

Require: Combination vector $\vec{w} \in \mathbb{R}^2$
Ensure: Pareto front approximation
1: Randomly initialize population P
2: **while** stopping criterion is not fulfilled **do**
3: Create a single offspring solution $\vec{q}$ using variation operators
4: $Q \leftarrow P \cup \{\vec{q}\}$
5: $\{R_1, \ldots, R_k\} \leftarrow$ nondominated-sorting(Q)
6: **if** $|R_k| > 1$ **then**
7: $z_i^{\text{max}} \leftarrow \max_{\vec{a} \in Q} a_i, i = 1, \ldots, m$
8: $z_i^{\text{min}} \leftarrow \min_{\vec{a} \in Q} a_i, i = 1, \ldots, m$
9: Normalize $\{R_j\}_{j=1,\ldots,k}$ using $\vec{z}^{\text{max}}$ and $\vec{z}^{\text{min}}$
10: $B \leftarrow \{\vec{b} \mid \text{IGD}^+(R_k \setminus \{\vec{b}\}, R_1) = \text{IGD}^+(R_k, R_1), \forall \vec{b} \in R_k\}$
11: **if** $|B| > 0$ **then**
12: $\vec{a}_{\text{worst}} \leftarrow \arg\max_{\vec{b} \in B} C_{E_s}(\vec{b}, R_k)$
13: **else**
14: $\vec{a}_{\text{worst}} \leftarrow \arg\min_{\vec{r} \in R_k} \text{ATCH}_{\vec{w}}(\vec{I}_s(R_k \setminus \{r\}, R_1))$
15: **end if**
16: **else**
17: $\vec{a}_{\text{worst}}$ is the sole solution in R_k
18: **end if**
19: $P \leftarrow Q \setminus \{\vec{a}_{\text{worst}}\}$
20: **end whilereturn** P

Algorithm 1 outlines PFI-EMOA that is a steady-state algorithm based on the framework of SMS-EMOA [14]. In contrast to CRI-EMOA that requires the parameters $\bar{\beta}$ and $\bar{\theta}$ to control the switching between the IB-DEs, PFI-EMOA only requires a weight vector $\vec{w} \in \mathbb{R}^2$ such that $w_1 + w_2 = 1$ and $w_1, w_2 > 0$ which control the importance of the indicators. Lines 2 to 17 encompass the main loop of PFI-EMOA where the population P (randomly initialized in line 1) is evolved to obtain a Pareto front approximation. At each iteration, the nondominated sorting algorithm [16] classifies P and an offspring solution generated through variation operators[4] in layers $R_1, \ldots, R_k$ according to the Pareto dominance relation. If the cardinality of R_k (which has the worst solutions according to the Pareto dominance) is greater than one, our density estimator is applied; otherwise, the sole solution in R_k is deleted. To execute the density estimator, it is first necessary to normalize the objective values of all solutions. In line 10, the set B of non-contributing solutions to the IGD^+ indicator is computed and if it has one or more solutions, the one having the worst Riesz s-energy contribution is selected. Otherwise, the solution in R_k having the worst trade-off between IGD^+ and E_s is selected in line 14. In both lines 10 and 14, R_1 is employed as the reference set of IGD^+. The final step of the loop is to delete in line 17 the selected solution. Finally, P is returned as the Pareto front approximation.

Regarding the computational cost of the new selection mechanism, we employed the efficient algorithms for calculating the individual contributions to IGD^+ and E_s, introduced in [18] and [19], respectively. For both indicators, given an approximation set of size N, the cost of computing all the individual contributions is $\Theta(N^2)$. In consequence, the runtime involved in lines 10 to 15 is described in the following. The identification of noncontributing solutions that shape B is in $\Theta(|R_k|^2)$. If B has one or more solutions, the identification of $\vec{a}_{\text{worst}}$ is $\Theta(|R_k|^2)$. Otherwise, the contributions to the combined indicator are calculated in $\Theta(m \cdot (|R_k|^2 + |R_k|^2)) = \Theta(m|R_k|^2)$, where the term m is due to the calculation of the ATCH value. Consequently, the use of the combined indicator in the selection mechanism does not aggregate a considerable

[4]We employed the simulated binary crossover (SBX) and the polynomial-based mutation (PBX) operators [16].

overhead.

