Les Cahiers du GERAD ISSN: 0711–2440

Performance indicators in multiobjective optimization

C. Audet, J. Bigeon, D. Cartier,
S. Le Digabel, L. Salomon

G–2018–90

October 2018

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2018
– Bibliothèque et Archives Canada, 2018

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2018
– Library and Archives Canada, 2018

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Performance indicators in multiobjective optimization

Charles Audet [a]
Jean Bigeon [b]
Dominique Cartier [c]
Sébastien Le Digabel [a]
Ludovic Salomon [a,b]

[a] *GERAD & Department of Mathematics and Industrial Engineering, Polytechnique Montréal (Québec) Canada, H3C 3A7*

[b] *Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France*

[c] *Collège de Maisonneuve, Montréal (Québec) Canada, H1X 2A2*

Charles.Audet@gerad.ca
jean.bigeon@grenoble-inp.fr
dcartier@cmaisonneuve.qc.ca
Sebastien.Le.Digabel@gerad.ca
ludovic.salomon@gerad.ca

October 2018
Les Cahiers du GERAD
G–2018–90

Abstract: In the recent years, the development of new algorithms for multiobjective optimization has considerably grown. A large number of performance indicators has been introduced to measure the quality of Pareto fronts approximations produced by these algorithms. In this work, we propose a review of a total of 57 performance indicators partitioned into four groups according to their properties: cardinality, convergence, distribution and spread. Applications of these indicators are presented as well.

Keywords: multiobjective optimization, quality indicators, performance indicators

1 Introduction

Since the eighties, a large number of methods has been developed to treat multiobjective optimization problems (e.g [1, 2, 3, 4, 5]). Given that conflicting objectives are provided, the set of solutions, the *Pareto front*, is described as the set of best trade-off points in the objective space. Knowledge of the Pareto front enables the decision maker to visualize the consequences of his/her choices in terms of performance for a criterion at the expense of one or other criteria, and to make appropriate decisions.

Formally, a feasible vector x is said to *(Pareto)-dominate* another feasible vector x' if x is at least as good as x' for all the objectives, and strictly better than x' for at least one objective. The decision vectors in the feasible set that are not dominated by any other feasible vector are called *Pareto optimal*. The set of non-dominated points in the feasible is the set of *Pareto solutions*, whose images (by the objective functions) constitute the *Pareto front*.

In single-objective minimization, the quality of a given solution is trivial to quantify: the smaller the objective function value, the better. However, evaluating the quality of an approximation of a Pareto set is non trivial. The question is important for the comparison of algorithms, the definition of stopping criteria, or even the design of better methods. According to [6], a Pareto front approximation should satisfy the following:

- The distance between the Pareto front and its approximation should be minimized.
- A good (according to some metric) distribution of the points of the approximated front is desirable.
- The extent of the approximated front should be maximized, i.e., for each objective, a wide range of values should be covered by the non-dominated points.

To answer this question, many metrics called *performance indicators* [7, 8] have been introduced. Performance indicators can be considered as mappings that assign scores to Pareto front approximations.

Surveys of performance indicators already exist. In [2, chapter 7], the authors list some performance indicators to measure the quality of a Pareto front approximation. In [7], an exhaustive survey is conducted on a vast number of performance indicators which are grouped according to their properties. Mathematical frameworks to evaluate performance indicators are proposed in [9, 10] and additional metrics and algorithms are listed in [11]. In [12], the authors review some performance indicators and analyze their drawbacks. In [13], an empirical study focuses on the correlations between different indicators with their computation time on concave and convex Pareto fronts. Finally, the usage of indicators proposed by the multiobjective evolutionary optimization community prior to 2013 is analyzed in [14].

Table 1 provides a panorama of existing indicators, classifies them based on their properties, and indicates the section in which they are discussed. The use of performance metrics targets three cases: comparison of algorithms, suggestion of stopping criteria for multiobjective optimization and identification of promising performance indicators to evaluate and improve the distribution of the points for a given solution.

This work is organized as follows. Section 2 introduces the notations and definitions related to multiobjective optimization and quality indicators. Section 3 is the core of this work, and is devoted to classification of the indicators according to their specific properties. Finally, Section 4 presents some applications.

Table 1: A summary of performance indicators

Category	Performance indicators	Sect.	[12]	[2]	[9]	[10]	[7]	[11]	[13]	[14]
Cardinality 3.1	C-metric/Two sets Coverage [15]	3.1.5	✓	✓	✓		✓		✓	✓
	Error ratio [16]	3.1.4		✓	✓	✓	✓		✓	
	Generational non dominated vector generation [17]	3.1.3		✓			✓		✓	✓
	Generational non dominated vector generation ratio [17]	3.1.3					✓		✓	✓
	Mutual domination rate [18]	3.1.6								✓
	Nondominated vector additional [17]	3.1.3		✓			✓		✓	
	Overall nondominated vector generation [16]	3.1.1		✓	✓	✓	✓	✓	✓	✓
	Overall nondominated vector generation ratio [16]	3.1.2		✓	✓	✓	✓		✓	✓
	Ratio of non-dominated points by the reference set [19]	3.1.5					✓		✓	
	Ratio of the reference points [19]	3.1.4					✓		✓	

Table 1 – continued from previous page

Category	Performance indicators	Sect.	[12]	[2]	[9]	[10]	[7]	[11]	[13]	[14]
Convergence 3.2	Averaged Hausdorff distance [20]	3.2.6							✓	
	Degree of Approximation [21]	3.2.10								
	D_R-metric [19]	3.2.1			✓	✓	✓			✓
	ϵ-family [10]	3.2.9				✓			✓	✓
	Generational distance [16]	3.2.1		✓	✓	✓	✓	✓	✓	✓
	γ-metric [22]	3.2.1	✓	✓			✓		✓	✓
	Inverted generational distance [23]	3.2.5	✓						✓	✓
	Maximum Pareto front error [16]	3.2.4		✓	✓	✓	✓	✓	✓	
	$M_1^\star$-metric [6]	3.2.1		✓		✓	✓		✓	✓
	Modified inverted generational distance [24]	3.2.7								
	Progression metric [16]	3.2.8		✓						
	Seven points average distance [25]	3.2.3					✓		✓	
	Standard deviation from the Generational distance [16]	3.2.2		✓						
Distribution and spread 3.3	Cluster [26]	3.3.17				✓	✓	✓	✓	✓
	Δ-index [22]	3.3.2	✓				✓	✓		✓
	Δ'-index [22]	3.3.2					✓		✓	✓
	$\Delta^\star$ spread metric [27]	3.3.2						✓	✓	✓
	Distribution metric [28]	3.3.12								
	Diversity comparison indicator [29]	3.3.17								
	Diversity indicator [30]	3.3.15								
	Entropy metric [31]	3.3.17					✓	✓	✓	✓
	Evenness [32]	3.3.7						✓		
	Extension [33]	3.3.14						✓		
	Γ-metric [34]	3.3.3				✓				
	Hole Relative Size [2]	3.3.4		✓		✓		✓		
	Laumanns metric [35]	3.3.16		✓						
	Modified Diversity indicator [36]	3.3.17								
	$M_2^\star$-metric [6]	3.3.5		✓		✓	✓	✓	✓	
	$M_3^\star$-metric [6]	3.3.5	✓	✓		✓	✓	✓	✓	✓
	Number of distinct choices [26]	3.3.17				✓	✓	✓	✓	
	Outer diameter [8]	3.3.11				✓				
	Overall Pareto Spread [26]	3.3.10				✓	✓	✓	✓	✓
	Sigma diversity metric [37]	3.3.17							✓	
	Spacing [25]	3.3.1		✓	✓	✓	✓	✓	✓	✓
	U-measure [38]	3.3.9						✓		
	Uniform assessment metric [39]	3.3.13								
	Uniform distribution [40]	3.3.5					✓		✓	
	Uniformity [41]	3.3.6						✓		
	Uniformity [33]	3.3.8						✓		
Convergence and distribution 3.4	Cone-based hypervolume [42]	3.4.4								
	Dominance move [43]	3.4.3								
	D-metric/Difference coverage of two sets [44]	3.4.4	✓			✓	✓		✓	✓
	Hyperarea difference [26]	3.4.4					✓	✓	✓	✓
	Hypervolume indicator (or S-metric) [6]	3.4.4	✓	✓	✓	✓	✓	✓	✓	✓
	G-metric [45]	3.4.2								
	Logarithmic hypervolume indicator [46]	3.4.4								
	R-metric [19]	3.4.1				✓	✓	✓	✓	✓

2 Notations and definitions

To apprehend quality indicators, the first part of this section describes the main concepts related to multiobjective optimization. The second part focuses on the theory of Pareto set approximations and quality indicators.

2.1 Multiobjective optimization and Pareto dominance

We consider the following continuous multiobjective optimization problem:

$$\min_{x\in\Omega} F(x) = [f_1(x)\ f_2(x)\dots f_m(x)]^\top$$

where $\Omega \subset \mathbb{R}^n$ is called the *feasible set*, and $f_i : \mathbb{R}^n \to \mathbb{R}$ are m objective functions for $i = 1, 2, \dots, m$, with $m \geq 2$. The image of the feasible set $\mathcal{F} = \{F(x) \in \mathbb{R}^m : x \in \Omega\}$ is called the *objective space*.

The following cone order relation is adopted [47]: given two vectors z and z' in the objective space $\mathcal{F}$, we have

$$z \leq z' \iff z' - z \in \mathbb{R}^m_+ \iff z_i \leq z'_i, \text{ for all } i = 1, 2, \dots, m.$$

In a similar way, we define the strict order relation $<$ in the objective space. We can now present the concept of dominance.

Definition 1 (Dominance relations) *Given two decision vectors x and x' in Ω, we write:*

- *$x \preceq x'$ (x weakly dominates x') if and only if $F(x) \leq F(x')$.*
- *$x \prec x'$ (x dominates x') if and only if $x \preceq x'$ and at least one component of $F(x)$ is strictly less than the corresponding one of $F(x')$.*
- *$x \prec\prec x'$ (x strictly dominates x') if and only if $F(v) < F(v')$.*
- *$x \parallel x'$ (x and x' are incomparable) if neither x weakly dominates x' nor x' weakly dominates x.*

With these relations, we now precise the concept of solution in the multiobjective optimization framework.

Definition 2 (Pareto optimality and Pareto solutions) *The vector $x \in \Omega$ is a* Pareto-optimal *solution if there is no other vector in Ω that dominates it. The set of Pareto-optimal solutions is called the* Pareto set*, denoted $\mathcal{X}_P$, and the image of the Pareto set is called the* Pareto front*, denoted $\partial\mathcal{F}$.*

In single-objective optimization, the set of optimal solutions is often composed of a singleton. In the multiobjective case, the Pareto front usually contains many elements (an infinity in continuous optimization and an exponential number in discrete optimization [47]). For a problem with m objectives, $\partial\mathcal{F}$ is of dimension $m-1$ or less. For example, with two objectives, $\partial\mathcal{F}$ is a curve, for three objectives, $\partial\mathcal{F}$ is a surface, and so on. Also, it is interesting to define some bounds on this set.

Definition 3 (Ideal and nadir points) *The* ideal point F^I *[2] is defined as the vector whose components are the solutions of each single-objective problem $\min_{x\in\Omega} f_i(x),\ i = 1, 2, \dots, m$. The* nadir point F^N *is defined as the vector whose components are the solutions of the single-objective problems $\max_{x\in\mathcal{X}_P} f_i(x), i = 1, 2, \dots, m$.*

For computation reasons, the nadir point is often approximated by $\widetilde{F}^N$ for which the coordinates are defined the following way: let $x_i^\star$ be the solution of the single-objective problem $\min_{x\in\Omega} f_i(x)$ for $i = 1, 2, \dots, m$. The ith coordinate of $\widetilde{F}^N$ is given by:

$$\widetilde{F}_i^N = \max_{k=1,2,\dots,m} f_i(x_k^\star).$$

For a biobjective optimization problem, F^N equals $\widetilde{F}^N$. It is not always the case when $m \geq 3$.

An illustration is given in Figure 1 where the Pareto front is piecewise continuous. To simplify the notation, continuous Pareto and piecewise continuous Pareto fronts will be respectively designed as continuous and discontinuous Pareto fronts.

Remark 1 *In a multiobjective optimization problem, objectives are not necessarily contradictory, and the set of Pareto solutions may be a singleton. In this study, we assume that this is not the case.*

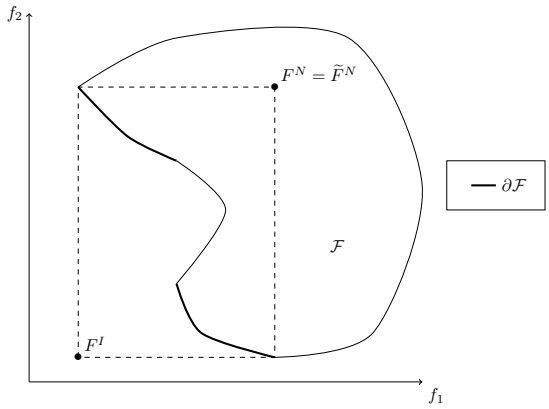

Figure 1: Objective space, ideal point and nadir point (inspired by [47])

2.2 Approximation sets and performance indicators

Generally, whether in the context of continuous or discrete optimization, it is not possible to find or enumerate all elements of the Pareto front. Hence to solve a multiobjective problem, one must look for the best discrete representation of the Pareto front. Evaluating the quality of a Pareto front approximation is not trivial. It itself involves several factors such as the closeness to the Pareto front and the coverage in the objective space. Measures should capture these factors. To compare multiobjective optimization algorithms, the choice of a good performance indicator is crucial [9]. Hansen and Jaszkiewicz [19] are the first to introduce a mathematical framework to evaluate the performance of metrics according to the comparison of methods. In their work, they define what can be considered as a good measure to evaluate the quality of Pareto front. This work has been extended in [8, 9, 10]. We next define the notion of an approximation.

Definition 4 (Pareto set approximation) *A set of vectors A in the decision space is called a* Pareto set approximation *if no element of this set is dominated any other. The image of such a set in the objective space is called a* Pareto front approximation. *The set of all Pareto set approximations is denoted Ψ.*

Remark 2 *We use the terms* Pareto set approximation *and* Pareto front approximation *in the remaining of the paper.*

Zitzler et al. [10] propose an extension of the relation order for decision vectors to Pareto set approximations. They are summarized in Table 2, and Figures 2 and 3 illustrate these concepts.

