On the Construction of Pareto-compliant Combined Indicators

Technical Report EVOCINV-01-2020

Jesús Guillermo Falcón-Cardona
Computer Science Department
CINVESTAV-IPN
Mexico City, 07360, Mexico
jfalcon@computacion.cs.cinvestav.mx

Michael T. M. Emmerich
LIACS
Leiden University
Leiden, 2333, The Netherlands
m.t.m.emmerich@liacs.leidenuniv.nl

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

Abstract

The most relevant property that a quality indicator (QI) is expected to have is Pareto compliance, which means that every time an approximation set is better than another in a Pareto sense, the indicator must reflect this. The hypervolume indicator and its variants are the only unary QIs known to be Pareto-compliant but there are many commonly used weakly Pareto-compliant indicators such as R2, IGD$^+$ and ϵ^+. Currently, an open research area is related to finding new Pareto-compliant indicators whose preferences are different to those of the hypervolume indicator. In this paper, we propose a theoretical basis to combine existing weakly Pareto-compliant indicators with at least one being Pareto-compliant, such that the resulting combined indicator is Pareto-compliant as well. Most importantly, we show that the combination of Pareto-compliant QIs with weakly Pareto-compliant indicators leads to indicators that inherit properties of the weakly compliant indicators in terms of optimal point distributions. The consequences of these new combined indicators are threefold: 1) to increase the variety of available Pareto-compliant QIs by correcting weakly Pareto-compliant indicators, 2) to introduce a general framework for the combination of QIs, and 3) to generate new selection mechanisms for multi-objective evolutionary algorithms where it is possible to achieve/adjust desired distributions on the Pareto front.

1 Introduction

Currently, there exists a plethora of multi-objective evolutionary algorithms (MOEAs) focused on solving multi-objective optimization problems (MOPs)

[10, 36]. An MOP involves minimizing a vector function $\vec{F}(\vec{x})$ composed of k objective functions $f_i(\vec{x}), i = 1, \ldots, k$ where $\vec{x}$ is the vector of decision variables, $\Omega \subseteq \mathbb{R}^n$ is the decision variable space and $\vec{F}(\vec{x})$ is the vector of m (≥ 2) objective functions, where $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$. Solving an MOP involves finding points that represent the best possible trade-offs among the objective functions. In this regard, the Pareto dominance relation[1] has been the commonly used optimality criterion to find a set of trade-off solutions. The particular set that yields that optimum solutions, when defined in decision variable space, is known as the Pareto Optimal Set and its image in objective space is known as the Pareto Optimal front ($\mathcal{PF}^*$).

Since each MOEA generates Pareto front approximations[2] with specific convergence and diversity properties, a critical aspect is to decide which one performs best. In the early days, approximation sets were visually compared, attending convergence, spread and uniformity [21]. However, as the dimensionality of the manifolds related to the Pareto front approximations increases, comparisons get more difficult. In consequence, new comparison methods have been proposed. First, Pareto dominance, which is the most general preference information, was extended to comparisons between sets as described in Table 1. A drawback of using the Pareto dominance relation is that it only captures convergence information but not the spread and uniformity of solutions. The other comparison method relies on the use of quality indicators (QIs) which are functions that assign a real value to one or more approximation sets simultaneously, according to specific preference information [43]. This way, they impose a total order on the set of approximation sets related to an MOP. The cornerstone of QIs can be traced back to the Ph.D. thesis of David Van Veldhuizen [35] who made a comprehensive review of the existing QIs. This study was followed by an important theoretical analysis of quality indicators, proposed by Ztizler *et al.* [43], relating their properties to the various preorders on approximation sets defined in Table 1. Since then, QIs have been employed to quantitatively compare MOEAs [27, 30, 29].

---

[1]Given $\vec{x}, \vec{y} \in \Omega$ and $\vec{F} : \mathbb{R}^n \to \mathbb{R}^m$, $\vec{x}$ Pareto dominates $\vec{y}$ (denoted as $\vec{F}(\vec{x}) \prec \vec{F}(\vec{y})$) iff $f_i(\vec{x}) \leq f_i(\vec{y})$ for all $i = 1, \ldots, m$ and there exists an index $j \in \{1, \ldots, m\}$ such that $f_j(\vec{x}) < f_j(\vec{y})$. In case $f_i(\vec{x}) \leq f_i(\vec{y})$ for all $i = 1, \ldots, m$, $\vec{x}$ is said to weakly Pareto dominate $\vec{y}$ (denoted as $\vec{F}(\vec{x}) \preceq \vec{F}(\vec{y})$). Assuming that $f_i(\vec{x}) < f_i(\vec{y})$ for all $i = 1, \ldots, m$, $\vec{x}$ is said to strongly Pareto dominate $\vec{y}$ (denoted as $\vec{F}(\vec{x}) \prec\prec \vec{F}(\vec{y})$).

[2]Let $\mathcal{A} \in \Psi$ be a finite set of m-dimensional objective vectors. $\mathcal{A}$ is called a Pareto front approximation or approximation set if $\forall \vec{u}, \vec{v} \in \mathcal{A}, \vec{u} \neq \vec{v}$ it holds that $\vec{u} \not\preceq \vec{v} \wedge \vec{v} \not\preceq \vec{u}$. The set of all approximation sets is denoted as Ψ.

Table 1: Relations on approximation sets based on Pareto dominance relations. $\mathcal{A} \prec\prec \mathcal{B} \Rightarrow \mathcal{A} \prec \mathcal{B} \Rightarrow \mathcal{A} \lhd \mathcal{B} \Rightarrow \mathcal{A} \preceq \mathcal{B}$.

Relation	Description
$\mathcal{A} \prec\prec \mathcal{B}$	$\forall \vec{b} \in \mathcal{B}, \exists \vec{a} \in \mathcal{A} : \vec{a} \prec\prec \vec{b}$
$\mathcal{A} \prec \mathcal{B}$	$\forall \vec{b} \in \mathcal{B}, \exists \vec{a} \in \mathcal{A} : \vec{a} \prec \vec{b}$
$\mathcal{A} \lhd \mathcal{B}$	$\forall \vec{b} \in \mathcal{B}, \exists \vec{a} \in \mathcal{A} : \vec{a} \preceq \vec{b} \wedge \mathcal{A} \neq \mathcal{B}$
$\mathcal{A} \preceq \mathcal{B}$	$\forall \vec{b} \in \mathcal{B}, \exists \vec{a} \in \mathcal{A} : \vec{a} \preceq \vec{b}$
$\mathcal{A} \parallel \mathcal{B}$	$\mathcal{A} \npreceq \mathcal{B} \wedge \mathcal{B} \npreceq \mathcal{A}$

In the specialized literature, there are several QIs that aim to measure the convergence, spread and uniformity of approximation sets [29]. Among QIs, those focused on assessing convergence have considerably attracted the attention of the evolutionary multi-objective optimization community [43, 30] since they can be used not only to assess the performance of MOEAs but also to design selection mechanisms [39, 4, 8, 16]. A remarkable property of convergence QIs[3] is Pareto compliance or Pareto compatibility which, in words, means that every time an approximation set dominates another, the result of the indicator when evaluating the dominating set is always better than the indicator value of the dominated set (see Property 1). Zitzler *et al.* [43] introduced a mathematical analysis where a number of QIs were analyzed regarding their compatibility and completeness with the set dominance relations of Table 1. From this study, the hypervolume indicator (HV) [41] was found to be the only unary QI that is Pareto-compliant. HV measures the extent of the volume dominated by an approximation set and bounded by an anti-optimal reference point, i.e., a point that is dominated by all points in the Pareto front approximation. An important drawback of HV is that under NP $\neq$ P, its computational cost increases super-polynomially as the dimension of the objective function does. Therefore, other QIs have been proposed, being less expensive but weakly Pareto-compliant (see Property 2) or not Pareto-compliant. For example, the most noteworthy weakly Pareto-compliant QIs are R2 [7], the Inverted Generational Distance plus (IGD$^+$) [25], and the unary additive ϵ indicator (ϵ^+) [43] while IGD [9] and

---

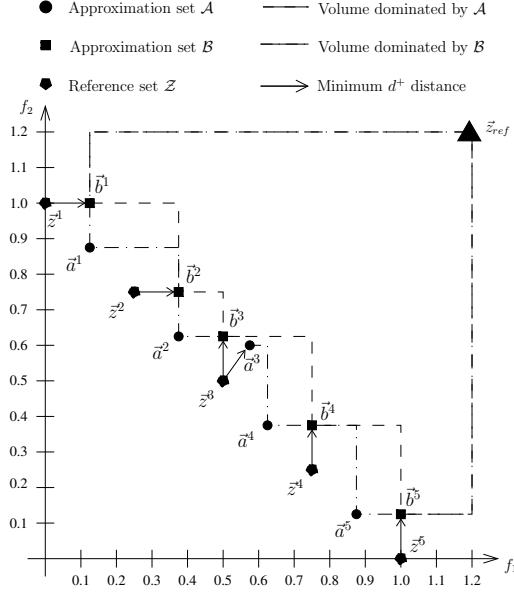[3]From this point onwards, convergence QIs will be denoted just as QIs.

Figure 1: Let $\mathcal{A} = \{(0.125, 0.875), (0.375, 0.625), (0.575, 0.6), (0.625, 0.375), (0.875, 0.125)\}$, $\mathcal{B} = \{(0.125, 1), (0.375, 0.75), (0.5, 0.625), (0.75, 0.375), (1, 0.125)\}$, and $\mathcal{Z} = \{(0, 1), (0.25, 0.75), (0.5, 0.5), (0.75, 0.25), (1, 0)\}$. Even though $\mathcal{A} \lhd \mathcal{B}$, IGD prefers $\mathcal{B}$ and IGD$^+$ assigns the same quality to both sets. HV prefers $\mathcal{A}$ since it is Pareto-compliant.

Generational Distance [35] are not Pareto-compliant indicators.