V. Experimental Results

In this section, we analyze the performance of PFI-EMOA[5] by performing three experiments. First, we turn off the Riesz s-energy indicator in the combined indicator to determine what is its effect on the Pareto front approximations. Then, we compare PFI-EMOA with CRI-EMOA [6], AR-MOEA [4], RVEA [20], SPEA2+SDE [21], GrEA [22], and Two_Arch2 [4]. We used the WFG [2] and WFG^{-1} [3] benchmark problems with 2 to 6 objective functions to test the convergence and diversity properties of the selected MOEAs. For each test instance, we performed 30 independent executions and to obtain statistical confidence, we performed the Wilcoxon rank-sum test with a confidence value of 95%. To assess the performance of PFI-EMOA, we employed four QIs: IGD$^+$ and E_s since PFI-EMOA aims to optimize them, and, as neutral QIs, we used HV and the Solow-Polasky-Diversity indicator (SPD) [23]. For each test instance, we merged all the solutions from the MOEAs to produce a subset of mutually nondominated solutions of size $100m$ (where m is the number of objective functions) as the reference set required by IGD$^+$. In all cases, $s = m-1$ for E_s as suggested in [24] and [6]. Regarding HV, the reference point was set as follows: $\vec{z}_{\text{ref}} = \{2i+1\}_{i=1,\ldots,m}$ for all the WFG problems, and $\vec{z}_{\text{ref}} = (10,\ldots,10)$ for all the WFG^{-1} instances. SPD uses $\theta = 10$ for its computation [23]. Finally, we analyze the three possible ways in which PFI-EMOA deletes a solution at each iteration, aiming to determine which combined indicator to adopt.

A. Parameters Settings

Let $(m;\mu;T)$ be a configuration tuple where m is the number of objective functions, μ is the population size, and T is the maximum number of function evaluations. For a fair comparison, PFI-EMOA and the other MOEAs used the same population size and the same number of function evaluations as their stopping criterion as follows: $(2;120;40,000)$, $(3;120;50,000)$, $(4;120;60,000)$, $(5;126;70,000)$, $(6;126;80,000)$. Since all the MOEAs employ SBX and PBX as their variation operators, we set the crossover probability (P_c), the crossover distribution index (N_c), the mutation probability (P_m), and the mutation distribution index (N_m) as follows. For MOPs having two and three objective functions $P_c = 0.9$ and $N_c = 20$, while for MOPs with $m > 3$, $P_c = 1.0$ and $N_c = 30$. In all cases, $P_m = 1/n$ where n is the number of decision variables and $N_m = 20$. The number of decision variables n and the number of position-related parameters k of both the WFG and WFG^{-1} test problems are $n = 24 + 2(m-2)$ and $k = 2(m-1)$. Regarding PFI-EMOA, we set $\vec{w} = (0.5, 0.5)$ for two- and three-dimensional MOPs while $\vec{w} = (0.9, 0.1)$ for MOPs having $m > 3$, where w_1 is related to IGD$^+$ and w_2 to E_s. The reason for this decision is that for many-objetive optimization problems, the number of nondominated solutions increases exponentially and, in some cases, E_s could reward solutions with low convergence degree. Hence, IGD$^+$ should increase the selection pressure, which is done by setting $w_1 = 0.9$. CRI-EMOA uses $T_w = \mu$, $\bar{\beta} = 0.1$ and $\bar{\theta} = 0.25$ for all MOPs. The size of the convergence archive of Two_Arch2 is equal to the population size and the fractional distance is set to $1/m$ for all the test instances. Concerning RVEA, the rate of change of the penalty is set to 2 and the frequency of employing the reference vector adaptation is equal to 0.1 in all cases. GrEA creates 45, 15, 10, 9, and 9 divisions of the objective space for 2, 3, 4, 5, and 6 objective functions, respectively. We employed PlatEMO 2.0 [25] to execute AR-MOEA, RVEA, GrEA, SPEA2+SDE, and Two_Arch2 while for CRI-EMOA, we used the source code available at http://computacion.cs.cinvestav.mx/~jfalcon/CRI-EMOA.html.

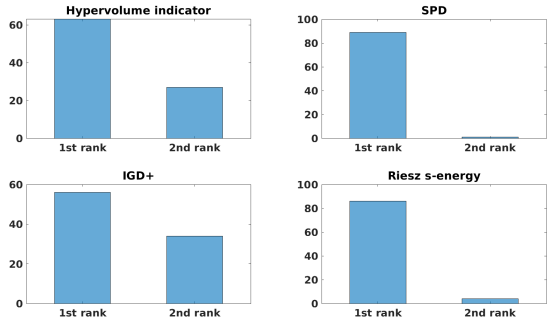


Fig. 3: For each considered indicator, the number of test instances for which PFI-EMOA is ranked first or second when compared to PFI-EMOA/E_s.