Table 2: Comparison relations between approximation sets [10]. Notice that $A \prec\prec B \Longrightarrow A \prec B \Longrightarrow A \lhd B \Longrightarrow A \preceq B$

Relation	Decision vectors x and x'		Approximation sets A and B	
Strictly dominates	$x \prec\prec x'$	x is better than x' in all objectives	$A \prec\prec B$	Every $x' \in B$ is strictly dominated by at least one $x \in A$
Dominates	$x \prec x'$	x is not worse than x' in all objectives and better in at least one objective	$A \prec B$	Every $x' \in B$ is dominated by at least one $x \in A$
Weakly dominates	$x \preceq x'$	x is not worse than x' in all objectives	$A \preccurlyeq B$	Every $x' \in B$ is weakly dominated by at least one $x \in A$
Is better			$A \lhd B$	Every $x' \in B$ is weakly dominated by at least one $x \in A$ and $A \neq B$
Is incomparable	$x \parallel x'$	Neither x weakly dominates x' nor x' weakly dominates x	$A \parallel B$	Neither A weakly dominates B nor A weakly dominates B

Measures are defined on approximation sets. They are designed as quality indicators or performance indicators [10].

Definition 5 (Quality indicator) *A* quality (unary) indicator *is a function $I : \Psi \rightarrow \mathbb{R}$ which assigns a real number to an Pareto set approximation.*

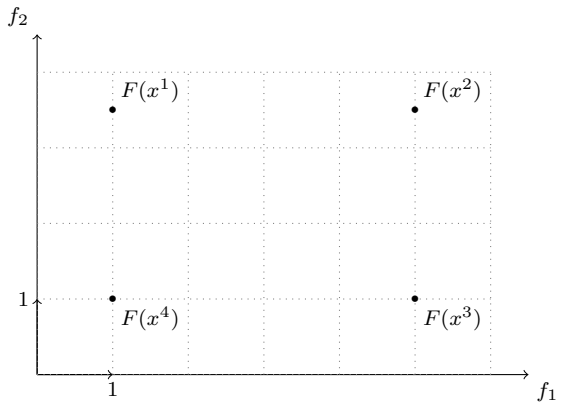

Figure 2: Example of the dominance relation for objective vectors for a biobjective problem (inspired by [10]): $x^4 \prec\prec x^2$, $x^4 \prec x^2$, $x^4 \preceq x^2$, $x^1 \prec x^2$, $x^3 \prec x^2$, $x^4 \preceq x^1$, $x^4 \prec x^1$, $x^4 \preceq x^3$, $x^4 \preceq x^1$, $x^1 \preceq x^1$, $x^2 \preceq x^2$, $x^3 \preceq x^3$, $x^4 \preceq x^4$ **and** $x^1 \parallel x^3$

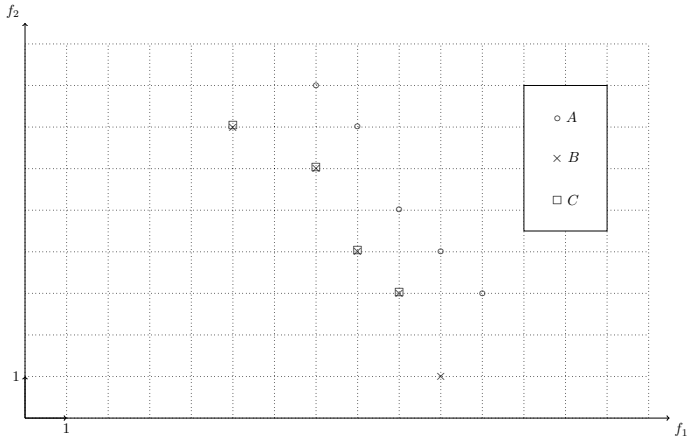

Figure 3: Example of the dominance relation for Pareto set approximations in the objective space for a biobjective problem (inspired by [10]): $C \prec A$, $B \prec A$, $B \prec\prec A$, $B \prec C$, $B \preceq C$, $C \preceq A$, $B \preceq A$, $A \preceq A$, $B \preceq B$, $C \preceq C$, $C \lhd A$, $B \lhd A$ **and** $B \lhd C$

A performance indicator may consider several Pareto set approximations. The most common ones are mappings that take only one or two Pareto set approximations as arguments. They are known respectively as unary and binary performance indicators. With such a quality indicator, one can define a relation order between different Pareto set approximations. The indicators that are interesting are the ones that capture the Pareto dominance.

Definition 6 (Monotonicity) *A quality indicator* $I: \Psi \to \mathbb{R}$ *is* monotonic *if and only if*

$$\text{For all } A, B \in \Psi,\ A \preceq B \implies I(A) \geq I(B).$$

Similarly, a quality indicator $I: \Psi \to \mathbb{R}$ *is* strictly monotonic *if and only if*

$$\text{For all } A, B \in \Psi,\ A \prec B \implies I(A) > I(B).$$

Once the notion of performance indicator is defined, the definition of comparison method can be introduced.

Definition 7 (Comparison method) *Let* $A, B \in \Psi$ *be two Pareto set approximations,* $I = (I_1, I_2, \ldots, I_k)$ *a combination of quality indicators and* $E: \mathbb{R}^k \times \mathbb{R}^k \to \{true,\ false\}$ *a Boolean function taking two vectors of size* k *as arguments. If all* I_i *for* $i = 1, 2, \ldots, k$ *are unary, the* comparison method $C_{I,E}(A, B)$ *is defined*

as a Boolean function by the following formula:

$$C_{I,E}(A,B) = E\left(I(A),\ I(B)\right)$$

where for all $Y \in \Psi,\ I(Y) = \ (I_1(Y), I_2(Y), \ldots, I_k(Y))$.

If every I_i *for* $i = 1, 2, \ldots, k$ *is binary, the comparison method* $C_{I,E}(A,B)$ *is defined as a Boolean function by*

$$C_{I,\ E}(A,B) = E\left(I(A,B), I(B,A)\right)$$

where for all $Y,\ Y' \in \Psi,\ I(Y,Y') = (I_1(Y,Y'),\ I_2(Y,Y'), \ldots, I_k(Y,Y'))$.

If I is composed of a single indicator I_0, we adopt the notation $C_{I_0,E}\left(A,B\right)$ instead of $C_{I,E}(A,B)$.

Informally, a comparison method is a true/false answer to: Is a Pareto front approximation better than another one according to the combination of quality indicators I ? A simple comparison method is the following: given an unary performance indicator I and two approximation sets $A, B \in \Psi$,

> if the proposition $(C_{I,E}(A,\ B) = (I(A) > I(B)))$ is true, then A is said to be better than B according to the indicator I.

To compare several Pareto set approximations, one can be interested in defining comparison methods that capture the Pareto dominance, i.e given two Pareto set approximations $A, B \in \Psi$,

> $(C_{I,E}(A,B)$ is true$) \Longrightarrow A$ weakly dominates/strictly dominates/is better than B.

More precisely, good comparison methods should capture the $\lhd$-relation between two Pareto set approximations, as "it represents the most general and weakest form of superiority" [10]. The following definition summarizes these points:

Definition 8 (Compatibility and completeness) *Let* $\mathcal{R}$ *be an arbitrary binary relation on Pareto set approximations (typically,* $\mathcal{R} \in \{\prec, \prec\prec, \preceq, \lhd\}$*). The comparison method* $C_{I,E}$ *is denoted as* $\mathcal{R}$-compatible *if for all* A, B *Pareto set approximations, we have:*

$$C_{I,E}(A,B) \Rightarrow A\mathcal{R}B \text{ or } C_{I,E}(A,B) \Rightarrow B\mathcal{R}A.$$

The comparison method is denoted as $\mathcal{R}$-complete *if for all* A, B *Pareto set approximations,*

$$A\mathcal{R}B \Rightarrow\ C_{I,E}(A,B) \text{ or } B\mathcal{R}A \Rightarrow C_{I,E}(A,\ B).$$

For any Pareto set approximations $A, B \in \Psi$, there are no combination I of unary quality indicators such that $A \lhd B \ \Leftrightarrow\ C_{I,E}(A,\ B)$ [10].

The mathematical properties of the performance indicators mentioned in this survey are summarized in Tables 3, 4 and 5 in the appendices.

Remark 3 *The remaining of the paper uses the notations from [7]. A discrete representation of the Pareto set is denoted by* P*, called the* Pareto optimal solution set*. The Pareto set approximation (or* optimal solution set *or* practical Pareto front *[2]) returned by an algorithm will be denoted by* S *and the Pareto set approximation at iteration* k *will be denoted by* $S(k)$*. In many cases, the Pareto set is unknown. The user needs to specify a set of points in the objective space, called a* reference set *and denoted by* R*. Note that a Pareto set (approximated or not) contains only feasible points, i.e. each element of an Pareto set approximation belongs to* Ω*. It implies that if algorithm that does not find any feasible points then* $S(k)$ *is empty. For the following definitions to apply, we impose that the iteration counter* k *is set to* 0 *at the iteration where a first feasible point has been found.*

3 A classification of performance indicators

We classify performance indicators into the four following groups [13, 7, 14]:

- *Cardinality indicators* 3.1: Quantify the number of non-dominated points generated by an algorithm.
- *Convergence indicators* 3.2: Quantify how close a set of non-dominated points is from the Pareto front in the objective space.
- *Distribution and spread indicators* 3.3: Can be classified into two sub-groups. The first one measures how well distributed the points are on the Pareto front approximation; the second focuses on the extent of the Pareto front approximation, i.e. if it contains the extreme points of the Pareto front.
- *Convergence and distribution indicators* 3.4: Capture both the properties of convergence and distribution.

3.1 Cardinality indicators

These metrics focus on the number of non-dominated points generated by a given algorithm. Some of them require the knowledge of the Pareto front.

3.1.1 Overall Non-dominated vector generation ($ONVG$) [16]

$ONVG$ is the cardinality of the Pareto front approximation generated by the algorithm:

$$\text{For all } S \in \Psi,\ ONVG(S) = |S|.$$

Nonetheless, this is not a pertinent measure. For example, consider a Pareto set approximation A composed of one million non-dominated points and a Pareto set approximation B with only one point, such as this point dominates all the other points of A. A outperforms B on this metric but B is clearly better than A [9].

3.1.2 Overall Non-dominated vector generation ratio ($ONVGR$) [16]

$ONVGR$ is given by the following formula:

$$ONVGR(S, P) = \frac{|S|}{|P|}$$

where $|P|$ is the cardinality of a Pareto optimal solution set and $|S|$ the number of points of the approximation Pareto set. Notice that this indicator is just $ONVG$ divided by a scalar. Consequently, it suffers from the same drawbacks as the previous indicator.

3.1.3 Generational indicators ($GNVG$, $GNVGR$ and NVA) [16]

$GNVG(S, k)$ (generational non-dominated vector generation) is the cardinality of the number of non-dominated points $|S(k)|$ generated at iteration k for a given iterative algorithm. $GNVGR(S, P, k)$ (generational non-dominated vector generation ratio) is the ratio of non-dominated points $|S(k)|$ generated at iteration k over the cardinality of P where P is a set of points from the Pareto set. $NVA(S, k)$ (non-dominated vector addition) represents the variation of non-dominated points generated between successive iterations. It is given by:

$$NVA\,(S, k) = |S(k)| - |S(k-1)| \ \text{ for } k > 0.$$

These metrics can be used to follow the evolution of the generation of non-dominated points along iterations of a given algorithm. It seems difficult to use them as a stopping criterion as the number of non-dominated points can evolve drastically between two iterations.

3.1.4 Error ratio (ER) [16]

This measure is given by the following formula:

$$E(S) = \frac{1}{|S|} \sum_{a \in S} e_a$$

where:

$$e_a = \begin{cases} 0 & \text{if } F(a) \text{ belongs to the Pareto front.} \\ 1 & \text{otherwise.} \end{cases}$$

A set of non-dominated points far from the Pareto front will have an error ratio close to 1. Authors of [16] do not mention the presence of rounding errors in their indicator. A suggestion should be to consider an external accuracy parameter ϵ, quantifying the belonging of an element of the Pareto set approximation to the Pareto front with ϵ near to correct rounding errors.

This indicator requires the analytical expression of the Pareto front. Consequently, an user can only use it on analytical benchmark tests. Moreover, this indicator depends mostly on the cardinality of the Pareto set approximation, which can misguide interpretations. [9] illustrates this drawback with the following example. Let consider two Pareto front approximations. The first one has 100 elements, one in the Pareto front and the others close to it. Its error ratio is equal to 0.99. The second one has only two elements, one in the Pareto front, the other far from it. Its ratio is equal to 0.5. It is obvious that the first Pareto front approximation is better, even if its error ratio is bad. However, it is straightforward to compute.

Similarly to the error ratio measure, [19] defines the C_{1R} metric (called also *ratio of the reference points*). Given a reference set R (chosen by the user) in the objective space, it is the ratio of the number of points found in R over the cardinality of the Pareto set approximation.

3.1.5 C-metric or coverage of two sets (C) [44]

Let A and B be two Pareto set approximations. The C-metric maps the ordered pair $(A,\ B)$ to the interval $[0; 1]$ and is defined by:

$$C(A, B) = \frac{|\{b \in B,\ \text{there exists } a \in A \text{ such that } a \preceq b\}|}{|B|}.$$

If $C(A, B) = 1$, all the elements of B are dominated by (or equal to) the elements of A. If $C(A, B) = 0$, all the elements in B strictly dominate the elements of the set A. Both orderings have to be computed, as $C(A, B)$ is not always equal to $1 - C(A, B)$. This metric captures the proportion of points in an Pareto set approximation A dominated by the Pareto set approximation B.

Knowles et al. [9] point out the limits of this metric. If $C(A, B) \neq\ 1$ and if $C(B, A) \neq\ 1$, the two sets are incomparable. If the distribution of the sets or the cardinality is not the same, it gives some unreliable results. Moreover, it does not give an indicator of 'how much' a Pareto set approximation strictly dominates another.

Similarly to the C-metric, given a reference set R, the C_{2R} metric (*Ratio of non-dominated points by the reference set*) introduced in [19] is given by:

$$C_{2R}(S, R) = \frac{|\{x \in S;\ \text{there does not exist } r \in R \text{ such that } x \succ r\}|}{|S|}.$$

This indicator has the same drawbacks as the C-metric.