Currently, due to the high impact of the so-called many-objective optimization problems (i.e., MOPs having more than three objective functions), weakly and not Pareto-compliant QIs have been widely employed to compare MOEAs [11, 37, 34, 28]. However, not using Pareto-compliant QIs to assess MOEAs could lead to misleading results. In the following, let's consider the approximation sets $\mathcal{A}$ and $\mathcal{B}$ of Figure 1 for which we will compute HV, IGD$^+$, and IGD. It is worth noting that a lower value of IGD$^+$ and IGD means higher quality in constrast to HV that aims to maximize the dominated volume. From the figure, it is clear that $\mathcal{A} \lhd \mathcal{B}$ which is reflected by HV since HV$(\mathcal{A}, \vec{z}_{ref}) = 0.781875$ and HV$(\mathcal{B}, \vec{z}_{ref}) = 0.671250$. On the other hand, the IGD$^+$ indicator cannot decide which set is better since they have the same value equal to 0.125. Regarding IGD, $\mathcal{B}$ is given more quality because IGD$(\mathcal{B}, \mathcal{Z}) = 0.125$ and IGD$(\mathcal{A}, \mathcal{Z}) = 0.167705$ even though it is dominated by $\mathcal{A}$. Based on the above results, it is evident that using QIs with

4

weaker compliance properties could lead to false conclusions when assessing MOEAs. Another critical drawback is related to the so-called deteriorative cycles in the search process of MOEAs [42, 2]. A set-based MOEA that uses IGD$^+$ or IGD as its selection mechanism could easily get stuck in a cyclic selection. In the example of Figure 1, the MOEA could select either $\mathcal{A}$ or $\mathcal{B}$ since they have the same IGD$^+$ value. This simple example illustrates the need to construct new Pareto-compliant QIs.

From the above discussion, it is clear the need to construct new Pareto-compliant QIs. To support this claim and to justify why weak Pareto compliance might be not enough to ask for, we will consider the following example that shows that weak Pareto compliance is not sufficient as a guideline for constructing meaningful indicators. The example is an indicator which we will call "Zero indicator". It is defined as $Z : \Psi \to \mathbb{R}$ with $Z \equiv 0$. Clearly, for every $\mathcal{A}, \mathcal{B} \in \Psi$ such that $\mathcal{A} \lhd \mathcal{B}$, it implies $Z(\mathcal{A}) = Z(\mathcal{B})$, i.e., Z is weakly Pareto-compliant. Although indicators such as R2, IGD$^+$ and ϵ^+ are more complex than Z in a mathematical sense, all of them are only weakly Pareto-compliant as the Zero indicator. Hence, we can see that is not enough to construct a weakly Pareto-compliant QI and that Pareto compliance is indeed desirable to avoid misleading results when comparing approximation sets.

Currently, an open research area is to find new Pareto-compliant QIs, having preferences that are different from those of HV. For instance, if a set of MOEAs is to be assessed based on HV and one of them uses an HV-based selection mechanism, then HV will reward this hypothetical MOEA. This situation can be avoided if the comparison is performed using a more neutral Pareto-compliant QI, i.e., one having different preferences. In this paper, we propose to construct new Pareto-compliant QIs based on the mathematical combination of one or more weakly Pareto-compliant indicators with at least one Pareto-compliant QI, using an order-preserving combination function. Under these conditions, the combined indicators preserve the Pareto compliance property. Additionally, our framework allows to create Pareto-compliant QIs with different preferences to those of HV in two ways: 1) exploiting the conflict that sometimes exists between the preferences of indicators, such that the combined indicator shows intermediate preferences, and 2) keeping the original preferences of the weakly Pareto-compliant QIs but using a correcting term derived from the Pareto-compliant QI being used. Overall, the contributions of this paper are the following:

1. We provide a guideline to the construction of new Pareto-compliant QIs whose preferences are essentially different from those of HV.

2. This is the first theoretical study of the combined properties of indicators.

3. Our proposed framework allows correcting weakly Pareto-compliant QIs, such as R2, IGD$^+$, and ϵ^+, making them Pareto-compliant.

4. We introduce an empirical study of the optimal distribution of solutions generated by a steady-state MOEA based on some selected new Pareto-compliant QIs.

The remainder of the paper is organized as follows. Section 2 introduces definitions related to QIs. Our mathematical framework for construction of new Pareto-compliant QIS is introduced in Section 3. The experimental results using the combined indicators is presented in Section 4. Finally, the main conclusions and future work are described in Section 5.

2   Quality Indicators

In the following, we introduce the formal definition of a unary quality indicator, the Pareto compliance properties, and the indicators HV, R2, IGD$^+$, and ϵ^+. In all cases, let $\mathcal{A}$ denote an approximation set and $\mathcal{Z}$ be a reference set that discretizes $\mathcal{PF}^*$.

Definition 1 (Unary Quality Indicator). *A unary quality indicator I is a function $I : \Psi \to \mathbb{R}$, which assigns a real value to each approximation set $\mathcal{A} \in \Psi$, where Ψ is the set of all approximation sets for an MOP.*

Hansen and Jaszkiewicz [18] defined when the evaluation of two approximation sets by a certain indicator is compatible with the result of a Pareto-based outperformance relation applied to these two sets. Hence, an indicator could be compliant or weakly compliant with the outperformance relation $\lhd$ that is defined in Table 1. Both properties are defined in the following. Without loss of generality, let us assume that a greater indicator value corresponds to a higher quality.

Property 1 (Pareto compliance). *Given two approximation sets $\mathcal{A}$ and $\mathcal{B}$, a unary indicator I is $\lhd$-compliant (Pareto-compliant) if $\mathcal{A} \lhd \mathcal{B} \Rightarrow I(\mathcal{A}) > I(\mathcal{B})$.*

Property 2 (Weakly Pareto compliance). *Given two approximation sets $\mathcal{A}$ and $\mathcal{B}$, a unary indicator I is weakly $\lhd$-compliant (weakly Pareto-compliant) if $\mathcal{A} \lhd \mathcal{B} \Rightarrow I(\mathcal{A}) \geq I(\mathcal{B})$.*

Definition 2 (Hypervolume indicator). *Let Λ denote the Lebesgue measure in $\mathbb{R}^m$, the hypervolume indicator (HV) is defined as follows:*

$$HV(\mathcal{A}, \vec{z}_{ref}) = \Lambda \left(\bigcup_{\vec{a} \in \mathcal{A}} \{\vec{x} \mid \vec{a} \prec \vec{x} \prec \vec{z}_{ref}\} \right), \tag{1}$$

where $\vec{z}_{ref} \in \mathbb{R}^m$ is a reference point which should be dominated by all points in $\mathcal{A}$.

HV is a convergence-diversity QI that measures the extent of volume jointly dominated by the points in $\mathcal{A}$ and bounded by $\vec{z}_{ref}$. Currently, HV and the closely related logarithmic HV [17], the weighted HV [1], and the free HV [15] are the only Pareto-compliant QIs. The two main drawbacks of HV are the following. First, under NP $\neq$ P, computational cost increases super-polynomially as the number of objective functions does [3]. The other issue is related to $\vec{z}_{ref}$ since the preferences of HV strongly depend on it [1, 23]. In other words, the specification of the reference is dependent on the Pareto front shape. It has been shown that the distribution of points is often concentrated on the boundary and in knee point regions.

Definition 3 (Unary R2 indicator). *The unary R2 indicator is defined as follows:*

$$R2(\mathcal{A}, W) = \frac{1}{|W|} \sum_{\vec{w} \in W} \min_{\vec{a} \in \mathcal{A}} \{u_{\vec{w}}(a)\}, \tag{2}$$

where W is a set of m-dimensional weight vectors and $u_{\vec{w}} : \mathbb{R}^m \to \mathbb{R}$ is a utility function, parameterized by $\vec{w} \in W$, that assigns a real value to each solution vector.

The R2 indicator is a convergence-diversity QI that measures the average minimum utility values of the approximation set with respect to a set of weight vectors. Its computational cost is in $\Theta(m|W| \cdot |\mathcal{A}|)$. Unlike the hypervolume indicator, the time complexity of R2 scales only linearly with the number of objectives. Its time complexity is, however, proportional to

the number of weight vectors[4], which has to grow exponentially in size, if the number of objectives increases and the same resolution of sampling is desired. A major conceptual difference with regard to the hypervolume indicator is that the R2 indicator does not require an anti-optimal reference point. Instead, it works with an ideal or utopian reference point. In many application problems, for instance, in error or cost minimization, there is a natural choice for an ideal point, but it is difficult to define an anti-optimal reference point. Hence it would be desirable to use the R2 indicator.

A problem, however, arises due to the fact that the R2 indicator is not Pareto-compliant, and it is only weakly Pareto-compliant. This makes it possible that a set might have equal R2 indicator values than another set, although it is dominated in the set order, or that sets degenerate if this indicator is used as a guideline in Pareto optimization. One might argue that these are rare cases, as they always involve shared coordinate values among points, and in most cases, the R2 indicator works well when comparing sets. In fact, in continuous unconstrained optimization, such cases might occur with low probability, but it is relatively likely in continuous optimization and in cases where box constraints are introduced.

Definition 4 (Inverted Generational Distance plus). *The IGD$^+$, for minimization, is defined as follows:*

$$IGD^+(\mathcal{A}, \mathcal{Z}) = \frac{1}{|\mathcal{Z}|} \sum_{\vec{z} \in \mathcal{Z}} \min_{\vec{a} \in \mathcal{A}} d^+(\vec{a}, \vec{z}) \tag{3}$$

where $d^+(\vec{a}, \vec{z}) = \sqrt{\sum_{k=1}^{m} (\max\{a_k - z_k, 0\})^2}$.

Ishibuchi *et al.* [25] proposed IGD$^+$ as an improvement of the IGD indicator [9]. Both QIs measure convergence and diversity of solutions simultaneously. However, IGD$^+$ is weakly Pareto-compliant while IGD is not Pareto-compliant [5]. IGD$^+$ measures the average distance from the reference set to the dominated space of the approximation set. Its computational cost is $\Theta(m|\mathcal{Z}| \cdot |\mathcal{A}|)$. A critical aspect is how to specify the reference set when no information is available at about $\mathcal{PF}^*$ [24].