B. Effect of Riesz s-energy

In this experiment, we analyze what is the effect of the Riesz s-energy indicator in the final quality of the Pareto front approximations generated by PFI-EMOA. To this aim, we turned off the E_s-contribution in PFI-EMOA, i.e., we set $\vec{w} = (1, 0)$, where the zero value is related to E_s, and we changed ATCH by the weighted sum function ($\text{WS}_{\vec{w}}(\vec{x}) = \sum_{i=1}^{k} w_i x_i$) in the combined indicator. The reason to use WS instead of ATCH is that even though we set $\vec{w} = (1, 0)$ in ATCH, E_s would be taken into account in the correction factor $\alpha \sum_{i=1}^{k} x_i$ of ATCH. We denote this modified PFI-EMOA as PFI-EMOA/E_s and we compare it with PFI-EMOA in the 90 test instances indicated before, using HV, SPD, IGD$^+$, and E_s. Due to space limitations,[6] in Figure 3 we show a summary of the comparison where we stress the number of test instances for which PFI-EMOA is ranked first or second by each QI. Regarding the SPD and E_s values in the figure, it is clear that E_s helps PFI-EMOA to significantly improve the diversity of the Pareto front approximations. On the other hand, PFI-EMOA performs better than PFI-EMOA/E_s in almost 60 out of the 90 test instances for both HV and IGD$^+$. According to

[5]The source code of PFI-EMOA is available at http://computacion.cs.cinvestav.mx/~jfalcon/PFI-EMOA.html.

[6]The complete numerical results are available at http://computacion.cs.cinvestav.mx/~jfalcon/PFI-EMOA.html.

the numerical results, PFI-EMOA/E_s is better than PFI-EMOA for WFG1, WFG6, WFG7, WFG8 and WFG9 in terms of HV. This is because PFI-EMOA/E_s is being guided by an IGD$^+$-DE which produces similar distributions to HV (see [18], [13]) and, hence, in the performance comparison HV is rewarding this behavior. In this light, IGD$^+$ prefers PFI-EMOA/E_s in some cases because its preferences are similar to those of HV. However, PFI-EMOA obtained the first rank in both indicators in almost 66.66% of the benchmark problems, which implies that E_s helps to improve the convergence quality as well.

C. Comparison with state-of-the-art MOEAs

We analyzed the convergence and diversity properties of PFI-EMOA with respect to six state-of-the-art MOEAs: GrEA and RVEA are two approaches that aim to balance convergence and diversity during the evolutionary process, and CRI-EMOA, AR-MOEA, SPEA2+SDE, and Two_Arch2 which were specifically designed to have a good performance regardless of the Pareto front shape of the MOP being tackled. Figure 4 presents the statistical ranks obtained by the considered MOPs for each quality indicator. The underlying numerical results are available at http://computacion.cs.cinvestav.mx/~jfalcon/PFI-EMOA.html. From the figure, it is clear that PFI-EMOA simultaneously optimizes IGD$^+$ and E_s, being the best-ranked algorithm. This is also clear from HV and SPD which are neutral indicators in the comparison, i.e., no algorithm aims to optimize such QIs. Figure 5 shows some Pareto front approximations generated by all the MOEAs, where it is possible to see that PFI-EMOA produces the best distributions. An important factor to discuss is the effect of the weight vector $\vec{w}$ in the performance of PFI-EMOA. Based on a wide range of experiments, we observed that for two- and three-dimensional objective spaces, IGD$^+$ and E_s can be equally important (i.e., $\vec{w} = (0.5, 0.5)$) because there are not that many mutually nondominated solutions. However, for many-objective problems, the number of mutually nondominated solutions drastically increases, and, in these cases, IGD$^+$ should apply more selection pressure while E_s helps PFI-EMOA just to refine the ordering of solutions. If we give E_s more importance in high-dimensional spaces, it is more likely to allow dominated solutions to survive.