3.1.6 Mutual domination rate (MDR) [18]

The authors of [18] use this quality indicator in combination with a Kalman filter to monitor the progress of evolutionary algorithms along iterations and thus providing a stopping criterion. Given two Pareto set approximations A and B, let introduce the function $\Delta\ (A,\ B)$ that returns the set of elements of A that are dominated by at least one element of B. It is given by:

$$MDR(S,k) = \frac{|\Delta\left(S(k-1), S(k)\right)|}{|S(k-1)|} - \frac{|\Delta\left(S(k),\ S(k-1)\right)|}{|S(k)|}$$

where $S(k)$ is the Pareto set approximation generated at iteration k. It captures how many non-dominated points at iteration $k-1$ are dominated by non-dominated points generated at iteration k and reciprocally. If $MDR(S,k) = 1$, the set of non-dominated points at iteration k totally dominates its predecessor at iteration $k-1$. If $MDR(S,k) = 0$, no significant progress has been observed. $MDR(S,k) = -1$ is the worst case, as it results in a total loss of domination at the current iteration.

Cardinality indicators have a main drawback. They fail to quantify how well-distributed the Pareto front approximation is, or to quantify how it converges during the course of an algorithm.

3.2 Convergence indicators

These measures require the knowledge of the Pareto Front to be evaluated. They evaluate the distance between a Pareto front and its approximation.

3.2.1 Generational distance (GD) [16]

This indicator is given by the following formula:

$$GD(S,P) = \frac{1}{|S|}\left(\sum_{s\in S}\min_{r\in P} \parallel F(s) - F(r) \parallel^p\right)^{\frac{1}{p}}$$

where $|S|$ is the number of points in an Pareto set approximation and P a discrete representation of the Pareto front. Generally, $p = 2$. In this case, it is equivalent to the $M_1^\star$-measure defined in [6]. When $p = 1$, it is equivalent to the γ-metric defined in [22].

Similarly to GD, given a reference set R, Dist_{1R} [48] is given by:

$$\text{Dist}_{1R}(S,R) = \frac{1}{|R|}\sum_{i=1}^{|R|}\min_{x\in S}\{c(r_i,\ x)\}$$

where $c(r_i,x) = \max_{j=1,2,\ldots,m}\{0,\ w_j\left(f_j(x) - f_j(r_i)\right)\}$ with w_j a relative weight assigned to objective j.

GD is straightforward to compute but very sensitive to the number of points found by a given algorithm. In fact, if the algorithm identifies a single point in the Pareto front, the generational distance will equal 0. An algorithm can then miss an entire portion of the Pareto front without being penalized by this indicator. This measure favors algorithms returning a few non-dominated points close to the Pareto front versus those giving a more distributed representation of the Pareto front. As suggested by Colette and Siarry [2], it could be used as a stopping criteria. A slight variation of the generational distance $GD\left(S(k), S(k+1)\right)$ between two successive iterations, as long as the algorithm is running, could mean a convergence towards the Pareto front. It can be applied on continuous and discontinuous Pareto front approximations.

3.2.2 Standard deviation from the generational distance ($STDGD$) [16]

It measures the deformation of the Pareto set approximation according to a Pareto optimal solution set. It is given by the following formula:

$$STDGD(S,P) = \frac{1}{|S|}\sum_{s\in S}\left(\min_{r\in P} \parallel F(s) - F(r) \parallel - GD(S,P)\right)^2.$$

The same critics than the generational distance apply.

3.2.3 Seven points average distance ($SPAD$) [25]

As it is not practical to obtain the Pareto front, an alternative is to use a reference set R in the objective space. The $SPAD$ indicator defined for biobjective optimization problems uses a reference set composed of seven points:

$$R = \left\{ \left(\frac{i}{3}\max_{x\in\Omega} f_1(x),\ \frac{j}{3}\max_{x\in\Omega} f_2(x)\right)_{0\leq i,j\leq 3} \right\}.$$

$SPAD$ is then given by:

$$SPAD(S,R) = \frac{1}{7}\sum_{k=1}^{7}\min_{s\in S} \parallel F(s) - F(r_k) \parallel$$

where $r_k \in R$.

This indicator raises same critics as above. Notice that the computation cost to solve the single-objective problems $\max_{x\in\Omega} f_i(x)$ for $i = 1, 2$ is not negligible. Also, the points in the reference set can fail to capture the whole form of the Pareto front. Its limitation to two objectives is also an inconvenient. Nonetheless, it does not require the knowledge of Pareto front.

3.2.4 Maximum Pareto front error ($MPFE$) [16]

This indicator defined in [16] is another measure that evaluates the distance between a discrete representation of the Pareto front and the Pareto set approximation obtained by a given algorithm. It is expressed with the following formula (generally, $p = 2$):

$$MPFE(S,P) = \max_{j\in P}\left(\min_{i\in S}\sum_{h=1}^{m}|f_h(j) - f_h(i)|^p\right)^{\frac{1}{p}}.$$

It corresponds to the largest minimal distance between elements of the Pareto front approximation and their closest neighbors belonging to the Pareto front. It is not relevant, as pointed out in [9]. Let consider two Pareto fronts approximations. The first possesses only one element in the Pareto front P. The second has ten elements: nine of them belong to the Pareto front and one is some distance away from it. As $MPFE$ considers only largest minimal distances, it favors the first Pareto front approximation. But the second is clearly better.

On the contrary, it is straightforward and cheap to compute. It can be used on continuous and discontinuous problems.

3.2.5 Inverted generational distance (IGD) [23]

IGD has a quite similar form than GD. It is given by

$$IGD(S,P) = \frac{1}{|P|}\left(\sum_{i=1}^{|P|} d_i^p\right)^{\frac{1}{p}}$$

where $d_i = \min_{x\in S}||F(x) - F(i)||$ and generally, $p = 2$. Pros and cons are the same as for the GD indicator.

3.2.6 Averaged Hausdorff distance (Δ_p) [20]

In [20], the authors combine IGD and GD into a new indicator, called the averaged Hausdorff distance Δ_p defined by

$$\Delta_p(S,P) = \max\left\{GD_p(S,P),\ IGD_p(S,P)\right\}$$

where GD_p and IGD_p are slightly modified versions of the GD and IGD indicators defined as

$$GD_p(S,P) = \left(\frac{1}{|S|}\sum_{s\in S}\text{dist}\,(s,P)^p\right)^{\frac{1}{p}} \text{ and } IGD_p\,(S,P) = \left(\frac{1}{|P|}\sum_{i=1}^{|P|}\text{dist}\,(i,S)^p\right)^{\frac{1}{p}}.$$

It is straightforward to compute and to understand. On the contrary, it requires the knowledge of the Pareto front. Authors of [20] introduce this new metric to correct the defaults of the GD and IGD indicators. It can be used to compare continuous and discontinuous approximations of Pareto fronts.

3.2.7 Modified inverted generational distance (IGD^+) [24]

Although the GD and IGD indicators are commonly used due to their low computation cost [14], one of their major drawbacks is that they are non monotone [24]. The Δ_p indicator has the same problem.

Also, the authors of [24] propose a slightly different version of the IGD indicator named IGD^+ computable in $\mathcal{O}(m\ |S| \times |P|)$ where P is a fixed Pareto optimal solution set. It is weakly Pareto compliant, i.e. :

$$IGD^+(A,P) \leq IGD^+(B,P) \text{ for } A \text{ and } B \text{ two Pareto set approximations.}$$

Let $d^+(z,a) = \sum_{i=1}^{m}\left(\max(0,a_i - z_i)\right)^2$ be the modified distance calculation for minimization problems. The IGD^+ indicator is defined by

$$IGD^+(S,P) = \frac{1}{|P|}\sum_{z\in P}\min_{s\in S} d^+\left(F(z),F(s)\right).$$

As opposed to the IGD indicator, only points dominated by $z \in P$ are taken into account. A reference set R can also be used instead of P: authors of [49] analyzes the choice of such reference points. This indicator can be used with discontinuous and continuous Pareto fronts.

3.2.8 Progress metric (P_g) [16]

This indicator introduced in [50] measures the progression of the Pareto front approximation given by an algorithm towards the Pareto front in function of the number of iterations. It is defined by:

$$P_g = \ln\sqrt{\frac{f_j^{best}(0)}{f_j^{best}(k)}}$$

where $f_j^{best}(k)$ is the best value of objective function j at iteration k. Author of [16] modifies this metric to take into account whole Pareto sets approximations:

$$RP_g(S,P,k) = \ln\sqrt{\frac{GD(S(0),P)}{GD(S(k),P)}}$$

where $GD(S(k),P)$ is the generational distance of the Pareto set approximation $S(k)$ at iteration k.

P_g is not always defined, for example when values of $f_{j_{\max}}(0)$ or $f_{j_{\max}}(k)$ are negative or null. As GD is still positive, RP_g is well defined, but it requires the knowledge of the Pareto front.

P_g, when it exists, provides an estimation of the speed of convergence of the associated algorithm. RP_g captures only the variations of the generational distance along the number of iterations. The drawbacks of the generational distance do not apply in this case. Finally, a bad measure of progression does not necessarily mean that an algorithm performs poorly. Some methods less deeply explore the objective space, but reach the Pareto front after a more important number of iterations.

3.2.9 ϵ-indicator (I_ϵ) [10]

A decision vector x^1 is ϵ-dominating, for $\epsilon > 0$, a decision vector x^2 if:

$$\text{For all } i = 1, 2, \ldots, m, \quad f_i(x^1) \leq \epsilon\ f_i(x^2).$$

The ϵ-indicator for two Pareto set approximations A and B is defined as

$$I_\epsilon(A, B) = \inf_{\epsilon > 0} \left\{ x^2 \in B :\ \exists x^1 \in A \text{ such that } x^1 \text{ is } \epsilon\text{-dominating } x^2 \right\}$$

It can be calculated the following way:

$$I_\epsilon(A, B) = \max_{x^2 \in B} \min_{x^1 \in A} \max_{1 \leq i \leq m} \frac{f_i(x^1)}{f_i(x^2)}.$$

Given a reference set P, the unary metric can be defined as $I_\epsilon(S) = I_\epsilon(P, S)$.

Similarly, Zitzler [10] defines an ϵ-additive indicator based on the following ϵ-domination. It is said that a decision vector x^1 is ϵ-dominating a decision vector x^2 for $\epsilon > 0$ if for all $i = 1, 2, \ldots, m, f_i(x^1) \leq \epsilon + f_i(x^2)$. This indicator is then calculated by:

$$I_\epsilon(A, B) = \max_{x^2 \in B} \min_{x^1 \in A} \max_{1 \leq i \leq m}\ f_i(x^1) - f_i(x^2).$$

The main problem with the ϵ-indicator is that it considers only one objective, that can lead to an information loss. Consider $F(x^1) = (0, 1, 1)$ and $F(x^2) = (1, 0, 0)$ in a tri-objective maximization problem, the additive ϵ-indicator is the same for both:

$$I_\epsilon\left(\{x^1\}, \{x^2\}\right) = I_\epsilon\left(\{x^2\}, \{x^1\}\right).$$

But x^1 as a decision vector is more interesting than x^2 (the three criteria are considered equivalent) in the objective space. On the contrary, it is straightforward to compute. It can be used for continuous and discontinuous approximations of Pareto fronts.

3.2.10 Degree of approximation (DOA) [21]

This indicator is proved to be $\prec$-complete (see Definition 8). It aims to compare algorithms when the Pareto fronts are known.

Given y a point belonging to P, the set $D_{y,\ S}$ in the objective space is defined as the subset of points belonging to the Pareto set approximation S dominated by the point y. If $D_{y,S}$ is not empty, the Euclidean distance between each point $s \in D_{y,S}$ and y is computed with

$$\mathrm{d}f(y, s) = \sqrt{\sum_{j=1}^{m} \left(f_j(s) - f_j(y)\right)^2}.$$

Then the minimum Euclidean distance between $y \in P$ and $s \in D_{y,S}$ is computed with

$$d(y, S) = \begin{cases} \min\limits_{s \in D_{y,S}} \mathrm{d}f(y, s) & \text{if } |D_{y,S}| > 0 \\ \infty & \text{if } |D_{y,S}| = 0. \end{cases}$$

Similarly, $r(y,S)$ is defined for $y \in P$ by considering the set of points that do not belong to $D_{y,S}$ as:

$$r(y,S) = \begin{cases} \min\limits_{x \in S \backslash D_{y,S}} \mathrm{r}f(y,x) & \text{if } |S \backslash D_{y,S}| > 0 \\ \infty & \text{if } |S \backslash D_{y,S}| = 0 \end{cases}$$

where $\mathrm{r}f(y,x) = \sqrt{\sum_{j=1}^{m} \max\left\{0, f_j(x) - f_j(y)\right\}^2}$.

The DOA indicator is finally given by

$$DOA(S,P) = \frac{1}{|P|} \sum_{y \in P} \min\left\{d(y,S), r(y,S)\right\}.$$

The value of DOA does not depend on the number of points of P, i.e. if $|P| \gg |S|$ [21]. In fact, this indicator partitions S into subsets in which each element is dominated by a point $y \in P$. Its computation cost is quite low (in $\mathcal{O}(m\ |S| \times |P|)$). It can be used for discontinuous and continuous approximations of Pareto fronts.

3.3 Distribution and spread indicators

According to [34], "the spread metrics try to measure the extents of the spread achieved in a computed Pareto front approximation". They are not really useful to **evaluate the convergence of an algorithm, or at comparing algorithms**, but rather the distribution of the points along Pareto front approximations. They only make sense when the Pareto set is composed of several solutions.

3.3.1 Spacing (SP) [25]

This indicator is computed with

$$SP(S) = \sqrt{\frac{1}{|S|-1} \sum_{i=1}^{|S|} \left(\bar{d} - d_i\right)^2}$$

where $d_i = \min_{(s_i,s_j) \in S,\ s_i \neq s_j} \parallel F(s_i) - F(s_j) \parallel_1$ is the l_1 distance between a point $s_i \in S$ and the closest point of the Pareto front approximation produced by the same algorithm, and $\bar{d}$ the mean of the d_i.

This method cannot account for holes in the Pareto front approximation as it takes into account the distance between a point and its closest neighbor. The major issue with this metric is it gives some limited information when points given by the algorithm are clearly separated, but spread into multiple groups. On the contrary, it is straightforward to compute.