---

[4]The Simplex-Lattice-Design method is usually employed to construct the set of weight vectors [38]. Using this method, the number of weight vectors is the following combinatorial number: $N = C_{m-1}^{H+m-1}$, where $H \in \mathbb{N}$ is a user-supplied parameter that determines the number of divisions of the space, and m is the number of objectives.

Definition 5 (Unary ϵ^+ indicator). *Mathematically, it is defined as follows:*

$$\epsilon^+(\mathcal{A}, \mathcal{Z}) = \max_{\vec{z} \in \mathcal{Z}} \min_{\vec{a} \in \mathcal{A}} \max_{1 \leq i \leq m} \{z_i - a_i\}. \tag{4}$$

The unary ϵ^+-indicator gives the minimum distance by which a Pareto front approximation needs to or can be translated in all dimensions at once in objective space such that a reference set is weakly dominated. In consequence, this QI exclusively measures convergence to $\mathcal{PF}^*$ and it is weakly Pareto-compliant. A remarkable aspect is that ϵ^+ does not require any parameters but, as in the case of IGD$^+$, a reference set has to be supplied. Additionally, ϵ^+ is not very sensitive to local changes in the solutions in $\mathcal{A}$ [6].

3   New Pareto-compliant indicators

The combination of quality indicators is a research topic that has not been widely studied. A remarkable exception is the averaged Hausdorff distance (Δ_p) proposed by Schütze *et al.* [33]. Δ_p combines modified versions of the indicators IGD and GD into a single value by selecting the maximum value between both QIs. It is worth noting that Δ_p inherits the not Pareto compliance of both IGD and GD. In this section, we propose the first systematic framework for combining QIs. Additionally, we provide the mathematical argumentation to ensure that when combining QIs with specific properties, the resulting combined indicator will be Pareto-compliant. This leads not only to new types of indicators but also proves to be a way to create new Pareto-compliant indicators with very different properties than the HV in terms of the distribution of points that they favor, and in terms of the parameters provided by the user. In the following, we present the mathematical framework for the combination of QIs.

Definition 6 (Combination function). *A combination function $C : \mathbb{R}^k \to \mathbb{R}$ assigns a real value to a vector $\vec{I} = (I_1, I_2, \ldots, I_k)$, where each I_j is the value of a unary indicator.*

Definition 7 (Combined Indicator). *Given an indicator vector $\vec{I} = (I_1, I_2, \ldots, I_k)$ and a combination function C, a combined indicator $\mathcal{I}$ is defined as follows: $\mathcal{I} = C(\vec{I})$.*
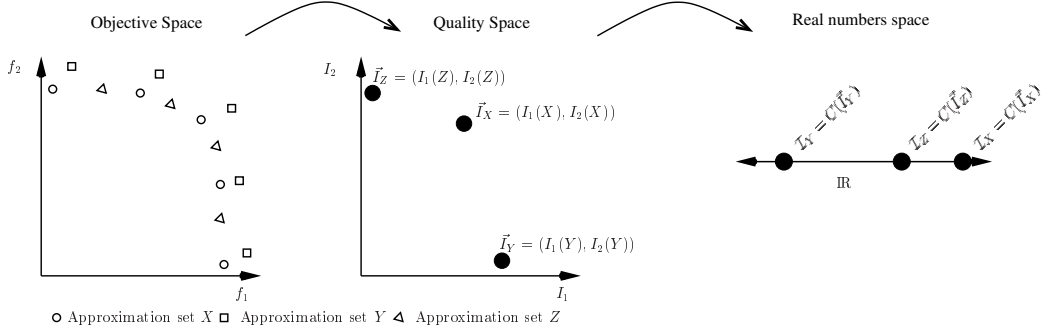
Figure 2: The objective space contains the approximation sets X, Y, and Z that are mapped to the quality space using an indicator vector. The points $\vec{I}_X, \vec{I}_Y$, and $\vec{I}_Z$ in quality space are then transformed each to a single real value by the combination function $\mathcal{C} : \mathbb{R}^2 \rightarrow \mathbb{R}$ to generate the real values $\mathcal{I}_X, \mathcal{I}_Y$, and $\mathcal{I}_Z$.

A general combined indicator $\mathcal{I}$ is a function that maps a vector of indicator values to a single real value as it is stated in Definitions 6 and 7. Figure 2 shows how to map Pareto front approximations (in objective space) to the quality space Q, where each axis correspond to a specific indicator. Then, indicator vectors in Q are assigned a real value, using the combination function C. Based on the above definitions, nothing can be said about the properties of $\mathcal{I}$ at this point. Hence, for getting more important theoretical results, we should say something about the properties of each $I_j, j = 1, \ldots, k$ and of the combination function. We are interested in analyzing the Pareto compliance of $\mathcal{I}$. Based on Properties 1 and 2, we construct a special vector of indicators that is necessary for the refinement of the combined indicator model.

Definition 8 (Compliant Indicator Vector). $\vec{I} = (I_1, I_2, \ldots, I_k) \in Q$ is called a compliant indicator vector (CIV) if $\forall j = 1, \ldots, k, I_j$ is weakly Pareto compliant and there exists at least one index $t \in \{1, \ldots, k\}$ such that I_t is Pareto compliant. $Q \subseteq \mathbb{R}^k$ is denoted as the quality space.

For the following Theorem, let us assume, without loss of generality, that the unary indicators $I_1, \ldots, I_k$ are to be maximized. Additionally, for the rest of this section, let us consider the Pareto dominance for maximization, i.e., given $\vec{x}, \vec{y} \in \mathbb{R}^k$, $\vec{x}$ Pareto dominates $\vec{y}$ (denoted as $\vec{x} \succ \vec{y}$) iff $x_i \geq y_i$ for all $i = 1, \ldots, k$ and there exists an index $j \in \{1, \ldots, k\}$ such that $x_j > y_j$.

10

Theorem 1 (Construction of Pareto-compliant combined indicators)**.** *Let* $I_1, \ldots, I_k$ *be unary indicators that form a compliant indicator vector* $\vec{I}$. *A combined indicator* $\mathcal{I}(\vec{I})$ *is* $\vartriangleleft$-*compliant if* $\mathcal{I}$ *has the order-preserving property:*

$$\forall \vec{u}, \vec{v} \in \mathbb{R}^k, \vec{u} \succ \vec{v} \Rightarrow \mathcal{I}(\vec{u}) > \mathcal{I}(\vec{v}).$$

Proof. Consider two approximation sets $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{A} \vartriangleleft \mathcal{B}$ and let $\vec{I}^{\mathcal{A}} := \vec{I}(\mathcal{A})$ and $\vec{I}^{\mathcal{B}} := \vec{I}(\mathcal{B})$, where $\vec{I}$ is a CIV. Then, $\mathcal{A} \vartriangleleft \mathcal{B} \Rightarrow \vec{I}^{\mathcal{A}} \succ \vec{I}^{\mathcal{B}}$ because the Pareto-compliant indicators get better and the weakly Pareto-compliant ones get better or stay equal. Moreover, by definition $\vec{I}^{\mathcal{A}} \succ \vec{I}^{\mathcal{B}} \Rightarrow \mathcal{I}(\vec{I}^{\mathcal{A}}) > \mathcal{I}(\vec{I}^{\mathcal{B}})$. Hence, by transitivity of $\Rightarrow$, it holds $\mathcal{A} \vartriangleleft \mathcal{B} \Rightarrow \mathcal{I}(\vec{I}^{\mathcal{A}}) > \mathcal{I}(\vec{I}^{\mathcal{B}})$, i.e., $\mathcal{I}$ is Pareto-compliant. $\square$

Theorem 1 provides a sufficient condition for constructing Pareto-compliant combined indicators on the basis of compliant indicator vectors. In other words, a combined indicator preserves the Pareto-compliant property because of the use of order-preserving combination functions.

Remark 1. *The condition of Theorem 1 is suffcient but not necessary. For instance, given* $\vec{I} = (I_1, I_2, \ldots, I_k)$ *where* I_1 *is Pareto-compliant and the* $I_j, j = 2, \ldots, k$ *are not Pareto-compliant, the "combined" indicator* $\mathcal{I}(\vec{I}) = I_1$ *is also Pareto-compliant. Hence, there is a large number of possibilities to construct combined and compliant indicators.*

There exist many combination functions that have the property of Theorem 1. However, in this paper, we focus on certain utility functions [31, 32] $u : \mathbb{R}^k \to \mathbb{R}$ that hold the desired property. A utility function (UF) is a model of the decision maker preferences that assigns to each k-dimensional vector a utility value. Thus, a combination function C can be defined in terms of these functions. Generally, UFs employ a convex weight vector $\vec{w} \in \mathbb{R}^k$ such that $\sum_{i=1}^{k} w_i = 1, w_i \geq 0$. However, for the combination of QIs, we need $w_i > 0$ for all $i = 1, \ldots, k$ such that all QIs contribute to the combined indicator value. Based on the above, a Pareto-compliant utility indicator (PCUI) is defined as follows:

Definition 9 (Utility indicator)**.** *Given a utility function* $u : \mathbb{R}^k \to \mathbb{R}$, *a CIV* $\vec{I} \in \mathbb{R}^k$ *that assesses an approximation set* $\mathcal{A}$ *and a weight vector* $\vec{w} \in \mathbb{R}^k$ *where* $w_i > 0$ *for all* $i = 1, \ldots, k$, *we denote a utility indicator as* $u_{\vec{w}}(\vec{I}(\mathcal{A}))$.

If u is also order-preserving as required in Theorem 1, $u_{\vec{w}}(\vec{I}(\mathcal{A}))$ is denoted as a Pareto-compliant utility indicator.

In this work, we focused our attention on two utility functions that are order-preserving, namely, the weighted sum function (WS) and the augmented Tchebycheff function (ATCH). However, there is a plethora of utility functions having this property. In the following, we prove that both WS and ATCH are order-preserving functions and, thus, can be employed to define PCUIs.

Definition 10. *The weighted sum (WS) is defined by the following formula:*

$$WS_{\vec{w}}(\vec{x}) = \sum_{i=1}^{k} w_i x_i, \tag{5}$$

where $\vec{x}, \vec{w} \in \mathbb{R}^k$ and $\sum_{i=1}^{k} w_i = 1, w_i \geq 0, \forall i = 1, \ldots, k$.