D. Selection analysis

At each iteration, PFI-EMOA deletes a solution $\vec{a}_{\text{worst}}$ from the population by: (1) selecting the worst-contributing solution to E_s from the set of non-contributing solutions to IGD$^+$, i.e., by line 12 of Algorithm 1, (2) using the ATCH$_{\vec{w}}(\vec{I}_s(\mathcal{A}, \mathcal{Z}))$-based density estimator in line 14, and (3) determining the worst solution in terms of the Pareto dominance relation in line 17. These three cases are denoted as C1, C2, and C3. In this section, we analyze the tendency of PFI-EMOA to use these three selection criteria. In Figure 6, we present statistical data of the utilization of C1, C2, and C3 for MOPs WFG1-WFG4 and their corresponding inverted instances for 2 to 6 objective functions. For all two-objective test instances (except for WFG1^{-1}), C3 is the most employed selection case, followed by C2, and C1 in that order. This is because in two-dimensional objective spaces, it is more likely that the nondominated sorting algorithm creates numerous layers and, therefore, the Pareto-based selection is executed more times. On the other hand, as the dimension of the objective space increases, the number of mutually nondominated solutions increases as well [26]. Consequently, for 3 to 6 objective functions, Figure 6 shows that C2 is the most employed selection criterion in a significant manner, followed by C3 and C1. Hence, this result shows that our proposed combined indicator is mostly guiding the search of PFI-EMOA for many-objective problems.

VI. Conclusions and Future Work

In this paper, we proposed to exploit the trade-off between the IGD$^+$ and Riesz s-energy indicators (that assess convergence and diversity, respectively) to design a selection mechanism that produces Pareto front approximations having both properties simultaneously. To this aim, we combined IGD$^+$ and Riesz s-energy in a single quality indicator that is then embedded in a density estimator of a steady-state MOEA, called PFI-EMOA. Our experimental results based on the hypervolume, Solow-Polasky Diversity, IGD$^+$, and Riesz s-energy indicators showed that PFI-EMOA outperforms several state-of-the-art MOEAs on the WFG and WFG^{-1} test problems. As part of our future work, we want to study the properties of the combined indicator and test other combinations of indicators.

References

[1] Carlos A. Coello Coello, Gary B. Lamont, and David A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, New York, second edition, September 2007. ISBN 978-0-387-33254-3.

[2] Simon Huband, Phil Hingston, Luigi Barone, and Lyndon While. A Review of Multiobjective Test Problems and a Scalable Test Problem Toolkit. *IEEE Transactions on Evolutionary Computation*, 10(5):477–506, October 2006.

[3] Hisao Ishibuchi, Yu Setoguchi, Hiroyuki Masuda, and Yusuke Nojima. Performance of Decomposition-Based Many-Objective Algorithms Strongly Depends on Pareto Front Shapes. *IEEE Transactions on Evolutionary Computation*, 21(2):169–190, April 2017.

[4] Handing Wang, Licheng Jiao, and Xin Yao. Two_Arch2: An Improved Two-Archive Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 19(4):524–541, August 2015.

[5] Y. Tian, R. Cheng, X. Zhang, F. Cheng, and Y. Jin. An indicator-based multiobjective evolutionary algorithm with reference point adaptation for better versatility. *IEEE Transactions on Evolutionary Computation*, 22(4):609–622, Aug 2018.

[6] Jesús Guillermo Falcón-Cardona, Carlos A. Coello Coello, and Michael Emmerich. CRI-EMOA: A Pareto-Front Shape Invariant Evolutionary Multi-Objective Algorithm. In *Evolutionary Multi-Criterion Optimization, 10th International Conference, EMO 2019*, pages 307–318. Springer. Lecture Notes in Computer Science Vol. 11411, East Lansing, Michigan, USA, March 10–13 2019. ISBN 978-3-030-12597-4.

[7] Miqing Li and Xin Yao. Quality evaluation of solution sets in multiobjective optimisation: A survey. *ACM Computing Surveys*, 52(2):26:1–26:38, March 2019.

[8] Jesús Guillermo Falcón-Cardona, Michael Emmerich, and Carlos A. Coello Coello. On the Cooperation of Multiple Indicator-based Multi-Objective Evolutionary Algorithms. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 2050–2057, June 2019.

[9] D. P. Hardin and E. B. Saff. Minimal Riesz energy point configurations for rectifiable d-dimensional manifolds. *Advances in Mathematics*, 193(1):174–204, 2005.