3.3.2 Delta indexes (Δ', Δ and $\Delta^\star$) [22, 27]

Deb [22] introduces the Δ' index for biobjective problems

$$\Delta'(S) = \sum_{i=1}^{|S|-1} \frac{|d_i - \bar{d}|}{|S|-1}$$

where d_i is the Euclidean distance between consecutive elements of the Pareto front approximation S, and $\bar{d}$ the mean of the d_i. As this indicator considers Euclidean distances between consecutive points, it can be misleading if the Pareto front approximation is piecewise continuous. The Δ' index does not generalize to more than 2 objectives, as it uses lexicographic order in the biobjective objective space to compute the d_i. In addition, it does not consider the extent of the Pareto front approximation, i.e. distances between extreme points of the Pareto front.

The Δ index is an indicator derived from the Δ' index to take into account the extent of the Pareto front approximation for biobjective problems:

$$\Delta(S,P) = \frac{d_f + d_l + \sum_{i=1}^{|S|-1} \left|d_i - \overline{d}\right|}{d_f + d_l + (|S|-1)\,\overline{d}}$$

where d_f and d_l are the Euclidean distances between the extreme solutions of the Pareto front P (i.e. solutions for one objective of the objective function) and the boundary solutions of the Pareto front approximation. The other notations remain the same as before. This metric requires the resolution of each single-objective optimization problem. This indicator is extended to Pareto fronts with more than two objectives by [27] to the generalized $\Delta^\star$-index:

$$\Delta^\star(S,P) = \frac{\sum_{j=1}^{m} d(e_j,S) + \sum_{i=1}^{|S|} \left|d_i - \overline{d}\right|}{\sum_{j=1}^{m} d(e_k,S) + |S|\ \ \overline{d}}$$

where $d(e_j,S) = \min\limits_{x\in S} ||F(e_j) - F(x)||$ with $e_j \in P$ the solution to the j-th single-objective problem and $d_i = \min_{(s_i,s_j)\in S,\ s_i\neq s_j} \parallel F(s_i) - F(s_j) \parallel$ the minimal Euclidean distance between two points of the Pareto front approximation. $\overline{d}$ is the mean of the d_i. As it considers consider the shortest distances between elements of the Pareto front approximation, the $\Delta^\star$ index suffers from the same drawbacks as the spacing metric. Moreover, it requires the knowledge of the extreme solutions of the Pareto front.

3.3.3 Two measures proposed by [34] (Γ and Δ)

Let assume that an algorithm computed a Pareto front approximation with N points, indexed by $1,2,\ldots,N$ to which two extreme points indexed by 0 and $N+1$ are added (for example, $s_0 = s_1$ and $s_{N+1} = s_N$). For each objective j for $j = 1,2,\ldots,m$, elements s_i for $i = 0,1,\ldots,N+1$ of the Pareto set approximation S are sorted such that for all $j = 1,2,\ldots,m$,

$$f_j(s_0) \leq f_j(s_1) \leq f_j(s_2) \leq \ldots \leq f_j(s_{N+1}).$$

Custódio et al. [34] introduces the following metric $\Gamma > 0$ defined by:

$$\Gamma(S) = \max_{j\in\{1,2,\ldots,m\}} \max_{i\in\{0,1,\ldots,N\}} \delta_{i,j}$$

where $\delta_{i,j} = f_j(s_{i+1}) - f_j(s_i)$. When considering a biobjective problem ($m = 2$), the metric reduces to consider the maximum distance in the infinity norm between two consecutive points in the Pareto front approximation as it is shown in Figure 4.

To take into account the extent of the Pareto front approximation, the authors of [34] define the following indicator by

$$\Delta(S) = \max_{j\in 1,2,\ldots,m} \left\{ \frac{\delta_{0,j} + \delta_{N,j} + \sum_{i=1}^{N-1} \left|\delta_{i,j} - \overline{\delta_j}\right|}{\delta_{0,j} + \delta_{N,j} + (N-1)\overline{\delta_j}} \right\}$$

where $\overline{\delta_j}$, for $j = 1,2,\ldots,m$, is the mean of the distances $\delta_{i,j}$ for $i = 1,2\ldots,N-1$.

The Γ and Δ indicators do not use the closest distance between two points in the objective space. Consequently, they do not have the same drawbacks as the spacing metric. However, the $\delta_{i,j}$ distance captures holes in the Pareto front if this one is piecewise discontinuous. These two metrics are more adapted to continuous Pareto front approximations.

Remark 4 *The authors of [34] suggest two ways to compute extreme points. For benchmark tests, the Pareto front is known and extreme points correspond to the ones of the Pareto front. Otherwise, the Γ and Δ indicators use the extreme points of the Pareto front approximation S.*

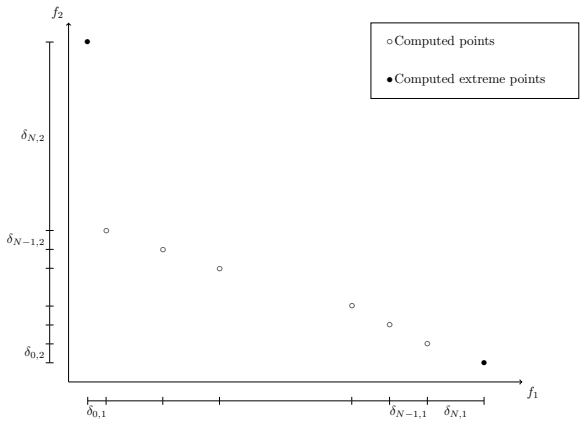

Figure 4: Illustration of the Γ metric for a biobjective problem (inspired by [34])

3.3.4 Hole relative size (HRS) [2]

This indicator identifies the largest hole in a Pareto front approximation S. It is given by

$$HRS(S) = (1/\bar{d}) \max_{i=1,2,\ldots,|S|} d_i$$

where $d_i = \min_{(s_i,s_j)\in S,\ s_i \neq s_j} \parallel F(s_i) - F(s_j) \parallel_1$ is the l_1 distance between point $s_i \in S$ and its closest neighbor, and $\bar{d}$ the mean of the d_i.

As the HRS indicator uses the minimum l_1 distance between two closest points, it has the same drawbacks as the spacing metric. It does not provide relevant information, as it does not even capture holes in the Pareto front approximation. For example, consider the following set of four non-dominated points S = $\{A(5,1),\ B(4,2),\ C(5,1),\ D(6,1)\}$ in the biobjective space. The largest gap in this Pareto front approximation in the l_1 norm is $d(B,C) = 6$; but $\max_{i=1,2,\ldots,|S|} d_i = 1$ and $HRS(S) = 1$.

3.3.5 Zitzler metrics $M_2^\star$ and $M_3^\star$ [2, 6]

The $M_2^\star$ metric returns a value in the interval $[0; |S|]$ where S is the Pareto set approximation. It reflects the number of subsets of the Pareto set approximation S of a certain size (σ). Its expression is given by

$$M_2^\star(S,\sigma) = \frac{1}{|S|-1} \sum_{x\in S} |\{y \in S, ||F(x) - F(y)|| > \sigma\}|.$$

If $M_2^\star(S) = |S|$, it means that for each objective vector, no other objective vector within the distance σ can be found. It is straightforward to compute but it can be difficult to interpret.

The authors of [40] introduce the Uniform distribution indicator, based too on the search of niches of size σ, given by

$$UD(S,\sigma) = \frac{1}{1 + D_{nc}(S,\sigma)}$$

where $D_{nc}(S,\sigma)$ is the standard deviation of the number of niches around all the points of the Pareto front approximation S defined as

$$D_{nc}(S,\sigma) = \sqrt{\frac{1}{|S|-1}\left(\sum_{i=1}^{|S|}\left(nc(s_i,\sigma) - \frac{1}{|S|}\sum_{j=1}^{|S|} nc(s_j,\sigma)\right)^2\right)}$$

with $nc(s,\sigma) = |t \in S, \parallel F(s) - F(t) \parallel < \sigma| - 1$.

Finally, the $M_3^\star$ metric defined by Zitzler [6], considers the extent of the front:

$$M_3^\star(S) = \sqrt{\sum_{i=1}^{m} \max \Big\{ \parallel F(u) - F(v) \parallel,\ u, v \in S \Big\}}.$$

The $M_3^\star$ metric only takes into account the extremal points of the computed Pareto front approximation. Consequently, it is sufficient for two different algorithms to have the same extremal points to be considered as equivalent according to this metric. It can be used on continuous and discontinuous approximations of Pareto fronts as it only gives information on the extent of the Pareto front.

3.3.6 Uniformity (δ) [41]

This is the minimal distance between two points of the Pareto front approximation. This measure is straightforward to compute and easy to understand. However, it does not really provide pertinent information on the repartition of the points along the Pareto front approximation.

3.3.7 Evenness (ξ) [32]

Given a point $F(s)$, $s \in S$, in the Pareto front approximation, and considering the closest neighbor at a distance d_s^l and the largest sphere of diameter d_s^u such that $F(s)$ and another point lie on the surface, we consider the set $D = \{d_s^u, d_s^l :\ s \in S\}$. ξ is then defined as

$$\xi(S) = \frac{\sigma_D}{\widehat{D}}$$

where σ_D is the standard deviation of D and $\widehat{D}$ its mean. The closest ξ is to 0, the better the uniformity is.

It can be considered as a coefficient of variation. It is straightforward to compute. In the case of continuous Pareto front, it cannot account for holes in the Pareto front approximation, as it considers only closest distances between two points in the objective space.

Reference [51] also defines the evenness as

$$E(S) = \frac{\max\limits_{s \in S} \min\limits_{t \in S, s \neq t} \parallel F(s) - F(t) \parallel}{\min\limits_{s \in S} \min\limits_{t \in S, s \neq t} \parallel F(s) - F(t) \parallel}.$$

The lower the value, the better the distribution with a lower bound $E(S) = 1$.

3.3.8 Binary uniformity (SPl) [33]

Contrary to others indicators, this indicator aims to compare the uniformity of two Pareto set approximations. This indicator is inspired by the wavelet theory.

Let A and B two Pareto set approximations. The algorithm is decomposed in several steps:

Let $l = 1$.

1. Firstly, for each set of non-dominated points, compute the distance between each point i of the set and its closest neighbor (for A and B respectively d_i^A and d_i^B) in the objective space.
2. Compute the mean of the d_i^A and d_i^B, i.e. $\overline{d_l^A} = \dfrac{1}{|A|} \displaystyle\sum_{i=1}^{|A|} d_i^A$ and $\overline{d_l^B} = \dfrac{1}{|B|} \displaystyle\sum_{i=1}^{|B|} d_i^B$

3. For each set, compute the following spacing measures:

$$SP_l^A = \sqrt{\sum_{i=1}^{|A|} \frac{\left(1 - \psi(d_i^A, \overline{d_l^A})\right)^2}{|A| - 1}} \quad \text{and} \quad SP_l^B = \sqrt{\sum_{i=1}^{|B|} \frac{\left(1 - \psi(d_i^B, \overline{d_l^B})\right)^2}{|B| - 1}}$$

with $\psi(a, b) = \begin{cases} \frac{a}{b} \text{ if } a > b \\ \frac{b}{a} \text{ else} \end{cases}$

4. If $SP_l^A < SP_l^B$, then A has better uniformity than B and reciprocally. If $SP_l^A = SP_l^B$ and $l \geq \min(|A| - 1, |B| - 1)$ then A has the same uniformity as B. Else if $SP_l^A = SP_l^B$ and $l < \min(|A| - 1, |B| - 1)$, then increment l by 1, and recompute the previous steps by removing the smallest distance d_i^A and d_i^B until the end.

The value of the binary uniformity indicator is difficult to interpret but can be computed easily. It does not take into account the extreme points of the Pareto front.

3.3.9 U-measure (U) [38]

The *U-measure* is given by

$$U(S) = \frac{1}{S} \sum_{i \in S} \frac{d'_i}{d_{\text{ideal}}} - 1$$

where d'_i is the distance from point i to its closest neighbor (the algorithm to find this closest neighbor is more precisely described in [38]) in the objective space translated from a distance of the extreme points of the Pareto front to their nearest neighbor and $d_{\text{ideal}} = \frac{1}{|S|} \sum_{i \in S} d'_i$.

$\frac{d'_i}{d_{\text{ideal}}} - 1$ can be interpreted as the percentage deviation from the ideal distance if it is multiplied by 100%. The U-measure is then the mean of this ratio along all points i of the Pareto front approximation. A small U can be interpreted as a better uniformity.

It attempts to quantify the uniformity of found points along the Pareto front approximation.

The same problems as for the previous metrics can be raised. Especially, the formula works only if there are several points. Moreover, this metric can take time to compute when computing the minimal distances. As for the spacing metric, this last one does not account for holes in the Pareto front approximation as it takes only into account closest neighbors. It is then more pertinent on continuous Pareto front approximations.

3.3.10 Overall Pareto spread (OS) [26]

This indicator only captures the extent of the front covered by the Pareto front approximation. The larger the better it is. It is given by

$$OS(S) = \prod_{i=1}^{m} \frac{\left| \max_{x \in S} f_i(x) - \min_{x \in S} f_i(x) \right|}{|f_i(P_B) - f_i(P_G)|}$$

where P_B is the nadir point (or an approximation) and P_G the ideal point (or an approximation).

This is an indicator for which the values are among the values 0 and 1. It needs the calculus of nadir and ideal points (so 2 m single-objective problems to preliminary solve). It does not take into account the distribution of points along the Pareto front approximation.

3.3.11 Outer diameter (I_{OD}) [8]

Analogously to the overall Pareto spread metric, the outer diameter indicator returns the maximum distance along all objective dimensions pondered by weights $w \in \mathbb{R}_+^m$ chosen by the user. It is given by:

$$I_{OD}(S) = \max_{1 \leq i \leq m} w_i \left(\max_{x \in S} f_i(x) - \min_{x \in S} f_i(x) \right).$$

The weights can be used to impose an order on criteria importance relatively to the modeling of a specific problem but it is not mandatory. Although this indicator is cheap to compute, it only takes into account the extend of the Pareto front approximation. By the way, it can result in an information loss of the extend of the Pareto front approximation, as it focuses only on the largest distance along a single dimension.

3.3.12 Distribution metric (DM) [28]

This indicator introduced by [28] aims to correct several defaults of the spacing measure [25] and add some information about the extent of the Pareto front. As it is mentioned, the "spacing metric does not adopt normalized distance, which may result in a bias conclusion, especially when the orders of magnitudes of the objectives differ considerably". Moreover, it cannot account for holes in the Pareto front, as it considers only closest neighbors. An example pointing out the defaults of the spacing metric is given in Figure 5.