Lemma 1. *Given two CIVs $\vec{x}, \vec{y} \in \mathbb{R}^k$ and a weight vector $\vec{w} \in \mathbb{R}^k$ such that $w_i > 0$ for all $i = 1, \ldots, k$, then if $\vec{x} \succ \vec{y} \Rightarrow WS_{\vec{w}}(\vec{x}) > WS_{\vec{w}}(\vec{y})$.*

Proof. Since $\vec{x} \succ \vec{y}$, then $x_i \geq y_i$ for all $i = 1, \ldots, k$ and there exists at least an index $j \in \{1, \ldots, k\}$ such that $x_j > y_j$. If $w_i > 0$ for all $i = 1, \ldots, k$, then $w_i x_i \geq w_i y_i$ for all i and $w_j x_j > w_j y_j$ for at least one j. Hence, $WS_{\vec{w}}(\vec{x}) > WS_{\vec{w}}(\vec{y})$.

$\square$

For the Augmented Tchebycheff function, we slighly modified its original defintion. We do not consider the absolute values such that the function is order-preserving in the whole $\mathbb{R}^k$.

Definition 11 (Augmented Tchebycheff). *Given $\vec{x}, \vec{w} \in \mathbb{R}^k$ with $w_i \geq 0$ and $\alpha > 0$, the Augmented Tchebycheff function (ATCH) is defined as follows:*

$$ATCH_{\vec{w}}(\vec{x}) = \max_{i=1,\ldots,k} \{w_i x_i\} + \alpha \sum_{i=1}^{k} x_i \tag{6}$$

Lemma 2. *Given two CIVs $\vec{x}, \vec{y} \in \mathbb{R}^k$ and a weight vector $\vec{w} \in \mathbb{R}^k, w_i > 0, i = 1, \ldots, k$, then if $\vec{x} \succ \vec{y} \Rightarrow ATCH_{\vec{w}}(\vec{x}) > ATCH_{\vec{w}}(\vec{y})$.*

Proof. Since $\vec{x} \succ \vec{y}$, then $x_i \geq y_i$ for all $i = 1, \ldots, k$ and there exists at least an index $j \in \{1, \ldots, k\}$ such that $x_j > y_j$. Let's suppose, without loss of generality, that $w_t x_t = w_t y_t$ is the resulting maximum value and $t \neq j$. In case $t = j$, the proof is trivial. Additionally, since $\alpha > 0$, then $\alpha \sum_{i=1}^{k} x_i > \alpha \sum_{i=1}^{k} y_i$ because $x_j > y_j$. This implies that $w_t x_t + \alpha \sum_{i=1}^{k} x_i > w_t y_t + \alpha \sum_{i=1}^{k} y_i$. Hence, $\vec{x} \succ \vec{y} \Rightarrow \mathrm{ATCH}_{\vec{w}}(\vec{x}) > \mathrm{ATCH}_{\vec{w}}(\vec{y})$. $\qquad\square$

Finally, we need to punctualize the relationship between the order-preserving functions and the PCUIs and we need to clarify what is the effect of the combination weight vector $\vec{w}$ that PCUIs require. First, PCUIs are invariant to the indicator scales because of the order-preserving combination function u. No matter what order-preserving function is used, if $\vec{x} \succ \vec{y} \Rightarrow u(\vec{x}) > u(\vec{y})$. However, if $\vec{x}$ and $\vec{y}$ are mutually non-dominated, we cannot say what will be the relation between $u(\vec{x})$ and $u(\vec{y})$ unless we know the definition of u. In consequence, each u expresses specific preferences when dealing with non-dominated solutions and such preferences depend on the landscape of u (see Figure 3). On the other hand, PCUIs require a weight vector $\vec{w} = (w_1, \ldots, w_k)$ where $w_i > 0$ for all $i = 1, \ldots, k$. Each w_j assigns a relative importance to its associated indicator I_j. Figure 3 shows the landscape of ATCH function for three different $\vec{w}$. Depending on $\vec{w}$, the PCUI exploits in a different way the trade-off among its baseline indicators. When all w_i are equal, the PCUI preferences are the intermediate point between the preferences of its baseline QIs. For the other two cases in Figure 3, the preferences of the PCUI will be biased to the indicator having the greatest w_i. Supposing a PCUI is integrated in the selection mechanism of an MOEA, we could control the final distribution of points by defining $\vec{w}$.

4 Experimental results

In this section, we present the results of two experiments that aim to provide empirical information on the following PCUIs: $\mathrm{WS}_{\vec{w}}(\mathrm{HV}, \mathrm{R2})$, $\mathrm{ATCH}_{\vec{w}}(\mathrm{HV}, \mathrm{R2})$, $\mathrm{WS}_{\vec{w}}(\mathrm{HV}, \mathrm{IGD}^+)$, $\mathrm{ATCH}_{\vec{w}}(\mathrm{HV}, \mathrm{IGD}^+)$, $\mathrm{WS}_{\vec{w}}(\mathrm{HV}, \epsilon^+)$, and $\mathrm{ATCH}_{\vec{w}}(\mathrm{HV}, \epsilon^+)$. These PCUIs are the Pareto-compliant versions of the indicators R2, IGD^+, and ϵ^+. The first experiment investigates the preferences of the adopted PCUIs by measuring the correlation of preferences between them and with their baseline QIs when assessing several MOEAs, having special distribution properties, on the Lamé and Mirror superspheres problems proposed by Em-
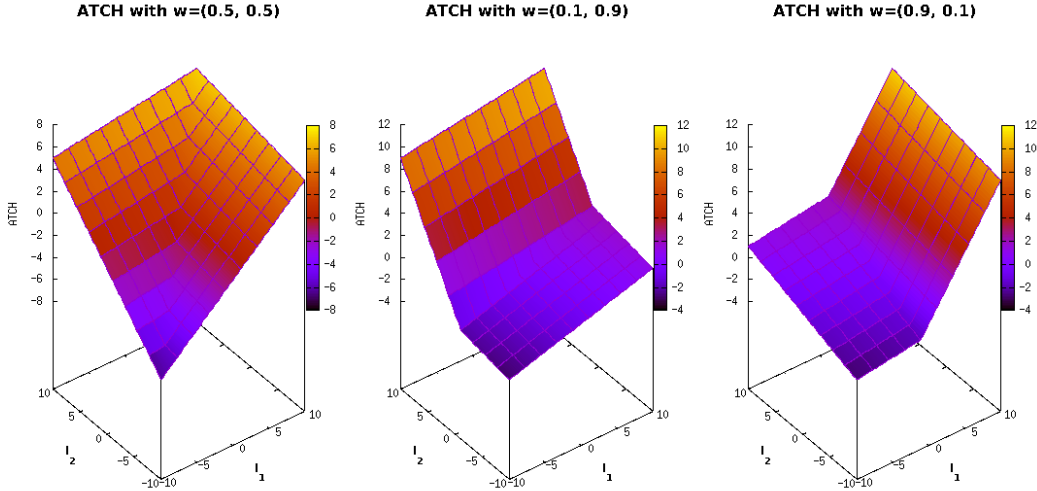
Figure 3: Landscapes of ATCH function varying the weight vector $\vec{w} = (w_1, w_2)$.

merich *et al.* [14]. On the other hand, the second experiment analyzes convergence and distribution properties of a steady-state MOEA, similiar to the S-Metric Selection Evolutionary Multi-Objective Algorithm (SMS-EMOA) [4], that uses the PCUIs as part of its density estimator. The proposed algorithm, denoted as PCUI-EMOA, is tested on MOPs from the benchmarks Deb-Thiele-Laumanns-Zitzler (DTLZ) [13], Walking-Fish-Group (WFG) [22], and their minus versions DTLZ^{-1} and WFG^{-1} [26], respectively.