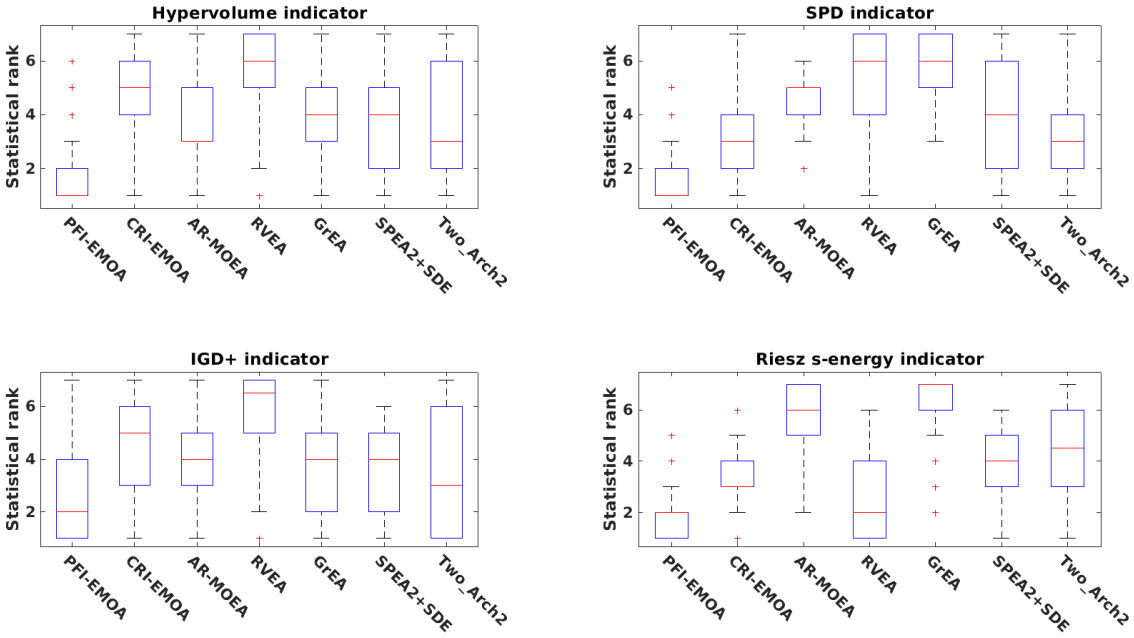


Fig. 4: Statistical ranks obtained by each algorithm over all benchmark functions with respect to each considered indicator. The lower, the better.