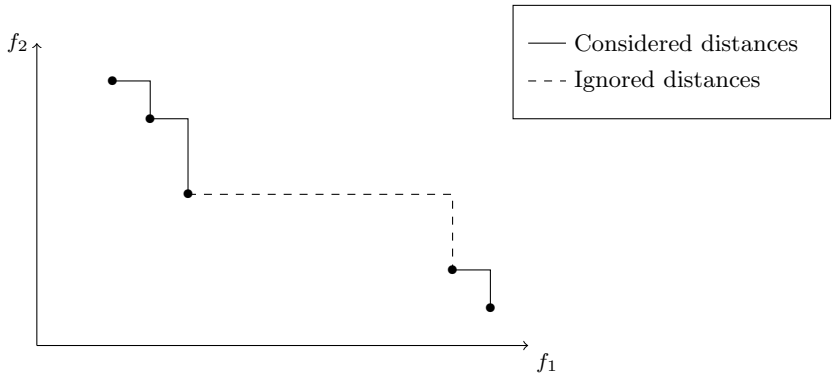

Figure 5: An example showing the weaknesses of the spacing metric (inspired by [28]): the spacing metric ignores the gap drawn in dashed lines

The DM indicator is given by

$$DM(S) = \frac{1}{|S|} \sum_{i=1}^{m} \left(\frac{\sigma_i}{\mu_i} \right) \left(\frac{|f_i(P_G) - f_i(P_B)|}{R_i} \right)$$

with $\sigma_i = \dfrac{1}{|S|-2} \displaystyle\sum_{e=1}^{|S|-1} \left(d_e^i - \overline{d^i} \right)^2$, $\mu_i = \dfrac{1}{|S|-1} \displaystyle\sum_{e=1}^{|S|-1} d_e^i$ and $R_i = \max_{s \in S} f_i(s) - \min f_i(s)$ where $|S|$ is the number of non-dominated points, $f_i(P_G)$ and $f_i(P_B)$ are the function values of design ideal and nadir points, respectively. d_e^i is the distance of the eth interval between two adjacent solutions corresponding to the ith objective, σ_i and μ_i are the standard deviation and mean of the distances relative to the ith objective, and $\dfrac{\sigma_i}{\mu_i}$ is the coefficient of variance relative to the ith objective.

A smaller DM indicates better distributed solutions. It takes into account the extent and repartition of the points along the Pareto front approximation. However, it requires the nadir and ideal points, which may be computationally expensive. As it accounts for holes, this indicator is more relevant for continuous Pareto front approximations.

3.3.13 Uniform assessment metric (I_D) [39]

Let S be a Pareto front approximation such that $|S| > 2$. The computation of this indicator is decomposed into several steps:

1. A minimum spanning tree T_G covering all the elements of S based on the euclidean distance in the objective space is built.
2. Each element $s \in S$ has at least one neighbor in the spanning set, i.e a vertex adjacent to s. Let $N_{T_G}(s)$ be the set of adjacent vertices to s in the spanning tree T_G.
   For each $v \in N_{T_G}(s)$, we define a "neighborhood" [39]
   $$N_v(s) = \{y \in S, \parallel F(y) - F(s) \parallel \leq \parallel F(v) - F(s) \parallel\}$$
   which corresponds to the subset of S contained in the closed ball of radius $\parallel F(v) - F(s) \parallel$ and centered in s. Notice that $\{s, v\} \in N_v(s)$. The neighborhoods that contain only two elements, i.e. s and v are not considered.
3. For all $s \in S$ and $v \in N_{T_G}(s)$, a distribution relation is defined by
   $$\psi(s,v) = \begin{cases} 0 & \text{if } |N_v(s)| = 2, \\ \displaystyle\prod_{y \in N_v(s),\, y \neq s} \frac{\parallel F(s) - F(y) \parallel}{\parallel F(s) - F(v) \parallel} & \text{otherwise.} \end{cases}$$
4. There are $2|S| - 2$ neighborhoods. Among them, N_r corresponds to the number of neighborhoods that only contain two elements. The uniform assessment metric is then defined by
   $$I_D(S) = \frac{1}{2|S| - N_r - 2} \sum_{s \in S} \sum_{v \in N_{T_G}(s)} \psi(s,v)$$
   which corresponds to the mean of the distribution relation for neighborhoods containing more than two elements.

This indicator does not require external parameters. Due to the definition of the neighborhood, it takes into account holes in the Pareto front. Indeed, contrary to the spacing metric, it does not consider only closest distances between objective vectors. The indicator is comprised between 0 and 1. The closest to 1, the better.

3.3.14 Extension measure (EX) [33]

This indicator aims to measure the extent of the Pareto front approximation. It is given by

$$EX(S) = \frac{1}{m} \sqrt{\sum_{i=1}^{m} d(f_i^\star, S)^2}$$

where $d(f_i^\star, S)$ is the minimal distance (norm) between the solution to the ith single-objective problem and the set of non-dominated points obtained by a given algorithm in the objective space.

This indicator requires the resolution of m single-objective optimization problems. It penalizes well-distributed Pareto front approximations neglecting the extreme values. It is straightforward to compute.

3.3.15 Diversity indicator based on reference vectors (DIR) [30]

Let $V = \{\lambda^1, \lambda^2, \ldots, \lambda^M\}$ be a set of uniformly generated reference vectors in $\mathbb{R}^m$. For each element of an approximation set $s \in S$, the closeness between s and the reference vector λ^i, for $i = 1, 2, \ldots, M$, is given by

$$\text{angle}(\lambda^i, F(s)) = \arccos \frac{(\lambda^i)^T (F(s) - F^I)}{\parallel \lambda^i \parallel \parallel F(s) - F^I \parallel}.$$

If a reference vector λ^i is the closest to an element s of S relatively to the closeness metric, it is said that s "covers the reference vector λ^i" [30]. The coverage vector c of size $|S|$ represents for each $s \in S$ the number of reference vectors that s covers. DIR is the normalized standard deviation of the coverage vector c, defined as

$$DIR = \sqrt{\frac{1}{|S|}\sum_{i=1}^{|S|}(c_i - \bar{c})^2} \ \div \ \left(\frac{M}{|S|}\sqrt{|S|-1}\right)$$

where $\bar{c}$ is the mean of the $(c_i)_{i=1,2,\ldots,|S|}$. The lower this indicator is, the better. It is intuitive to understand and cheap to compute (in $\mathcal{O}(m \times M \times |S|)$ [30]). It captures both the distribution and the spreading. Nonetheless, it requires the knowledge of the ideal point. The number of reference vectors to choose (at least greater than $|S|$ to be more pertinent) equally plays an important role. It can be biased when the Pareto front is piecewise continuous.

3.3.16 Laumanns metric (I_L) [52, 35]

Given a vector y in the objective space $\mathcal{F}$, let $D(y) = \{y' \in \mathcal{F},\ y \prec y'\}$ be the set of vectors dominated by y in the objective space. Given a Pareto front approximation S, $\mathcal{D}(S)$ is designed as the dominated space by the set S and is defined as

$$\mathcal{D}(S) = \bigcup_{y \in S} D(y).$$

Let $y^{\star i}$ be the ith outer point of the Pareto front approximation S defined by

$$(y^{\star i})_{1 \le j \le m} = \begin{cases} \max\{y_j : y \in S\} & \text{if } i \neq j, \\ \min\{y_i : y \in S\} & \text{otherwise.} \end{cases}$$

We introduce the hypercube $H(S) = \left\{ y \in \mathbb{R}^m : y = F^I + \sum_{i=1}^{m} a_i(y^{\star i} - F^I), a_i \in [0,\ 1] \right\}$ where F^I is the ideal point. The Laumanns metric is defined as the ratio of the Lebesgue measure of the intersection of $\mathcal{D}$ and H, with the Lebesgue measure of H:

$$I_L(S) = \frac{\lambda(\mathcal{D}(S) \cap H(S))}{\lambda(H(S))}$$

where $\lambda(A)$ is the Lebesgue measure of the bounded set A. The metric returns a value between 0 and 1. The higher the better. An illustration is given in Figure 6.

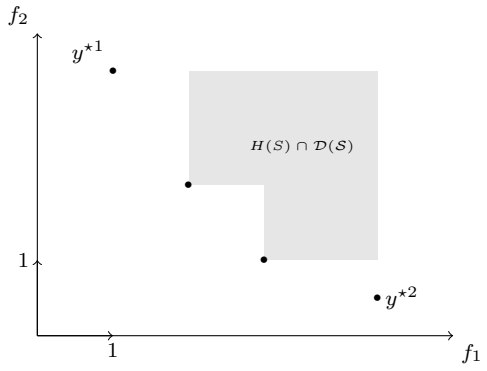

Figure 6: The intersection of $H(S)$ and $\mathcal{D}(\mathcal{S})$ for a biobjective minimization problem

This indicator is biased in favor of convex and extended fronts. Moreover, its computation complexity in $\mathcal{O}(|S|^{\frac{m}{2}} \log |S|)$ [53] explodes when the objective space dimension increases: in fact, it is similar to the hypervolume indicator when the reference point is chosen such as $\widetilde{F}^N$.

3.3.17 Other distribution indicators

Some other metrics are mentioned in this subsection. They require external parameters chosen by the user that can be crucial to their performance. The reader can consult the provided references.

1. Entropy measure [31]: For each point of S, an influential function (a Gaussian function centered in $F(s)$ for $s \in S$) is defined, which enables the creation of a density function considered as the sum of influential functions for each element $s \in S$. Peaks and valleys in the objective space are considered as places where information can be measured. A "good" Pareto front approximation should have an uniform density function in the objective space. The objective space bounded by the nadir and ideal points is firstly normalized, then divided into boxes, whose the number is decided by the user. Based on this discretization of the objective space, the measure is computed using the values of the density function for each center of each box and the Shannon formula of entropy [54].
2. Cluster CL_μ and Number of Distinct Choices NDC_μ [26]: Given two respective good (ideal point) and bad (nadir point) points P_G and P_B, the objective (preliminary normalized) is divided into hyperboxes of size μ ($\in (0;\ 1]$). NDC_μ is defined as the number of hyperboxes containing elements of the Pareto front approximation. CL_μ is then defined as $CL_\mu(S) = \frac{|S|}{NDC_\mu}$.
3. Sigma diversity metrics σ and $\overline{\sigma}$ [37]: The objective space is divided into zones delimited by uniformly distributed reference lines starting from the origin whose the number equals $|S|$. The metric value is the ratio of the number of lines that are sufficiently close to the reference lines according to the Euclidean norm with a threshold d chosen by the user, with the total number of reference lines.
4. Diversity comparison indicator DCI [29]: It is a k-ary spread indicator. The zone of interest in the objective space delimited by lower and upper bounds is divided into a number of hyperboxes. For each Pareto front approximation, a contribution coefficient is computed relatively to the hyperboxes where non-dominated points are found. For each Pareto front approximation, DCI returns the mean of contribution coefficients relatively to all hyperboxes of interest. A variant is the $M-DI$ indicator [36] (Modified Diversity Indicator) which considers a distributed reference set in the objective space instead of the set of non-dominated points from the union of the k Pareto front approximations.

A drawback of these metrics is the choice of external parameters (d threshold, μ size, number of hyperboxes) that can wrongly favor Pareto front approximations over others. σ and CL_μ can be considered as cardinal indicators too and therefore suffer from the same drawbacks as the above cardinal indicators.

3.4 Convergence and distribution indicators

These indicators are of two types: some enable to compare several approximated sets in term of distribution and Pareto dominance. The others give a value that capture distribution, spreading and convergence at the same time.

3.4.1 R_1 and R_2 indicators [19]

Let A and B be two Pareto set approximations, U a set of utility functions $u : \mathbb{R}^m \to \mathbb{R}$ mapping each point in the objective space into a measure of utility, and p a probability distribution on the set U. For each $u \in U$, let associate $u^\star(A) = \max_{s\in A} u\left(F(s)\right)$ and $u^\star(B) = \max_{s\in B} u\left(F(s)\right)$. The two indicators measure to which extent A is better than B over the set of utility functions U. The R_1 indicator is given by

$$R_1(A,B,U,p) = \int_{u\in U} C(A,B,u)p(u)\mathrm{du}$$

where

$$C(A,B,u) = \begin{cases} 1 & \text{if } u^\star(A) > u^\star(B), \\ 1/2 & \text{if } u^\star(A) = u^\star(B), \\ 0 & \text{if } u^\star(A) < u^\star(B). \end{cases}$$

The R_2 indicator defined as

$$R_2(A,B,U,p) = E\left(u^\star(A)\right) - E\left(u^\star(B)\right) = \int_{u\in U} \left(u^\star(A) - u^\star(B)\right) p(u)\mathrm{du}.$$

is the expected difference in the utility of an approximation Pareto front A with another one B. In practice, these two indicators use a discrete and finite set U of utility functions associated with an uniform distribution over U [8]. The two indicators can then be rewritten as

$$R_1(A,B) = \frac{1}{|U|}\sum_{u\in U} C(A,B,u) \text{ and } R_2(A,B,U) = \frac{1}{|U|}\sum_{u\in U} u^\star\left(A\right) - u^\star(B).$$

If $R_2(A,B,U) > 0$, then A is considered as better than B. Else if $R_2(A,B,U) \geq 0$, A is considered as not worse than B.

The authors of [19] recommend to use the utility set $U_\infty = (u_\lambda)_{\lambda\in\Lambda}$ of weighted Tchebycheff utility functions, with

$$u_\lambda(s) = -\max_{j=1,2,\ldots,m}\left(\lambda_j\left|\left(F(s)\right)_j - r_j\right|\right)$$

for $s \in A$ where r is a reference vector chosen so that any objective vector of a feasible space does not dominate r (or as an approximation of the ideal point [55, 56, 8]) and $\lambda \in \Lambda$ a weight vector such that for all $\lambda \in \Lambda$ and $j = 1, 2, \ldots, m$,

$$\lambda_j \geq 0 \text{ and } \sum_{j=1}^{m}\lambda_j = 1.$$

Zitzler [8] suggests using the set of augmented weighted Tchebycheff utility functions defined by

$$u_\lambda(s) = -\left(\max_{j=1,2,\ldots,m}\lambda_j\left|\left(F(s)\right)_j - r_j\right| + \rho\sum_{j=1}^{m}\left|\left(F(s)\right)_j - r_j\right|\right)$$

where ρ is a sufficiently small positive real number.