4.1   Analysis of preferences

We analyzed the correlation of preferences of the six adopted PCUIs when assessing the Pareto fronts of several Lamé and Mirror superspheres problems [14]. The Pareto front geometry of such MOPs is controlled by a parameter γ. Regarding the Lamé problems, when $\gamma \in (0, 1)$, the Pareto front is convex; a linear Pareto front is related to $\gamma = 1$; and when $\gamma \in (1, \infty)$, the Pareto front is concave. In case of the Mirror problems, $\gamma \in (0, 1)$ and $\gamma \in (1, \infty)$ are related to concave and convex Pareto fronts, respectively. For both Lamé and Mirror problems, we employed $\gamma = 0.25, 0.50, 0.75, 1.00, 1.50, 2.00, 6.00$ for 2, 3, and 4 objective functions. Regarding the construction of the Pareto front approximations, we employed MOEAs that exhibit particular distribution
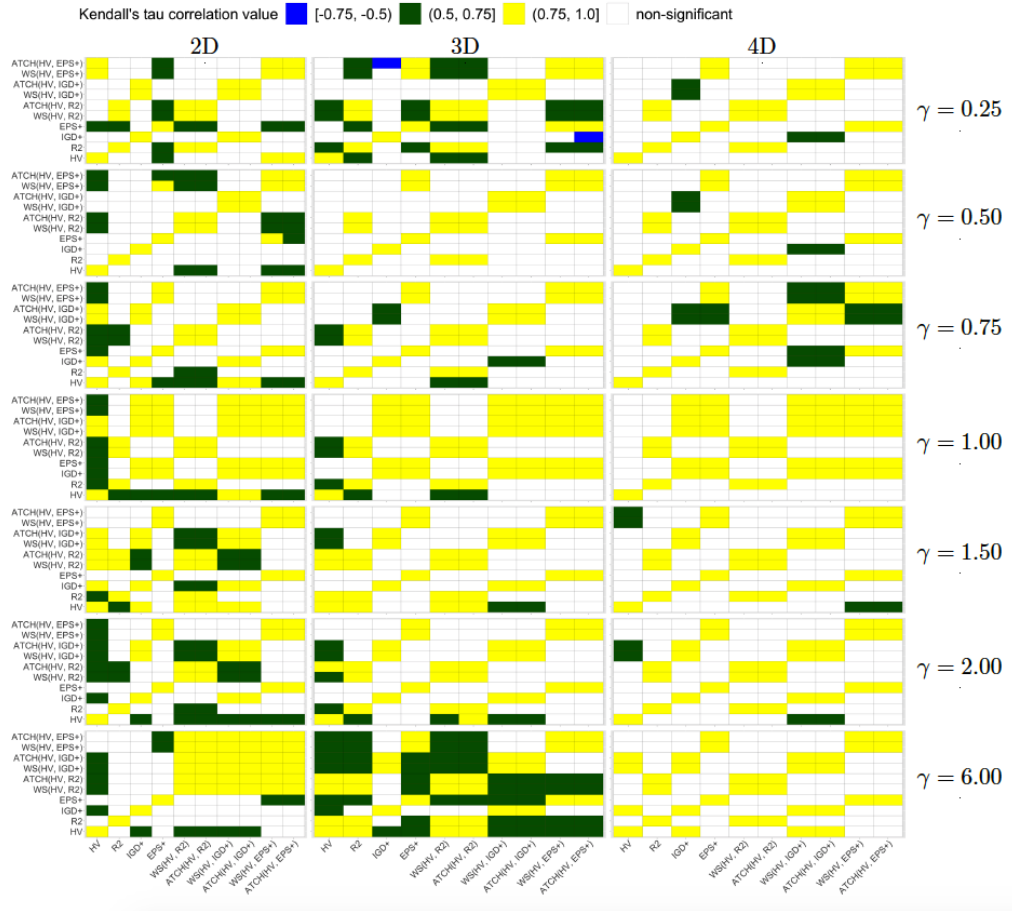
Figure 4: Heatmap Kendall rank correlation τ for each pair of set quality indicators, each Lamé problem on different dimensions of the objective space.
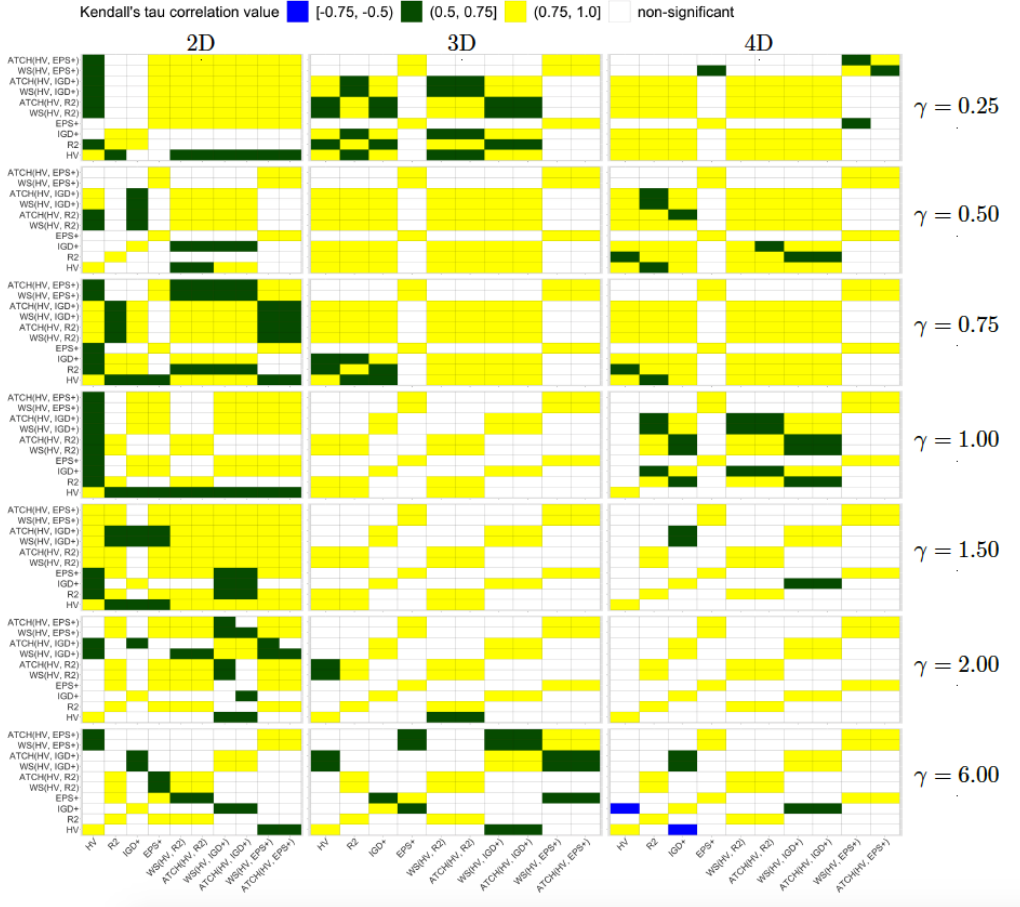
Figure 5: Heatmap Kendall rank correlation τ for each pair of set quality indicators, each Mirror problem on different dimensions of the objective space.

characteristics to have a representative sample of the set Ψ. For each test instance, 30 independent executions were produced by each MOEA, where all algorithms shared the same settings. For all the objective functions, the size of all the produced approximation sets was 120. The adopted MOEAs are classified in five classes as follows:

- Indicator-based MOEAs: SMS-EMOA [4], MOMBI2 [19], IGD$^+$-MaOEA [16] and Δ_p-MaOEA[5].

- Pareto-based MOEAs: NSGA-II [12] and SPEA2 [40].

- Reference set-based MOEAs: NSGA-III [11].

- Decomposition-based MOEAs: MOEA/D [38].

- Image analysis-based MOEAs: MOVAP [20].

Regarding the assessment of the Pareto front approximations generated by the adopted MOEAs, we used the following settings on the QIs and PCUIs. We assumed that we know nothing about the true Pareto fronts of the adopted MOPs to perform a fair quality comparison. Hence, we employed a very bad reference point for HV, i.e., $\vec{z}_{ref} = \{2i + 1\}_{i=1,...,m}$. A set of convex weight vectors (constructed by the Simplex-Lattice-Design method [38]) was employed as the set W for $R2$ and as the reference set for IGD$^+$ and ϵ^+. Since the set of weight vectors is in $[0, 1]^m$, we traslated all the approximation sets to $[1, 2]^m$. Since a PCUI requires all its baseline QIs to be maximized, we consider $-R2$, $-$IGD$^+$, and $-\epsilon^+$. For all the six adopted PCUIs, the weight vector was set as $\vec{w} = (0.1, 0.9)$, where 0.1 is the weight associated with HV and 0.9 is related to the weakly Pareto-compliant QI. This setting was employed to mostly preserve the preferences of the weakly Pareto-compliant QIs while producing Pareto-compliant results due to the use of HV as a correction factor.

Concerning the correlation analysis, we employed a similar methodology to the one adopted by Liefooghe and Derbel [30]. We aim to correlate the rankings of MOEAs within each indicator, i.e., by how much do the PCUIs and QIs rank the MOEAs (i.e., the characteristic Pareto front approximations) similarly. For each test instance and QI, the MOEAs are ranked by

---

[5]We proposed this algorithm based on the framework of IGD$^+$-MaOEA [16] but using the Δ_p indicator as the density estimator.

their mean indicator value. The ranks of MOEAs are then analyzed for correlation with the remaining QIs using the Kendall's τ nonparametric measure of association with a significance value $\alpha = 0.05$. It is worth emphasizing that Kendall's τ quantifies the difference between the proportion of concordant and discordant pairs among all possible pairwise MOEAs. Since $\tau \in [-1, 1]$, where $\tau = -1$ means perfect disagreement and $\tau = 1$ means perfect agreement of ranks, we decided to create intervals of τ values in order to represent them using Heatmaps. Such intervals are the following: $[-1, -0.75), [-0.75, -0.5), [-0.5, -0.25), [-0.25, 0.25], (0.25, 0.5], (0.5, 0.75],$ and $(0.75, 1]$. Figures 4 and 5 show the results of the nonparametric statistical test for the correlation, using heatmaps for all the adopted Lamé and Mirror test instances, respectively. It is worth noting that although we defined seven intervals for τ, the results employed only three of them, i.e., $[-0.75, -0.5), (0.5, 0.75],$ and $(0.75, 1.0]$.

4.1.1 Correlation between PCUIs and baseline QIs

Regarding the correlation analysis on Lamé problems in Fig. 4, we have the following conclusions. For 2 objective functions, the more linear the Pareto front (i.e., $\gamma = 0.75, 1.00, 1.50$), the more correlated are the PCUIs with their baseline QIs. For highly convex Pareto fronts there is more correlation with the weakly Pareto-compliant QI (i.e., R2, IGD$^+$, and ϵ^+) meanwhile for highly concave MOPs, the correlation is stronger with HV. Regarding 3-dimensional MOPs, each class of PCUI shows different behaviors. $\text{WS}_{\vec{w}}(\text{HV}, \text{R2})$ and $\text{ATCH}_{\vec{w}}(\text{HV}, \text{R2})$ are correlated with both HV and R2 for all test problems. For $\text{WS}_{\vec{w}}(\text{HV}, \text{IGD}^+)$ and $\text{ATCH}_{\vec{w}}(\text{HV}, \text{IGD}^+)$, there is independence with both HV and IGD$^+$ for problems with $\gamma = 0.25$ and 0.50 which means that the PCUIs have preferences completely different. For $\gamma = 0.75$ and 1.00, there is only correlation with IGD$^+$ and, in the concave cases, the PCUIs are correlated with both HV and IGD$^+$. Concerning $\text{WS}_{\vec{w}}(\text{HV}, \epsilon^+)$ and $\text{ATCH}_{\vec{w}}(\text{HV}, \epsilon^+)$, both are correlated with ϵ^+ in all cases. A noteworthy aspect is that for 4-dimensional MOPs, there is a strong tendency of all PCUIs to be exclusively correlated with their weakly Pareto-compliant QI while there is independence with HV. Hence, we could expect that as the dimensionality of the objective space increases, the PCUIs will be more correlated with the weakly Pareto-compliant QI. However, the evaluation results of the PCUIs will be Pareto-compliant.