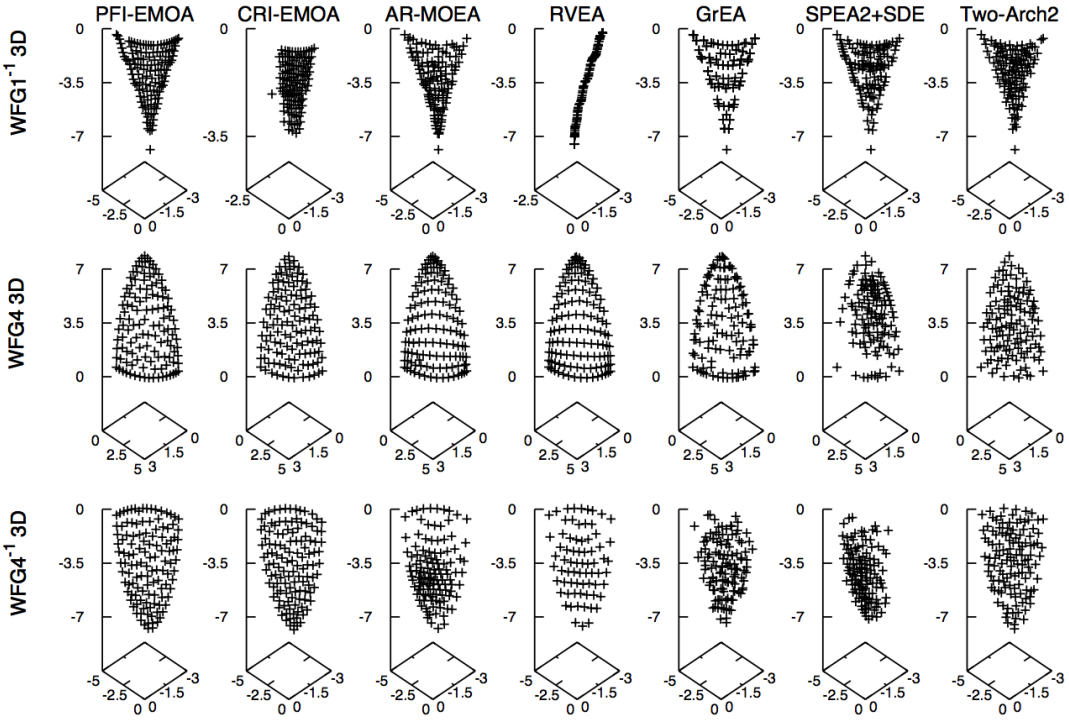


Fig. 5: Pareto front approximations generated by PFI-EMOA and the selected MOEAs. Each front corresponds to the median of the hypervolume indicator.

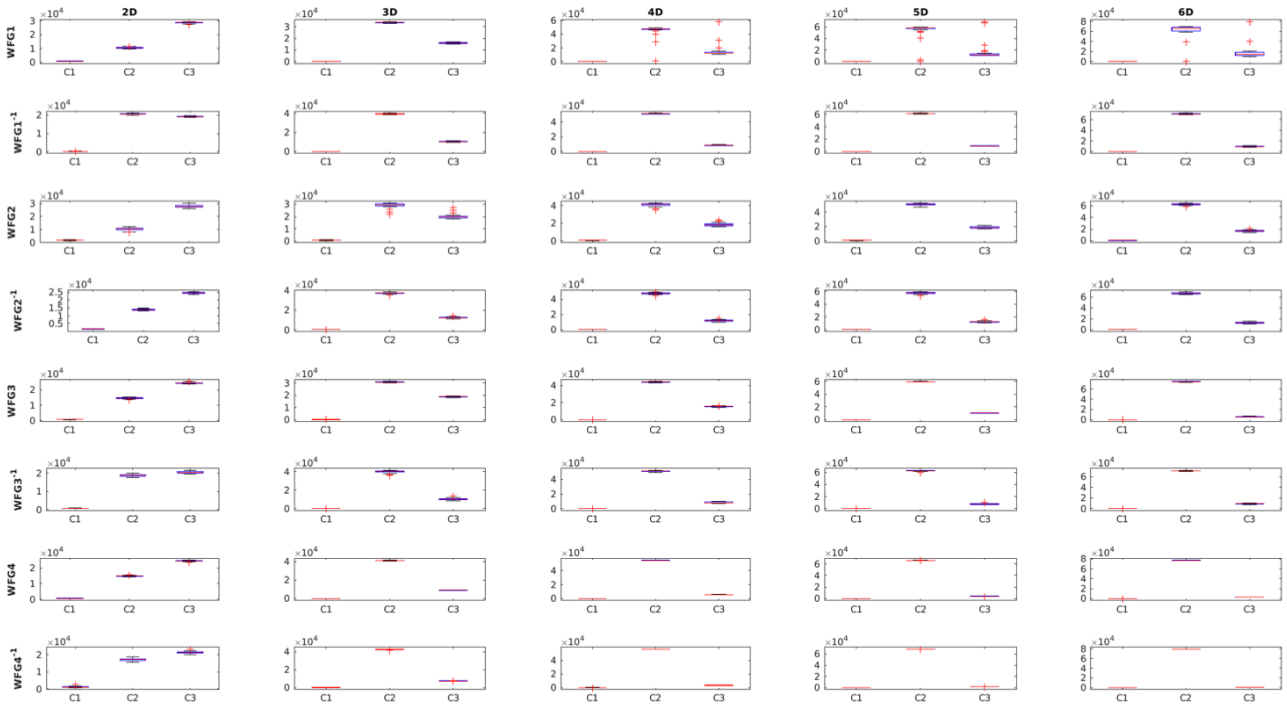


Fig. 6: Utilization of the three ways to select the worst solution of the population in Algorithm 1.

[10] Hisao Ishibuchi, Hiroyuki Masuda, Yuki Tanigaki, and Yusuke Nojima. Modified Distance Calculation in Generational Distance and Inverted Generational Distance. In António Gaspar-Cunha, Carlos Henggeler Antunes, and Carlos Coello Coello, editors, *Evolutionary Multi-Criterion Optimization, 8th International Conference, EMO 2015*, pages 110–125. Springer. Lecture Notes in Computer Science Vol. 9019, Guimarães, Portugal, March 29 - April 1 2015.