As given in [55], for $m = 2$ objectives, Λ can be chosen such that:

1. $\Lambda = \left\{(0,1), \left(\frac{1}{k-1},\ 1-\frac{1}{k-1}\right), \left(\frac{2}{k-1},\ 1-\frac{2}{k-1}\right), \ldots, (1,0)\right\}$ is a set of k weights uniformly distributed in the space $[0;1]^2$.
2. $\Lambda = \left\{\left(\frac{1}{1+\tan\varphi}, \frac{\tan\varphi}{1+\tan\varphi}\right), \varphi \in \Phi_k\right\}$ where $\Phi_k = \left\{0, \frac{\pi}{2(k-1)}, \frac{2\pi}{2(k-1)}, \ldots, \frac{\pi}{2}\right\}$ is a set of weights uniformly distributed over the trigonometric circle.

The I_{R_2} indicator [55] is an unary indicator derived from R_2 defined as (in the case of weighted Tchebycheff utility functions)

$$I_{R_2}(A,\Lambda) = \frac{1}{|\Lambda|}\sum_{\lambda\in\Lambda}\min_{s\in A}\left\{\max_{j=1,2,\ldots,m}\left(\lambda_j\left|\left(F(s)\right)_j - r_j\right|\right)\right\}.$$

The higher this index, the better.

As J. Knowles [9] remarks, *"the application of R_2 depends up on the assumption that it is meaningful to add the values of different utility functions from the set U. This simply means that each utility function in U must be appropriately scaled with respect to the others and its relative importance. By the way, R-metrics are only weakly monotonic, i.e. $I(A) \geq I(B)$ in A weakly dominates B"*. They do not require important computations as the number of objectives increase. The reference point has to be chosen carefully. Studies concerning the properties of the R_2 indicator can be found in [55, 56, 57].

3.4.2 G-metric [45]

This measure enables to compare k Pareto set approximations based on two criteria: their repartition of points in the space and the level of domination in the objective space. It is compatible with the weak dominance as defined below. Basically, its computation decomposes into several steps: given k Pareto set approximations $(A_1, A_2, \ldots, A_k)$:

1. Scale the values of the vectors in the k sets, i.e take the union $\bigcup_{i=1}^{k} A_i$, then normalize according to the extreme values of the objective vectors of this set.
2. Group the Pareto set approximations according to their degree of dominance. In level L_1 will be put all Pareto set approximations that strictly dominate all the others and are incomparable; we remove them then from the considered Pareto set approximations; then in L_2, will be put the Pareto set approximations that dominate all the other sets, and so on.
3. For each level of dominance L_q for $q = 1, 2, \ldots, Q$, where Q is the number of levels, dominated points belonging in the set $\bigcup_{A \in L_q} A$ are removed. Each non-dominated point in each set of the same level possesses a zone of influence. It is a ball of radius U centered in this last one. The radius U considers distances between neighbors points [38] for the k Pareto front approximations. For each Pareto set approximation belonging to the same level of dominance, a mesure of dispersion is computed. This last one takes into account the zone of influence that union of non-dominated elements of the set cover. The smaller the value, the closer the points are.
4. The G-metric associated to an Pareto set approximation is the summation of the dispersion measure of this set and the largest dispersion measure of Pareto approximated sets of lower dominance degree for each level. The bigger, the better.

The computation cost is quite important (in $\mathcal{O}(k^3 \times \max_{i=1,2,\ldots,k} |A_i|^2)$ [45]) but the cost can be decreased when one considers a small number of Pareto set approximations. Note that this indicator highly depends on the computation of the radius U when defining zones of influence. This metric can be used for continuous and discontinuous Pareto fronts, especially to compare two Pareto set approximations, in terms of dominance and distribution into the objective space.

3.4.3 Dominance move (DoM) [43]

This measure introduced by [43] was conceived to rectify the main default of the ϵ-indicator.

Definition 9 *[43] Let A be a set of points $a_1, a_2, \ldots, a_h$ and B be a set of points $b_1, b_2, \ldots, b_l$. The dominance move of A to B (denoted as $DoM(A, B)$ is the minimum total distance of moving points of A such that any point in B is weakly dominated [10] by at least one point in P. That is, we move $(a_1, a_2, \ldots, a_h)$ to positions $(a'_1, a'_2, \ldots, a'_h)$ thus constituting A' such that:*

1. *A' weakly dominates B.*
2. *The total of the moves from $a_1, a_2, \ldots, a_h$ to $a'_1, a'_2, \ldots, a'_h$ is minimized.*

Formally, the dominance move indicator is defined as

$$DoM(A, B) = \min_{A' \preceq B} \sum_{i=1}^{h} d(a_i,\ a'_i)$$

where $d(a_i, a'_i) = \| a_i - a'_i \|_1$ is the Manhattan distance between a_i and a'_i.

$DoM(A, B) \geq 0$ and if $A \preceq B$, $DoM(A, B) = 0$. Authors of [43] give an algorithm to compute this measure for biobjective problems. This relation can be used to compare sets between them. To the best of our knowledge, an algorithm for more than two objectives has not been proposed yet.

The notion of dominance move is also used in the construction of the performance comparison indicator PCI [58]. The PCI indicator evaluates the quality of multiple approximation sets by constructing a reference set thanks to them. Points in this reference set are divided into clusters (using a threshold σ). The PCI indicator measures the minimum move distance (according to the l_2 norm) of an approximation set to weakly dominate all points in a cluster.

3.4.4 Hyperarea/hypervolume metrics (HV) [44]

Named also *S-metric*, the hypervolume indicator is described as the volume of the space in the objective space dominated by the Pareto front approximation S and delimited from above by a reference point $r \in \mathbb{R}^m$ such that for all $z \in S,\ z \prec r$. The hypervolume indicator is given by

$$HV(S, r) = \lambda_m(\bigcup_{z \in S} [z; r])$$

where λ_m is the m-dimensional Lebesgue measure. An illustration is given in Figure 7 for the biobjective case ($m = 2$).

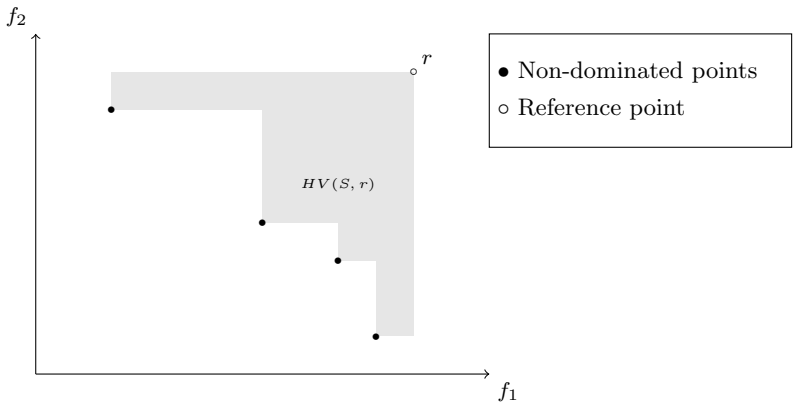

Figure 7: Illustration of the hypervolume indicator for a biobjective problem

If the Pareto front is known, *the Hyperarea ratio* is given by

$$HR(S, P, r) = \frac{HV(S, r)}{HV(P, r)}.$$

The lower the ratio is (converges toward 1), the better the approximation is.

The hypervolume indicator is the only known unary indicator to be strictly monotonic [8], i.e. if an Pareto set approximation A strictly dominates another Pareto front approximation B, $HV(A, r) > HV(B, r)$. The two main defaults are a complexity cost in $\mathcal{O}(|S|^{\frac{m}{2}} \log |S|)$ [53] and the choice of the reference point as illustrated in Figure 8.

If the origin is far from the Pareto front, the precision of the measure can decrease [7]. Recently, a practical guide was proposed to specify the reference point [59]. Besides, this measure privileges the convex parts of the Pareto front approximation over its concave parts. Other theoretical results can be found in [60, 61]. Due to its properties, it is widely used in the evolutionary community in the search of potential interesting new points or to compare algorithms.

Similarly, [44] introduces the *Difference D of two sets* S_1 and S_2. $D(S_1, S_2)$ enables to measure the size of the area dominated by S_1 not by S_2.

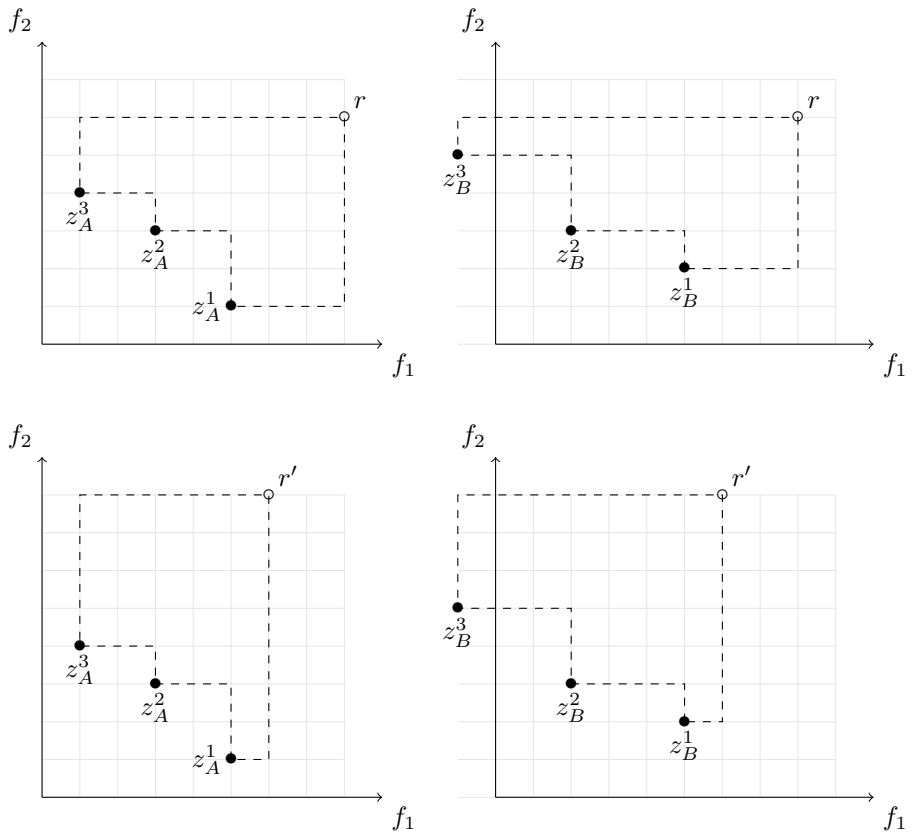

Figure 8: The relative value of the hypervolume metric depends on the chosen reference point r or r'. On the top, two non-dominated A and B sets are shown, with $HV(A, r) > HV(B, r)$. On the bottom, $HV(B, r') > HV(A, r')$

The *Hyperarea Difference* was suggested by [26] to compensate the lack of information about the theoretical Pareto front. Given a good point P_g and a bad point P_b, we can approximate the size of the area dominated by the Pareto front (or circumvent the objective space by a rectangle). The Hyperarea Difference is just the normalization of the dominated space by the approximation Pareto front over the given rectangle.

More recently, a pondered hyper-volume by weights was introduced by [62] to give a preference of an objective according to another. More volume indicators can be found in [26]. Some other authors [63] (for biobjective optimization problems) suggest to compute the hyper-volume defined by a reference point and the projection of the points belonging to the Pareto front approximation on the line delimited by the two extreme points. This measure enables to better estimate the distribution of the points along the Pareto front (in fact, it can be shown that for a linear Pareto front, an uniform distribution of points maximizes the hyper-volume indicator: see [64, 65] for more details about the properties of the hyper-volume indicator). A logarithmic version of the hypervolume indicator called the logarithmic hypervolume indicator [46] is defined by

$$\log HV(S, r) = \lambda_m \left(\bigcup_{z \in S} [\log z; \log r] \right)$$

with the same notations as previously. Notice that this indicator can only be used with positive vectors in $\mathbb{R}^m$. Finally, we can mention a generalization of the hyper-volume indicator called the cone-based hyper-volume indicator that was introduced recently by [42].

4 Some usages of performance indicators

This section focuses on three applications of performance indicators: comparison of algorithms for multiobjective optimization, definition of stopping criteria, and the use of relevant distribution and spread indicators for assessing the diversity characterization of a Pareto front approximation.

4.1 Comparison of algorithms

The first use of performance indicators is to evaluate the performance of an algorithms on a multiobjective problem. In single-objective optimization, the most used graphical tools to compare algorithms include performance profiles [66] and data profiles [67] (see also [68] for a detailed survey on the tools to compare single-optimization algorithms). More specifically, let $\mathcal{S}$ be a set of solvers and $\mathcal{P}$ the set of benchmarking problems. Let $t_{p,s} > 0$ be a performance measure of solver $s \in \mathcal{S}$ on problem $p \in \mathcal{P}$: the lower, the better. Performance and data profiles combine performance measures of solvers $t_{p,s}$ to enable a general graphic representation of the performance of each solver relatively to each other on the set of benchmarking problems $\mathcal{P}$.

To the best of our knowledge, Custódio and al [34] are the first to use data and performance profiles for multiobjective optimization. For each problem $p \in \mathcal{P}$, they build an Pareto set approximation $F_p = \bigcup_{s \in \mathcal{S}} F_{p,s}$ composed of the union of all Pareto set approximations $F_{p,s}$ generated by each solver $s \in \mathcal{S}$ for the problem p. All dominated points are then removed. Pareto approximation sets and relative Pareto front approximation are then compared using cardinality and γ and Δ metrics proposed by [34].

One of the critics we can make with this approach is the use of distribution and cardinality indicators that do not **capture order relations between two different sets**. The choice of (weakly) monotonic indicators or ($\prec$-complete / $\prec$-compatible) $\lhd$-complete / $\lhd$-compatible comparisons methods is more appropriated in this context ([19, 9, 10, 8]). Among them, dominance move, G-metric, binary ϵ-indicator and volume-space metrics have properties corresponding to these criteria. Mathematical proofs can be found in [64, 55, 56, 9, 43, 45, 10]) and are synthesized in Appendices. An example of data profile using the hypervolume indicator can be found in [69, 70]. The use of performance indicators such as GD or IGD as it is done in [71, 72] is not a pertinent choice due to their inability to capture dominance relation. Instead, we suggest to use their weakly monotonic counterpart IGD^+ or DOA, that can be cheaper to compute than for example the hypervolume indicator when the number of objectives is high.