For the Mirror problems in Fig. 5, we have some similar results. In

general, for 2 objective functions, the PCUIs are strongly correlated with both baseline QIs except in highly convex and concave MOPs where their preferences are correlated with either HV or the weakly Pareto-compliant QI. It is worth noting that this result is similar to the Lamé problems having 2 objectives. $\text{WS}_{\vec{w}}(\text{HV}, \text{R2})$ and $\text{ATCH}_{\vec{w}}(\text{HV}, \text{R2})$ are correlated with both QIs in all problems except for $\gamma = 6.00$ for MOPs with 3 objectives and for four-objective MOPs having $\gamma \in \{1.00, 1.50, 2.00, 6.00\}$. On the other hand, $\text{WS}_{\vec{w}}(\text{HV}, \text{IGD}^+)$ and $\text{ATCH}_{\vec{w}}(\text{HV}, \text{IGD}^+)$ are only correlated with both baseline QIs for convex problems and for linear and concave ones, the correlation is stronger with IGD^+ and there is independence with HV in 3- and 4-dimensional problems. Regarding the PCUIs based on HV and ϵ^+, in all cases, there is only correlation with the latter QI in MOPs with 3 and 4 objective functions.

In summary, there are two important points to emphasize. As the dimension of the objective space increases, the PCUIs tend to be strongly correlated with the preferences of their baseline weakly Pareto-compliant QIs and the independence with the preferences of HV gets more accentuated. Since the results of the PCUIs are Pareto-compliant, it is relevant that in high-dimensional objective spaces the PCUIs show preferences independent to those of HV because this could encourage the design of new selection mechanisms of MOEAs that would produce approximation sets with different distributions to those of SMS-EMOA but retaining the Pareto-compliance property. In other words, PCUIs could be employed to manipulate the distribution properties of MOEAs while maintaining the Pareto-compliance property. On the other hand, in general, a PCUI inherits from its baseline weakly Pareto-compliant QI the correlation with HV. A reason for this fact is that we are using a combination vector in the PCUIs that favors the weakly Pareto-compliant QI.

4.1.2 Correlation between PCUIs

We analyzed the correlation between the preferences of all PCUIs to ensure that the combination does produce different indicators. Concerning both the Lamé and Mirror problems, the correlation analysis indicates that the PCUIs based on the same weakly Pareto-compliant QI are strongly correlated between them. In consequence, the use of WS or ATCH is basically producing the same PCUI although they have different landscapes. In the next section, we give the reason for this behavior that, in a few words, is due to the
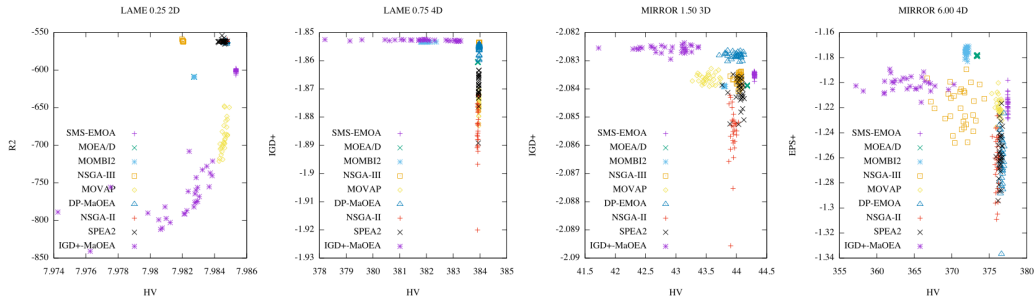
Figure 6: From left to right, it is shown the Quality Spaces: HV-R2 for Lamé $\gamma = 0.25$ 2D, HV-IGD$^+$ for Lamé $\gamma = 0.75$ 4D, HV-IGD$^+$ for Mirror $\gamma = 1.50$ 3D, and HV-ϵ^+ for Mirror $\gamma = 6.00$ 4D. All cases tend to show a Pareto front in Quality Space.

formation of a convex Pareto front in the Quality space. Since both WS and ATCH are able to find solutions on convex Pareto fronts, thus, both will present almost the same preferences when they are employed in PCUIs. As a result of this observation, in this correlation analysis, we investigated the Pareto-compliant versions of the R2, IGD$^+$, and ϵ^+ indicators.

Another remarkable conclusion is that the preferences of PCUIs based on a different weakly Pareto-compliant QI are, in general, independent. Hence, each class of PCUIs are presenting distinct preferences. This is explained by the analysis of the correlation between R2-IGD$^+$, R2-ϵ^+, and IGD$^+$-ϵ^+ that are mostly independent as shown in Figures 4 and 5. Additionally, due to the use of $\vec{w} = (0.1, 0.9)$, each PCUI inherits the preferences of its weakly Pareto-compliant QI. Hence, the PCUI will behave in a similar way to its weakly Pareto-compliant QI but maintaining the Pareto-compliance property.

4.1.3 Pareto fronts in Quality Space

In objective space, we find Pareto fronts that represent the solution to an MOP. These Pareto fronts are formed due to the conflict among objective functions. In Quality Space (see Fig. 2), it is also possible to find Pareto fronts when the preferences of an indicator are in conflict with the preferences of other QI. Based on the correlation analysis previously explained, we found that when there is independence of preferences between two QIs or when the preferences are negatively correlated (as in the case of HV and IGD$^+$ for the Mirror problem with $\gamma = 6.00$ in 4D), a Pareto front in the Quality Space Q

is formed. Fig. 6 shows four examples where it is possible to see the tendency to a Pareto front in quality space. These plots present the indicator vectors associated with each execution of the adopted MOEAs in the correlation study for a specific test instance. Since we are maximizing HV, R2, IGD$^+$, and ϵ^+ to use the PCUIs, it is possible to see that all plots introduce convex Pareto front shapes. Hence, this fact supports the observation that there is no critical difference when using WS or ATCH for constructing PCUIs. The rest of the cases of independence on both heatmaps in Figures 4 and 5 present convex Pareto fronts. In case a PCUI is employed in the selection mechanism of an MOEA, a compromise between the indicators will be found, resulting in new distributions on the Pareto fronts that represent the solution to an MOP. In conclusion, this result supports the fact that PCUIs could be employed to better control the diversity of an MOEA but maintaining the Pareto-compliance.

4.2   Steady-state selection

Algorithm 1 PCUI-EMOA general framework

Require: PCUI $u_{\vec{w}}(\vec{I})$, where $\vec{I}(I_1, I_2, \ldots, I_k)$.
Ensure: Approximation to the Pareto front
 1: Randomly initialize population P
 2: **while** stopping criterion is not fulfilled **do**
 3:     $q \leftarrow Variation(P)$
 4:     $Q \leftarrow P \cup \{q\}$
 5:     $\{R_1, \ldots, R_t\} \leftarrow NondominatedSorting(Q)$
 6:     **if** $|R_t| > 1$ **then**
 7:         $\vec{p}_{\text{worst}} = \arg\min_{\vec{p} \in R_t} \{u_{\vec{w}}(\vec{I}(R_t)) - u_{\vec{w}}(\vec{I}(R_t \setminus \{p\}))\}$
 8:     **else**
 9:         Let $\vec{p}_{\text{worst}}$ be the sole solution in R_t
10:      **end if**
11:      $P \leftarrow Q \setminus \{\vec{p}_{\text{worst}}\}$
12: **end while**
13: **return** P

In this section, we investigate the effect of using PCUIs in the selection mechanism of an MOEA. For this purpose, we considered the framework of SMS-EMOA that uses a density estimator (DE) based on HV but, in our case, a PCUI is employed in the DE. Algorithm 1 presents the general framework of our proposed PCUI-EMOA whose main loop is in lines 2 to 12. At each generation a new solution is created using genetic operators and, then, this newly created solution is added to the population P to create the temporary population Q. Then, in line 5, a set of ranks $R_1, \ldots, R_t$ are

created using the nondominated sorting algorithm [12], where R_t has the worst solutions according to the Pareto dominance relation. If R_t has more than one solution, the individual contributions to the PCUI are computed to delete the worst-contributing solution in line 11. Finally, the Pareto front approximation is returned when the stopping criterion is fulfilled.

We focused our attention on studying the final distribution properties of PCUI-EMOA in comparison with four steady-state MOEAs based on the indicators HV, R2, IGD$^+$, and ϵ^+, i.e., SMS-EMOA, R2-EMOA, IGD$^+$-MaOEA, and ϵ^+-MaOEA. The latter is similar to IGD$^+$-MaOEA. Regarding PCUI-EMOA, we employed the six PCUIs of the previous section. Since all the adopted indicator-based MOEAs (IB-MOEAs) share the same structure, the parameter settings are the following. For all objective functions, the population size is 120. All MOEAs use simulated binary crossover and polynomial-based mutation as their genetic operators [12], where, for all cases, the crossover probability is set to 0.9, the mutation probability is $1/n$ (n is the number of decision variables), and both the crossover and mutation distribution indexes are set to 20. PCUI-EMOA employs the combination vector as $\vec{w} = (0.5, 0.5)$ to look for the knee point on the Pareto front in quality space, i.e., to generate distributions similar to both baseline QIs. We tested the adopted MOEAs on 14 MOPs from the benchmarks DTLZ, WFG, DTLZ^{-1}, and WFG^{-1} for 2, 3 and 4 objective functions. We employed the problems DTLZ1, DTLZ2, DTLZ5, DTLZ7, WFG1, WFG2, WFG3, and their minus versions. We selected these MOPs since they possess Pareto fronts with different geometries, namely, linear, concave, convex, degenerate, mixed, disconnected, correlated with the simplex shape and not correlated with it [26].