[11] Eckart Zitzler and Lothar Thiele. Multiobjective Optimization Using Evolutionary Algorithms—A Comparative Study. In A. E. Eiben, editor, *Parallel Problem Solving from Nature V*, pages 292–301, Amsterdam, September 1998. Springer-Verlag.

[12] Kalyanmoy Deb, Lothar Thiele, Marco Laumanns, and Eckart Zitzler. Scalable Test Problems for Evolutionary Multiobjective Optimization. In Ajith Abraham, Lakhmi Jain, and Robert Goldberg, editors, *Evolutionary Multiobjective Optimization. Theoretical Advances and Applications*, pages 105–145. Springer, USA, 2005.

[13] Hisao Ishibuchi, Ryo Imada, Naoki Masuyama, and Yusuke Nojima. Comparison of hypervolume, igd and igd+ from the viewpoint of optimal distributions of solutions. In Kalyanmoy Deb, Erik Goodman, Carlos A. Coello Coello, Kathrin Klamroth, Kaisa Miettinen, Sanaz Mostaghim, and Patrick Reed, editors, *Evolutionary Multi-Criterion Optimization*, pages 332–345, Cham, 2019. Springer International Publishing.

[14] Nicola Beume, Boris Naujoks, and Michael Emmerich. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 181(3):1653–1669, 16 September 2007.

[15] Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M. Fonseca, and Viviane Grunert da Fonseca. Performance Assessment of Multiobjective Optimizers: An Analysis and Review. *IEEE Transactions on Evolutionary Computation*, 7(2):117–132, April 2003.

[16] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA–II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, April 2002.

[17] Jesús Guillermo Falcón-Cardona, Michael T.M. Emmerich, and Carlos A. Coello Coello. On the Construction of Pareto-Compliant Quality Indicators. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO 19, pages 2024–2027, New York, NY, USA, 2019. ACM.

[18] Jesús Guillermo Falcón-Cardona and Carlos A. Coello Coello. Towards a More General Many-objective Evolutionary Optimizer. In *Parallel Problem Solving from Nature – PPSN XV, 15th International Conference, Proceedings, Part I*, pages 335–346. Springer. Lecture Notes in Computer Science Vol. 11101, Coimbra, Portugal, September 8–12 2018. ISBN: 978-3-319-99258-7.

[19] Jesús Guillermo Falcón-Cardona, Hisao Ishibuchi, and Carlos A. Coello Coello. Riesz s-energy-based reference sets for multi-objective optimization. In *2020 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8, 2020.

[20] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff. A reference vector guided evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 20(5):773–791, Oct 2016.

[21] Miqing Li, Shengxiang Yang, and Xiaohui Liu. Shift-Based Density Estimation for Pareto-Based Algorithms in Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 18(3):348–365, June 2014.

[22] Shengxiang Yang, Miqing Li, Xiaohui Liu, and Jinhua Zheng. A Grid-Based Evolutionary Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 17(5):721–736, October 2013.

[23] Vitor Basto-Fernandes, Iryna Yevseyeva, André Deutz, and Michael Emmerich. A Survey of Diversity Oriented Optimization: Problems, Indicators, and Algorithms. In Michael Emmerich, André Deutz, Oliver Schütze, Pierrick Legrand, Emilia Tantar, and Alexandru-Adrian Tantar, editors, *EVOLVE – A Bridge between Probability, Set Oriented Numerics and Evolutionary Computation VII*, pages 3–23. Springer, Cham, Switzerland, 2017. ISBN 978-3-319-49325-1.

[24] Raquel Hernández Gómez and Carlos A. Coello Coello. A Hyper-Heuristic of Scalarizing Functions. In *2017 Genetic and Evolutionary Computation Conference (GECCO'2017)*, pages 577 –584, Berlin, Germany, July 15-19 2017. ACM Press. ISBN 978-1-4503-4920-8.

[25] Y. Tian, R. Cheng, X. Zhang, and Y. Jin. Platemo: A matlab platform for evolutionary multi-objective optimization [educational forum]. *IEEE Computational Intelligence Magazine*, 12(4):73–87, Nov 2017.

[26] M. Farina and P. Amato. On the Optimal Solution Definition for Many-criteria Optimization Problems. In *Proceedings of the NAFIPS-FLINT International Conference'2002*, pages 233–238, Piscataway, New Jersey, June 2002. IEEE Service Center.