4.2 Stopping criteria of multiobjective algorithms

To generate a Pareto front approximation, two approaches are currently considered. The first category, named as *scalarization methods*, consists in aggregating the objective functions and to solve a series of single-objective problems. Surveys about scalarization algorithms can be found for example in [73]. The second class, designed as *a posteriori articulations of preferences* [34] methods, aims at obtaining the whole Pareto front without combining any objective function in a single-objective framework. Evolutionary algorithms, Bayesian optimization methods [74] or deterministic algorithms such as DMS [34] belong to this category.

For scalarization methods, under some assumptions, solutions to single-objective problems can be proved to belong to the Pareto front or a local one. So, defining stopping criteria results in choosing the number of single-objective problems to solve via the choice of parameters and a single-objective stopping criterion for each of them. Stopping at a predetermined number of function evaluations is often used in the context of blackbox optimization [75]. The use of performance indicators also is not relevant.

A contrario, a posteriori methods consider a set of points in the objective space (a population) that is brought to head for the Pareto front along iterations. Basically, a number of maximum evaluations is still given as a stopping criterion but it remains crucial to give an estimation to how far from a (local) Pareto front the approximation set is. For multi-objective Bayesian optimization [74], the goal is to find at next iteration the

point that maximizes the hyperarea difference between old non-dominated set of points and the new one. The performance indicator is directly embedded into the algorithm and could be used as a stopping criterion. For evolutionary algorithms, surveys on stopping criteria for multiobjective optimization can be found in [18, 76]. The approach is to measure the progression of the current population combining performance indicators (hypervolume, MDR, etc.) and statistic tools (Kalman filter [18], χ^2-variance test [77], etc.) These last ones enable to detect a stationary state reached by the evolving population.

We believe that the use of monotonic performance indicators or binary ones that capture the dominance property seems to be the most efficient one in the years to come to follow the behavior of population-based algorithms along iterations.

4.3 Distribution and spread

The choice of spread and distribution metrics has only a sense when one wants to measure the repartition of points in the objective space, no matter how close from the Pareto front the approximated set is. Spread and distribution metrics can put forward global properties (for example statistics on the repartition of the points or extent of the front) or local properties such as the largest distance between closest non-dominated points that can be used to conduct search such as Γ indicator. Typically, the construction of a distribution or spread indicator requires two steps. The first consists in defining a distance between two points in the objective space. Many distribution metrics in the literature use minimum Euclidean or Manhattan distance between points such as the SP metric, the Δ index, HRS, and so on. The DM and Γ-metric indicators use a "sorting distance"; I_D a "neighborhood distance" based on a spanning tree, and so on. Once this is done, many of the existing distribution indicators are built by using statistic tools on this distance: mean (Δ index, U measure, DM for example), mean square (SP, D_{nc}), and so on.

To use a distribution or spread indicator, it should satisfy the following properties:

1. The support of scaled functions, which enables to compare all objectives in an equivalent way ($DM, OS, I_{OD}, \Delta, \Gamma$).
2. For piecewise continuous or discontinuous Pareto front approximations, a good distribution indicator should not be based on the distance between closest neighbors, as it can hide some holes [28]. Some indicators possess this property such as DM, Γ, Δ or evenness indicators.
3. Distribution and spread performance indicators should not be based on external parameters, such as Zitzler metric $M_2^\star$, UD, or entropy measure.
4. An easy interpretation: a value returned by an indicator has to be 'intuitive' to understand. For example, the binary uniformity is extremely difficult to interpret and should not be used. This remark applies for all types of performance indicators.

One could directly include spread control parameters in the design of new algorithms. The Normal Boundary Intersection method [78] controls the spread of a Pareto front approximation. This method is also used in the context of blackbox optimization [79].

5 Discussion

In this work, we give a review of performance indicators for the quality of Pareto front approximations in multiobjective optimization, as well as some usages of these indicators.

The most important application of performance indicators is to allow comparison and analysis of results of different algorithms. In this optic, among all these indicators, the hypervolume metric and its binary counterpart, the hyperarea difference can be considered until now as the most relevant. The hypervolume indicator possesses good mathematical properties, it can capture dominance properties and distribution and does not require the knowledge of the Pareto front. Empirical studies [13, 7] have confirmed its efficiency compared to other performance indicators. That is why it has been deeply used in the evolutionary community [14]. However, it has some limitations: the exponential cost as the number of objectives increases and

the choice of the reference point. To compare algorithms, it can be replaced with other indicators capturing lower dominance relation such as dominance move, G-metric, binary ϵ-indicator, modified inverted generated distance or degree of approximation whose computation cost is less important.

Future research can focus on the discovery of new performance indicators that correct some drawbacks of the hypervolume indicator but keeps its good properties, and the integration of performance indicators directly into algorithms for multiobjective optimization.

Appendix A A summary of performance indicators

Table 3 draws a summary of all indicators described in Section 3. Most of complexity cost indications for computing metrics are drawn from [13]. P corresponds to the *Pareto optimal solution set* and S is a Pareto set approximation returned by a given algorithm. The symbol "✗" indicates that the performance indicator does not satisfy the monotony property. The "-" symbol corresponds to binary indicators, for which monotonicity has no meaning.

Table 3: A summary of performance indicators

Category	Performance indicators	Sect.	Symbol	Parameters	Comparison sets	Computational complexity	Monotone
Cardinality 3.1	C-metric/Two sets Coverage [15]	3.1.5	C	None	Binary indicator	$\mathcal{O}(m \ \|S_1\| \times \|S_2\|)$	-
	Error ratio [16]	3.1.4	ER	None	Pareto front P	Low	✗
	Generational non dominated vector generation [17]	3.1.3	$GNVG$	None	Pareto front P	Low	✗
	Generational non dominated vector generation ratio [17]	3.1.3	$GNVGR$	None	Pareto front P	Low	✗
	Mutual domination rate [18]	3.1.3	MDR	None	None	Low	✗
	Nondominated vector additional [17]	3.1.3	NVA	None	None	Low	✗
	Overall nondominated vector generation [16]	3.1.1	$ONVG$	None	None	Low	✗
	Overall nondominated vector generation ratio [16]	3.1.2	$ONVGR$	None	Pareto front P	Low	✗
	Ratio of non-dominated points by the reference set [19]	3.1.5	C_{2R}	None	Reference set R	$\mathcal{O}(m \ \|S\| \times \|R\|)$	✗
	Ratio of the reference points [19]	3.1.4	C_{1R}	None	Reference set R	$\mathcal{O}(m \ \|S\| \times \|R\|)$	✗
Convergence 3.2	Averaged Hausdorff distance [20]	3.2.6	Δ_q	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	✗
	Degree of Approximation [21]	3.2.10	DOA	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	Not strictly
	D_R-metric [19]	3.2.1	-	None	Reference set R	$\mathcal{O}(m \ \|S\| \times \|R\|)$	Not strictly
	ϵ-family [10]	3.2.9	I_ϵ	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	Not strictly
	Generational distance [16]	3.2.1	GD	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	✗
	γ-metric [22]	3.2.1	γ	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	✗
	Inverted generational distance [23]	3.2.5	IGD	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	✗
	Maximum Pareto front error [16]	3.2.4	$MPFE$	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	✗
	Modified inverted generational distance [24]	3.2.7	IGD^+	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	Not strictly
	$M_1^\star$-metric [6]	3.2.1	$M_1^\star$	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	✗
	Progression metric [16]	3.2.8	-	None	None	$\mathcal{O}(m \ \|S\|)$	✗
	Seven points average distance [25]	3.2.3	$SPAD$	None	Reference set R	$\mathcal{O}(m \ \|S\|)$	✗
	Standard deviation from the Generational distance [16]	3.2.2	$STDGD$	None	Pareto front P	$\mathcal{O}(m \ \|S\| \times \|P\|)$	✗

Table 3 – continued from previous page

Category	Performance indicators	Sect.	Symbol	Parameters	Comparison sets	Computational complexity	Monotone
Distribution and spread 3.3	Cluster [26]	3.3.17	CL_μ	A parameter μ	None	High	✗
	Δ-index [22]	3.3.2	Δ	None	Pareto front P	$\mathcal{O}(m\ \lvert S\rvert^2 + m\ \lvert S\rvert \times \lvert P\rvert)$	✗
	Δ'-index [22]	3.3.2	Δ'	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	$\Delta^\star$ spread metric [27]	3.3.2	$\Delta^\star$	None	Pareto front P	$\mathcal{O}(m\ \lvert S\rvert^2 + m\ \lvert S\rvert \times \lvert P\rvert)$	✗
	Distribution metric [28]	3.3.12	DM	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	Diversity comparison indicator [29]	3.3.17	DCI	A parameter div	k-ary indicator comparing $S_1, S_2, \ldots, S_k$ non-dominated sets	$\mathcal{O}\left(m(k\lvert S_{\max}\rvert)^2\right)$	✗
	Diversity indicator [30]	3.3.15	DIR	Number of weights vectors M	None	$\mathcal{O}(mM\lvert S\rvert)$	✗
	Entropy metric [31]	3.3.17	-	A parameter $grids$	None	High	✗
	Evenness [32]	3.3.7	ξ	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	Extension [33]	3.3.14	EX	None	Pareto front P	$\mathcal{O}(m\ \lvert S\rvert \times \lvert P\rvert)$	✗
	Γ-metric [34]	3.3.3	Γ	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	Hole Relative Size [2]	3.3.4	HRS	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	Laumanns metric [35]	3.3.16	-	None	None	$\mathcal{O}(\lvert S\rvert^{\frac{m}{2}} \log \lvert S\rvert)$	✗
	Modified Diversity indicator [36]	3.3.17	$M-DI$	A parameter δ	Reference set R	$\mathcal{O}(m\lvert R\rvert \times \lvert S\rvert^2)$	
	$M_2^\star$-metric [6]	3.3.5	$M_2^\star$	Niche radius σ	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	$M_3^\star$-metric [6]	3.3.5	$M_3^\star$	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	Number of distinct choices [26]	3.3.17	NDC_μ	A parameter μ	None	High	✗
	Outer diameter [8]	3.3.11	I_{OD}	None	None	$\mathcal{O}(m\ \lvert S\rvert)$	✗
	Overall Pareto Spread [26]	3.3.10	OS	None	Good Point P_G and Bad Point P_B	$\mathcal{O}(m\ \lvert S\rvert)$	✗
	Sigma diversity metric [37]	3.3.17	σ	A parameter $lines$	None	High	✗
	Spacing [25]	3.3.1	SP	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	U-measure [38]	3.3.9	U	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	Uniform assessment metric [39]	3.3.13	I_D	None	None	$\mathcal{O}\left(m\lvert S\rvert^2\right)$	✗
	Uniform distribution [40]	3.3.5	UD	Niche radius σ	None	$O(m\ \lvert S\rvert^2)$	✗
	Uniformity [41]	3.3.6	δ	None	None	$\mathcal{O}(m\ \lvert S\rvert^2)$	✗
	Uniformity [33]	3.3.8	-	None	Binary	Quadratic	✗
Convergence and distribution 3.4	Cone-based hypervolume [42]	3.4.4	-	Angle γ	Reference set R	$\mathcal{O}(\lvert S\rvert^{\frac{m}{2}} \log \lvert S\rvert)$	Strictly
	Dominance move [43]	3.4.3	DoM	None	Binary indicator	$\mathcal{O}(\lvert S\rvert \log \lvert S\rvert)$	-
	D-metric/Difference coverage of two sets [44]	3.4.4	-	None	Reference set R Binary indicator	$\mathcal{O}(\lvert S\rvert^{\frac{m}{2}} \log \lvert S\rvert)$	-
	Hyperarea difference [26]	3.4.4	HD	None	Reference set R	$\mathcal{O}(\lvert S\rvert^{\frac{m}{2}} \log \lvert S\rvert)$	Strictly
	Hypervolume indicator (or S-metric) [6]	3.4.4	HV	None	Reference set R	$\mathcal{O}(\lvert S\rvert^{\frac{m}{2}} \log \lvert S\rvert)$	Strictly
	G-metric [45]	3.4.2	-	None	k-ary indicator comparing $S_1, S_2, \ldots, S_k$ non-dominated sets	$\mathcal{O}(k^3\ \lvert S_{\max}\rvert^2)$	Not strictly
	Logarithmic hypervolume indicator [46]	3.4.4	$\log HV$	None	Reference set R	$\mathcal{O}(\lvert S\rvert^{\frac{m}{2}} \log \lvert S\rvert)$	Strictly
	R-metric [19]	3.4.1	R	A set Λ of weights vectors	Reference set R	$\mathcal{O}(m\ \lvert S\rvert \times \lvert R\rvert \times \lvert\Lambda\rvert)$	Not strictly

Appendix B Compatibility and completeness

Table 4 and 5 summarize compatibility and completeness properties. Only the strongest relationships are kept. Some of them are drawn from [10]. All spread and distribution indicators are not compatible with approximation sets relations.