The distribution analysis is focused on determining if the Pareto front approximations produced by the six PCUI-EMOAs are similar to the IB-MOEAs that use their baseline indicators. For each test instance, the MOEAs were executed $N = 30$ independent times. Thus, each one produced N approximation sets for each MOP. We investigate the similarity between two sets of approximation sets produced by two MOEAs, using a similarity measure based on the Hausdorff distance that we propose in the following:

Definition 12 (Hausdorff similarity measure). *Given two sets $\mathbb{A} = \{A_1, \ldots, A_N\}$ and $\mathbb{B} = \{B_1, \ldots, B_N\}$, each one consisting of N Pareto front approximations, the Hausdorff similarity measure S is given as follows:*

$$S(\mathbb{A}, \mathbb{B}) = \frac{1}{N} \sum_{i=1}^{N} median(A_i, \mathbb{B}), \tag{7}$$

where $median(A_i, \mathbb{B})$ computes all the Hausdorff distances from A_i to every element in $\mathbb{B}$ and returns the median value.

S calculates the degree of similarity between two sets. However, if we are given three sets of approximation sets $\mathbb{A}, \mathbb{B}$, and $\mathbb{C}$ and we would like to know if $\mathbb{A}$ is similar to $\mathbb{B}$, to $\mathbb{C}$, to both or to none of them, a classification function is required. Such classifier is given as follows.

Definition 13 (Classifier). *Given three sets of approximation sets $\mathbb{A}, \mathbb{B}$, and $\mathbb{C}$ and a threshold $\epsilon > 0$, the classifier function is given as follows:*

$$C_\epsilon(S(\mathbb{A}, \mathbb{B}), S(\mathbb{A}, \mathbb{C})) = \begin{cases} -1, & S(\mathbb{A}, \mathbb{B}) \leq \epsilon \wedge S(\mathbb{A}, \mathbb{C}) > \epsilon \\ 0, & S(\mathbb{A}, \mathbb{B}) \leq \epsilon \wedge S(\mathbb{A}, \mathbb{C}) \leq \epsilon \\ 1, & S(\mathbb{A}, \mathbb{B}) > \epsilon \wedge S(\mathbb{A}, \mathbb{C}) > \epsilon \\ 2, & S(\mathbb{A}, \mathbb{C}) \leq \epsilon \wedge S(\mathbb{A}, \mathbb{B}) > \epsilon \end{cases}$$

where -1 means that $\mathbb{A}$ is exclusively similar to $\mathbb{B}$; 0 means that $\mathbb{A}$ is similar to both $\mathbb{B}$ and $\mathbb{C}$; 1 means that $\mathbb{A}$ is not similar to $\mathbb{B}$ nor $\mathbb{C}$; and, 2 means that $\mathbb{A}$ is exclusively similar to $\mathbb{C}$.

Based on the classification function, we analyzed the similarities between the approximation sets produced by the PCUI-EMOAs and their corresponding IB-MOEAs that use the baseline indicators for the construction of the PCUI. Table 2 shows the results for all the considered test instances using $\epsilon = 0.1$. Since all PCUI-EMOAs use $\vec{w} = (0.5, 0.5)$ as the combination weight vector for the order-preserving utility functions, our hypothesis is that the Pareto front approximations should be similar to both IB-MOEAs that employ the baseline indicators. This hypothesis is true for several cases related to the DTLZ and DTLZ^{-1} problems. Nevertheless, for most of the WFG and WFG^{-1} problems, the PCUI-EMOA tends to produce approximation sets with particular distributions that are not similar to the baseline IB-MOEAs. This fact could be explained by the independence of preferences between HV and the weakly Pareto-compliant indicators on these MOPs. Considering the linear problems DTLZ1 and DTLZ1^{-1}, it is clear that in most cases the PCUI-EMOAs produce approximation sets similar to the IB-MOEAs using

their baseline indicators. This result is explained by the correlation analysis of Section 4.1 where in almost all cases HV, R2, IGD$^+$, and ϵ^+ are strongly correlated. The most important observation is related to the PCUI-EMOAs based on $\text{WS}_{\vec{w}}(\text{HV}, \text{R2})$ and $\text{ATCH}_{\vec{w}}(\text{HV}, \text{R2})$. On the one hand, SMS-EMOA produces uniformly distributed solutions in convex and linear Pareto fronts and there is a bias towards the knee and boundaries of concave Pareto fronts. Additionally, SMS-EMOA presents good results in degenerate problems such as DTLZ5 and WFG3. On the other hand, R2-EMOA Brockhoff *et al.* [8] does not produce uniformly distributed solutions in convex Pareto fronts, but it does in linear and concave ones. Regarding, degenerate MOPs, R2-EMOA does not produce good results since its weight vectors do not completely intersect the Pareto front shape. Hence, SMS-EMOA and R2-EMOA have specific strengths and weaknesses depending on the MOP being tackled. Regarding DTLZ2 having two and three objective functions and concave Pareto fronts, it is possible to see that the distribution of the PCUI-EMOAs based on $\text{WS}_{\vec{w}}(\text{HV}, \text{R2})$ and $\text{ATCH}_{\vec{w}}(\text{HV}, \text{R2})$ are similar to the preferences of R2, i.e., R2-EMOA. When we analyze the minus version DTLZ2^{-1} for the same objective funtions, the distributions are similar to those of SMS-EMOA. This also happens for DTLZ5 3D which is degenerate where the distributions are similar to those of SMS-EMOA as well. Hence, we have empirical evidence on the compensation of weaknesses of one indicator with the strengths of the other baseline indicator when employing PCUI-EMOA. Fig 7 shows some examples of this remarkable compensation.

5   Conclusions and Future Work

In this paper, we proposed to construct new Pareto-compliant indicators by combining existing QIs, under specific conditions. To the authors' best knowledge, this is the first work that proposes such a combination to obtain Pareto-compliant QIs. To ensure the Pareto-compliance property, it is mandatory to combine at least one Pareto-compliant indicator (such as the hypervolume indicator) with as many weakly Pareto-compliant QIs, using an order-preserving function. Regarding these functions, we proposed to use the weighted-sum and the augmented Tchebycheff utility functions. Based on these utility functions, we denoted the combined indicators as Pareto-compliant Utility Indicators (PCUIs). As part of our experimental results, we analyzed the preferences of six PCUIs based on the combination of the hy-

Table 2: Distribution similarities between each PCUI-EMOA and the IB-MOEAs based on the indicators HV, R2, IGD$^+$ and ϵ^+. For each test instance, it is shown if the distribution of the PCUI-EMOA is similar to one or other baseline indicator, to both or none of them.

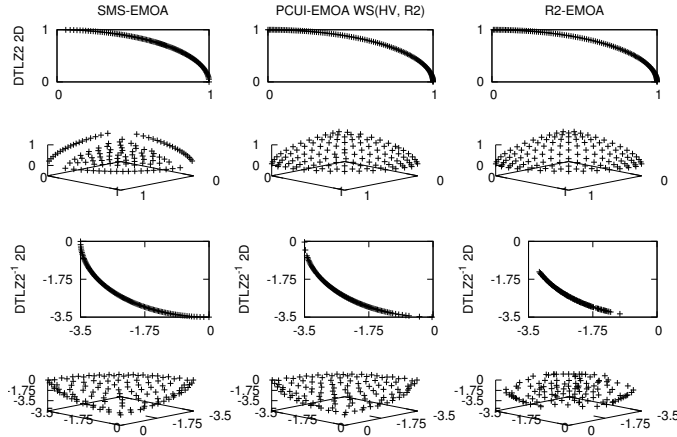| MOP	Dim.	WS$_{\vec{w}}$(HV, R2)	ATCH$_{\vec{w}}$(HV, R2)	WS$_{\vec{w}}$(HV, IGD$^+$)	ATCH$_{\vec{w}}$(HV, IGD$^+$)	WS$_{\vec{w}}$(HV, ϵ^+)	ATCH$_{\vec{w}}$(HVϵ^+)
DTLZ1	2	Both	Both	Both	Both	Both	Both
	3	Both	Both	Both	Both	Both	Both
	4	Both	Both	Both	Both	Both	Both
DTLZ1^{-1}	2	Both	Both	Both	Both	Both	Both
	3	Both	Both	Both	Both	Both	Both
	4	HV	HV	Both	Both	HV	None
DTLZ2	2	Both	Both	Both	Both	Both	Both
	3	R2	R2	Both	Both	Both	Both
	4	R2	R2	None	None	None	None
DTLZ2^{-1}	2	HV	HV	Both	Both	Both	Both
	3	HV	HV	None	None	None	None
	4	None	None	None	None	None	None
DTLZ5	2	R2	R2	Both	Both	Both	Both
	3	HV	HV	Both	Both	Both	Both
	4	HV	HV	IGD$^+$	IGD$^+$	ϵ^+	ϵ^+
DTLZ5^{-1}	2	HV	HV	Both	Both	Both	Both
	3	HV	HV	IGD$^+$	IGD$^+$	None	None
	4	None	None	None	None	None	None
DTLZ7	2	Both	Both	Both	Both	Both	Both
	3	None	None	IGD$^+$	IGD$^+$	ϵ^+	None
	4	None	None	None	None	None	None
DTLZ7^{-1}	2	Both	Both	Both	Both	Both	Both
	3	R2	R2	Both	Both	Both	Both
	4	None	None	IGD$^+$	None	ϵ^+	Both
WFG1	2	None	None	None	None	None	None
	3	None	None	None	None	None	ϵ^+
	4	None	None	IGD$^+$	IGD$^+$	ϵ^+	None
WFG1^{-1}	2	R2	R2	None	None	None	None
	3	None	None	None	None	None	None
	4	None	None	None	None	None	None
WFG2	2	None	None	None	None	None	None
	3	None	None	None	None	None	None
	4	None	None	None	None	None	None
WFG2^{-1}	2	None	None	None	None	None	None
	3	R2	R2	IGD$^+$	IGD$^+$	ϵ^+	ϵ^+
	4	None	None	None	None	None	None
WFG3	2	None	None	None	None	None	None
	3	Both	Both	Both	Both	Both	Both
	4	None	None	IGD$^+$	IGD$^+$	ϵ^+	ϵ^+
WFG3^{-1}	2	R2	R2	None	None	None	None
	3	Both	Both	Both	Both	Both	Both
	4	HV	HV	Both	Both	HV	HV

Figure 7: Pareto fronts that show the compensation of weaknesses of one indicator with the strengths of other when coupled to PCUI-EMOA.

pervolume indicator with the R2, IGD$^+$, and ϵ^+ indicators which are weakly Pareto-compliant. The results indicated that the PCUIs show different preferences to those of the hypervolume indicator which implies that PCUIs can be used as an alternative to HV while ensuring the Pareto-compliance property. Additionally, we analyzed the distribution properties of an MOEA based on the PCUIs, denoted as PCUI-EMOA. We empirically showed that when solving MOPs, PCUI-EMOA is able to compensate for the weaknesses of one of its baseline indicators with the strengths of the other. As part of our future work, we want to use other weakly Pareto-compliant indicators to define Pareto-compliant ones and we look for making a more in-depth theoretical analysis of the properties of PCUIs.

Acknowledgements

References

[1] Anne Auger, Johannes Bader, Dimo Brockhoff, and Eckart Zitzler. Theory of the Hypervolume Indicator: Optimal $\{\mu\}$-Distributions and the Choice of the Reference Point. In *FOGA '09: Proceedings of the tenth ACM SIGEVO workshop on Foundations of genetic algorithms*, pages 87–102, Orlando, Florida, USA, January 2009. ACM.

[2] Rudolf Berghammer, Tobias Friedrich, and Frank Neumann. Convergence of set-based multi-objective optimization, indicators and deteriorative cycles. *Theoretical Computer Science*, 456:2–17, October 19 2012.

[3] Nicola Beume, Carlos M. Fonseca, Manuel Lopez-Ibanez, Luis Paquete, and Jan Vahrenhold. On the Complexity of Computing the Hypervolume Indicator. *IEEE Transactions on Evolutionary Computation*, 13(5):1075–1082, October 2009.

[4] 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.

[5] Leonardo C. T. Bezerra, Manuel López-Ibáñez, and Thomas Stützle. An empirical assessment of the properties of inverted generational distance on multi- and many-objective optimization. In Heike Trautmann, Günter Rudolph, Kathrin Klamroth, Oliver Schütze, Margaret Wiecek, Yaochu Jin, and Christian Grimme, editors, *Evolutionary Multi-Criterion Optimization: 9th International Conference, EMO 2017, Münster, Germany, March 19-22, 2017, Proceedings*, pages 31–45, Cham, 2017. Springer International Publishing.

[6] Karl Bringmann, Tobias Friedrich, Frank Neumann, and Markus Wagner. Approximation-Guided Evolutionary Multi-Objective Optimization. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2011)*, pages 1198–1203, Barcelona, Spain, 16-22 July 2011. AAAI Press.

[7] Dimo Brockhoff, Tobias Wagner, and Heike Trautmann. On the Properties of the $R2$ Indicator. In *2012 Genetic and Evolutionary Computation Conference (GECCO'2012)*, pages 465–472, Philadelphia, USA, July 2012. ACM Press. ISBN: 978-1-4503-1177-9.

[8] Dimo Brockhoff, Tobias Wagner, and Heike Trautmann. R2 Indicator-Based Multiobjective Search. *Evolutionary Computation*, 23(3):369–395, Fall 2015.

[9] Carlos A. Coello Coello and Nareli Cruz Cortés. Solving Multiobjective Optimization Problems using an Artificial Immune System. *Genetic Programming and Evolvable Machines*, 6(2):163–190, June 2005.

[10] 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.

[11] 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):577–601, August 2014.

[12] 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.

[13] 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.

[14] Michael T.M. Emmerich and André H. Deutz. Test Problems Based on Lamé Superspheres. In Shigeru Obayashi, Kalyanmoy Deb, Carlo Poloni, Tomoyuki Hiroyasu, and Tadahiko Murata, editors, *Evolutionary Multi-Criterion Optimization, 4th International Conference, EMO 2007*, pages 922–936, Matshushima, Japan, March 2007. Springer. Lecture Notes in Computer Science Vol. 4403.

[15] Michael T.M. Emmerich, Andr H. Deutz, and Iryna Yevseyeva. On reference point free weighted hypervolume indicators based on desirability functions and their probabilistic interpretation. *Procedia Technology*, 16:532 – 541, 2014. CENTERIS 2014 - Conference on ENTERprise Information Systems / ProjMAN 2014 - International Conference on Project MANagement / HCIST 2014 - International Conference on Health and Social Care Information Systems and Technologies.

[16] Jesús Guillermo Falcón-Cardona and Carlos A. Coello Coello. Towards a More General Many-Objective Evolutionary Optimizer. In *2018 Parallel Problem Solving from Nature (PPSN XV)*, Coimbra, Portugal, September 7–12 2018. ACM Press.

[17] Tobias Friedrich, Karl Bringmann, Thomas Voß, and Christian Igel. The Logarithmic Hypervolume Indicator. In Hans-Georg Beyer and William B. Langdon, editors, *Proceedings of the 2011 ACM/SIGEVO Foundations of Genetic Algorithms XI (FOGA'2011)*, pages 81–92. ACM Press, Schwarzenberg, Austria, January 5–9 2011.

[18] Michael Pilegaard Hansen and Andrzej Jaszkiewicz. Evaluating the quality of approximations to the non-dominated set. Technical Report IMM-REP-1998-7, Technical University of Denmark, March 1998.

[19] 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)*, pages 679–686, Madrid, Spain, July 11-15 2015. ACM Press. ISBN 978-1-4503-3472-3.

[20] Raquel Hernández Gómez, Carlos A. Coello Coello, and Enrique Alba Torres. A Multi-Objective Evolutionary Algorithm based on Parallel Coordinates. In *2016 Genetic and Evolutionary Computation Conference (GECCO'2016)*, pages 565–572, Denver, Colorado, USA, 20-24 July 2016. ACM Press. ISBN 978-1-4503-4206-3.

[21] Jeffrey Horn, Nicholas Nafpliotis, and David E. Goldberg. A Niched Pareto Genetic Algorithm for Multiobjective Optimization. In *Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, volume 1, pages 82–87, Piscataway, New Jersey, June 1994. IEEE Service Center.

[22] 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.

[23] Hisao Ishibuchi, Ryo Imada, Yu Setoguchi, and Yusuke Nojima. Reference Point Specification in Hypervolume Calculation for Fair Comparison and Efficient Search. In *2017 Genetic and Evolutionary Computation Conference (GECCO'2017)*, pages 585–592, Berlin, Germany, July 15-19 2017. ACM Press. ISBN 978-1-4503-4920-8.

[24] Hisao Ishibuchi, Hiroyuki Masuda, Yuki Tanigaki, and Yusuke Nojima. Difficulties in Specifying Reference Points to Calculate the Inverted Generational Distance for Many-Objective Optimization Problems. In *2014 IEEE Symposium on Computational Intelligence in Multi-Criteria Decision-Making*

(MCDM'2014), pages 170–177, Orlando, Florida, USA, 9-12 December 2014. IEEE Press. ISBN 978-1-4799-4467-5.

[25] 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.

[26] 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.

[27] Siwei Jiang, Yew-Soon Ong, Jie Zhang, and Liang Feng. Consistencies and Contradictions of Performance Metrics in Multiobjective Optimization. *IEEE Transactions on Cybernetics*, 44(12):2391–2404, December 2014.

[28] Fei Li, Ran Cheng, Jianchang Liu, and Yaochu Jin. A Two-Stage R2 Indicator Based Evolutionary Algorithm for Many-Objective Optimization. *Applied Soft Computing*, 67:245–260, June 2018.

[29] 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.

[30] Arnaud Liefooghe and Bilel Derbel. A Correlation Analysis of Set Quality Indicator Values in Multiobjective Optimization. In *2016 Genetic and Evolutionary Computation Conference (GECCO'2016)*, pages 581–588, Denver, Colorado, USA, 20-24 July 2016. ACM Press. ISBN 978-1-4503-4206-3.

[31] Kaisa Miettinen. *Nonlinear Multobjective Optimization*. Kluwer Academic Publishers, Boston, 1999.

[32] 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 Heike Trautmann, Günter Rudolph, Kathrin Klamroth, Oliver Schütze, Margaret Wiecek, Yaochu Jin, and Christian Grimme, editors, *Evolutionary Multi-Criterion Optimization, 9th International Conference, EMO 2017*, pages 499–513. Springer. Lecture Notes in Computer Science Vol. 10173, Münster, Germany, March 19-22 2017. ISBN 978-3-319-54156-3.

[33] Oliver Schütze, Xavier Esquivel, Adriana Lara, and Carlos A. Coello Coello. Using the Averaged Hausdorff Distance as a Performance Measure in Evolutionary Multiobjective Optimization. *IEEE Transactions on Evolutionary Computation*, 16(4):504–522, August 2012.

[34] 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.

[35] David A. Van Veldhuizen. *Multiobjective Evolutionary Algorithms: Classifications, Analyses, and New Innovations.* PhD thesis, Department of Electrical and Computer Engineering. Graduate School of Engineering. Air Force Institute of Technology, Wright-Patterson AFB, Ohio, USA, May 1999.

[36] Christian von Lücken, Benjamin Baran, and Carlos Brizuela. A survey on multi-objective evolutionary algorithms for many-objective problems. *Computational Optimization and Applications*, 58(3):707–756, July 2014.

[37] Yuan Yuan, Hua Xu, Bo Wang, and Xin Yao. A New Dominance Relation-Based Evolutionary Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 20(1):16–37, February 2016.

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

[39] Eckart Zitzler and Simon Künzli. Indicator-based Selection in Multiobjective Search. In Xin Yao et al., editor, *Parallel Problem Solving from Nature - PPSN VIII*, pages 832–842, Birmingham, UK, September 2004. Springer-Verlag. Lecture Notes in Computer Science Vol. 3242.

[40] Eckart Zitzler, Marco Laumanns, and Lothar Thiele. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. In K. Giannakoglou, D. Tsahalis, J. Periaux, P. Papailou, and T. Fogarty, editors, *EUROGEN 2001. Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems*, pages 95–100, Athens, Greece, 2001.

[41] 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.

[42] Eckart Zitzler, Lothar Thiele, and Johannes Bader. On Set-Based Multi-objective Optimization. *IEEE Transactions on Evolutionary Computation*, 14(1):58–79, February 2010.

[43] 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.