Table 4: Compatibility and completeness of unary performance indicators

Category	Performance indicators	Sect.	Symbol	Boolean function	Compatible	Complete
Cardinality 3.1	Error ratio [16]	3.1.4	ER	$ER(A) < ER(B)$	✗	✗
	Generational non dominated vector generation [17]	3.1.3	$GNVG$	-	-	-
	Generational non dominated vector generation ratio [17]	3.1.3	$GNVGR$	-	-	-
	Mutual domination rate [18]	3.1.6	MDR	-	-	-
	Nondominated vector additional [17]	3.1.3	NVA	-	-	-
	Overall nondominated vector generation [16]	3.1.1	$ONVG$	$ONVG(A) > ONVG(B)$	✗	✗
	Overall nondominated vector generation ratio [16]	3.1.2	$ONVGR$	$ONVGR(A,P) > ONVGR(B,P)$	✗	✗
	Ratio of non-dominated points by the reference set [19]	3.1.5	C_{2R}	$C_{2R}(A,\ R) > C_{2R}(B,\ R)$	✗	✗
	Ratio of the reference points [19]	3.1.4	C_{1R}	$C_{1R}(A,R) > C_{1R}(B,\ R)$	✗	✗
Convergence 3.2	Averaged Hausdorff distance [20]	3.2.6	Δ_q	$\Delta_q(A,P) < \Delta_q(B,P)$	✗	✗
	Degree of Approximation [21]	3.2.10	DOA	$DOA(A,P) < DOA(B,P)$	Not better than	$\prec$
	D_R-metric [19]	3.2.1	-	$D_R(A,R) < D_R(B,R)$	Not better than	$\prec\prec$
	Generational distance [16]	3.2.1	GD	$GD(A,P) < GD(B,P)$	✗	✗
	γ-metric [22]	3.2.1	γ	$\gamma(A,P) < \gamma(B,P)$	✗	✗
	Inverted generational distance [23]	3.2.5	IGD	$IGD(A,P) < IGD(B,P)$	✗	✗
	Maximum Pareto front error [16]	3.2.4	$MPFE$	$MPFE(A,P) < MPFE(B,P)$	✗	✗
	Modified inverted generational distance [24]	3.2.7	IGD^+	$IGD^+(A,P) < IGD^+(B,P)$	Not better than	$\preceq$
	$M_1^\star$-metric [6]	3.2.1	$M_1^\star$	$M_1^\star(A,P) < M_1^\star(B,P)$	✗	✗
	Progression metric [16]	3.2.8	-	-	-	-
	Seven points average distance [25]	3.2.3	$SPAD$	$SPAD(A,P) < SPAD(B,P)$	✗	✗
	Standard deviation from the Generational distance [16]	3.2.2	$STDGD$	-	-	-
Distribution and spread 3.3	Cluster [26]		CL_μ	-	-	-
	Δ-index [22]	3.3.2	Δ	$\Delta(A) < \Delta(B)$	✗	✗
	Δ'-index [22]	3.3.2	Δ'	$\Delta'(A) < \Delta'(B)$	✗	✗
	$\Delta^\star$ spread metric [27]	3.3.2	$\Delta^\star$	$\Delta^\star(A) < \Delta^\star(B)$	✗	✗
	Distribution metric [28]	3.3.12	DM	$DM(A) < DM(B)$	✗	✗
	Diversity indicator [30]	3.3.15	DIR	$DIR(A) < DIR(B)$	✗	✗
	Entropy metric [31]	3.3.17	-	-	-	-
	Evenness [32]	3.3.7	ξ	$\xi(A) < \xi(B)$	✗	✗
	Extension [33]	3.3.14	EX	$EX(A) < EX(B)$	✗	✗
	Γ-metric [34]	3.3.3	Γ	$\Gamma(A) < \Gamma(B)$	✗	✗
	Hole Relative Size [2]	3.3.4	HRS	$HRS(A) < HRS(B)$	✗	✗
	Laumanns metric [35]	3.3.16	-	$I_L(A) > I_L(B)$	✗	✗
	Modified Diversity indicator [36]	3.3.17	$M-DI$	$M-DI(A,R) > M-DI(B,R)$	✗	✗
	$M_2^\star$-metric [6]	3.3.5	$M_2^\star$	$M_2^\star(A,\sigma) > M_2^\star(B,\sigma)$	✗	✗
	$M_3^\star$-metric [6]	3.3.5	$M_3^\star$	$M_3^\star(A) > M_3^\star(B)$	✗	✗
	Number of distinct choices [26]	3.3.17	NDC_μ	$NDC_\mu(A) > NDC_\mu(B)$	✗	✗
	Outer diameter [8]	3.3.11	I_{OD}	$I_{OD}(A) > I_{OD}(B)$	✗	✗
	Overall Pareto Spread [26]	3.3.10	OS	$OS(A) > OS(B)$	✗	✗
	Sigma diversity metric [37]	3.3.17	σ	$\sigma(A) > \sigma(B)$	✗	✗
	Spacing [25]	3.3.1	SP	$SP(A) < SP(B)$	✗	✗

Table 4 – continued from previous page

Category	Performance indicators	Sect.	Symbol	Boolean function	Compatible	Complete
	U-measure [38]	3.3.9	U	$U(A) < U(B)$	✗	✗
	Uniform assessment metric [39]	3.3.13	I_D	$I_D(A) > I_D(B)$	✗	✗
	Uniform distribution [40]	3.3.5	UD	$UD(A,\sigma) < UD(B,\sigma)$	✗	✗
	Uniformity [41]	3.3.6	δ	$\delta(A) < \delta(B)$	✗	✗
Convergence and distribution 3.4	Cone-based hypervolume [42]	3.4.4	-	$\chi(A) > \chi(B)$	Not better than	$\lhd$
	Hyperarea difference [26]	3.4.4	HD	$HD(A) < HD(B)$	Not better than	$\lhd$
	Hypervolume indicator (or S-metric) [6]	3.4.4	HV	$HV(A,r) > HV(B,r)$	Not better than	$\lhd$
	Logarithmic hypervolume indicator [46]	3.4.4	$\log HV$	$\log HV(A,r) > \log HV(B,r)$	Not better than	$\lhd$

Table 5: Compatibility and completeness of binary indicators (inspired by [10]): a - means there is no comparison method which is complete and compatible for the given relation, a ✗ that the indicator is not even monotone

Category	Performance indicators	Sect.	Symbol	Relation			
				$\lhd$	$\preceq$	$=$	$\|$
Cardinality 3.1	C-metric/Two sets Coverage [15]	3.1.5	C	$C(A,B)=1$ $C(B,A)<1$	$C(A,B)=1$	$C(A,B)=1$ $C(B,A)=1$	$C(A,B)>1$ $C(B,A)>1$
Convergence 3.2	Additive ϵ-indicator [10]	3.2.9	I_ϵ	$I_\epsilon(A,B)\leq 0$ $I_\epsilon(B,A)>0$	$I_\epsilon(A,B)\leq 0$	$I_\epsilon(A,B)=0$ $I_\epsilon(B,A)=0$	$I_\epsilon(A,B)>0$ $I_\epsilon(B,A)>0$
Distribution and spread 3.3	Diversity comparison indicator [29]	3.3.17	DCI	✗	✗	✗	✗
	Uniformity [33]	3.3.8	-	✗	✗	✗	✗
Convergence and distribution 3.4	Dominance move [43]	3.4.3	DoM	$DoM(A,B)=0$ $DoM(B,A)>0$	$DoM(A,B)=0$ $DoM(B,A)\geq 0$	$DoM(A,B)=0$ $DoM(B,A)=0$	$DoM(A,B)>0$ $DoM(A,B)>0$
	D-metric/Difference coverage of two sets [44]	3.4.4	-	$D(A,B)>0$ $D(B,A)=0$	$D(A,B)\geq 0$ $D(B,A)=0$	$D(A,B)=0$ $D(B,A)=0$	$D(A,B)>0$ $D(B,A)>0$
	G-metric [45]	3.4.2	-	-	-	-	-
	R-metric [19]	3.4.1	R	-	-	-	-

References

[1] J. Branke, K. Deb, K. Miettinen, R. Slowiński, Multiobjective optimization: Interactive and evolutionary approaches, Vol. 5252, Springer Science & Business Media, 2008.

[2] Y. Collette, P. Siarry, Optimisation multiobjectif: Algorithmes, Editions Eyrolles, 2011.

[3] A. Custódio, M. Emmerich, J. A. Madeira, Recent developments in derivative-free multiobjective optimization, Computational Technology Reviews 5 (2012) 1–30.

[4] K. Deb, Multi objective optimization using evolutionary algorithms, John Wiley and Sons, 2001.

[5] R. Slowinski, J. Teghem (Eds.), Stochastic vs. Fuzzy Approaches to Multiobjective Mathematical Programming Under Uncertainty, Kluwer Academic Publishers, Norwell, MA, USA, 1990.

[6] E. Zitzler, K. Deb, L. Thiele, Comparison of multiobjective evolutionary algorithms: Empirical results, Evolutionary computation 8 (2) (2000) 173–195.

[7] T. Okabe, Y. Jin, B. Sendhoff, A critical survey of performance indices for multi-objective optimisation, in: Evolutionary Computation, 2003. CEC'03. The 2003 Congress on, Vol. 2, IEEE, 2003, pp. 878–885.

[8] E. Zitzler, J. Knowles, L. Thiele, Quality assessment of pareto set approximations, Multiobjective Optimization (2008) 373–404.

[9] J. Knowles, D. Corne, On metrics for comparing nondominated sets, in: Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on, Vol. 1, IEEE, 2002, pp. 711–716.

[10] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, V. G. Da Fonseca, Performance assessment of multiobjective optimizers: An analysis and review, IEEE Transactions on evolutionary computation 7 (2) (2003) 117–132.

[11] S. L. Faulkenberg, M. M. Wiecek, On the quality of discrete representations in multiple objective programming, Optimization and Engineering 11 (3) (2010) 423–440.

[12] S. Cheng, Y. Shi, Q. Qin, On the performance metrics of multiobjective optimization, in: International Conference in Swarm Intelligence, Springer, 2012, pp. 504–512.

[13] S. Jiang, Y. Ong, J. Zhang, L. Feng, Consistencies and contradictions of performance metrics in multiobjective optimization, IEEE Transactions on Cybernetics 44 (12) (2014) 2391–2404.

[14] N. Riquelme, C. Von Lücken, B. Baran, Performance metrics in multi-objective optimization, in: Computing Conference (CLEI), 2015 Latin American, IEEE, 2015, pp. 1–11.

[15] E. Zitzler, L. Thiele, Multiobjective optimization using evolutionary algorithms—a comparative case study, in: international conference on parallel problem solving from nature, Springer, 1998, pp. 292–301.

[16] D. Van Veldhuizen, Multiobjective evolutionary algorithms: classifications, analyses, and new innovations. 1999, School of Engineering of the Air Force Institute of Technology, Dayton, Ohio.

[17] D. Van Veldhuizen, G. B. Lamont, On measuring multiobjective evolutionary algorithm performance, in: Evolutionary Computation, 2000. Proceedings of the 2000 Congress on, Vol. 1, IEEE, 2000, pp. 204–211.

[18] L. Martí, J. García, A. Berlanga, J. M. Molina, A stopping criterion for multi-objective optimization evolutionary algorithms, Information Sciences 367 (2016) 700–718.

[19] M. P. Hansen, A. Jaszkiewicz, Evaluating the quality of approximations to the non-dominated set, IMM, 1998.

[20] O. Schutze, X. Esquivel, A. Lara, C. A. Coello Coello, Using the averaged hausdorff distance as a performance measure in evolutionary multiobjective optimization, IEEE Transactions on Evolutionary Computation 16 (4) (2012) 504–522.

[21] E. Dilettoso, S. A. Rizzo, N. Salerno, A weakly pareto compliant quality indicator, Mathematical and Computational Applications 22 (1) (2017) 25.

[22] K. Deb, S. Agrawal, A. Pratap, T. Meyarivan, A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii, in: International Conference on Parallel Problem Solving From Nature, Springer, 2000, pp. 849–858.

[23] C. A. Coello Coello, N. Cortés, Solving multiobjective optimization problems using an artificial immune system, Genetic Programming and Evolvable Machines 6 (2) (2005) 163–190.

[24] H. Ishibuchi, H. Masuda, Y. Tanigaki, Y. Nojima, Modified distance calculation in generational distance and inverted generational distance, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2015, pp. 110–125.

[25] J. R. Schott, Fault tolerant design using single and multicriteria genetic algorithm optimization., Tech. rep., Air force Institute Of Tech Wright-Patterson AFB OH (1995).

[26] J. Wu, S. Azarm, Metrics for quality assessment of a multiobjective design optimization solution set, Journal of Mechanical Design 123 (1) (2001) 18–25.

[27] A. Zhou, Y. Jin, Q. Zhang, B. Sendhoff, E. Tsang, Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion, in: Evolutionary Computation, 2006. CEC 2006. IEEE Congress on, IEEE, 2006, pp. 892–899.

[28] K. Zheng, R. Yang, H. Xu, J. Hu, A new distribution metric for comparing pareto optimal solutions, Structural and Multidisciplinary Optimization 55 (1) (2017) 53–62.

[29] M. Li, S. Yang, X. Liu, Diversity comparison of pareto front approximations in many-objective optimization, IEEE Transactions on Cybernetics 44 (12) (2014) 2568–2584.

[30] X. Cai, H. Sun, Z. Fan, A diversity indicator based on reference vectors for many-objective optimization, Information Sciences 430 (2018) 467–486.

[31] A. Farhang-Mehr, S. Azarm, An information-theoretic entropy metric for assessing multi-objective optimization solution set quality, Journal of Mechanical Design 125 (4) (2003) 655–663.

[32] A. Messac, C. A. Mattson, Normal constraint method with guarantee of even representation of complete pareto frontier, AIAA journal 42 (10) (2004) 2101–2111.

[33] H. Meng, X. Zhang, S. Liu, New quality measures for multiobjective programming, Advances in Natural Computation (2005) 431–431.

[34] A. Custódio, J. A. Madeira, A. F. Vaz, L. N. Vicente, Direct multisearch for multiobjective optimization, SIAM Journal on Optimization 21 (3) (2011) 1109–1140.

[35] M. Laumanns, E. Zitzler, L. Thiele, A unified model for multi-objective evolutionary algorithms with elitism, in: Evolutionary Computation, 2000. Proceedings of the 2000 Congress on, Vol. 1, IEEE, 2000, pp. 46–53.

[36] M. Asafuddoula, T. Ray, H. Singh, Characterizing pareto front approximations in many-objective optimization, in: Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, ACM, 2015, pp. 607–614.

[37] S. Mostaghim, J. Teich, A new approach on many objective diversity measurement, in: Dagstuhl Seminar Proceedings, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2005.

[38] Y. Leung, Y. Wang, U-measure: a quality measure for multiobjective programming, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 33 (3) (2003) 337–343.

[39] M. Li, J. Zheng, G. Xiao, Uniformity assessment for evolutionary multi-objective optimization, in: Evolutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational Intelligence). IEEE Congress on, IEEE, 2008, pp. 625–632.

[40] K. Tan, T. H. Lee, E. Khor, Evolutionary algorithms for multi-objective optimization: Performance assessments and comparisons, Artificial intelligence review 17 (4) (2002) 251–290.

[41] S. Sayın, Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming, Mathematical Programming 87 (3) (2000) 543–560.

[42] M. Emmerich, A. Deutz, J. Kruisselbrink, P. K. Shukla, Cone-based hypervolume indicators: construction, properties, and efficient computation, in: International conference on evolutionary multi-criterion optimization, Springer, 2013, pp. 111–127.

[43] M. Li, X. Yao, Dominance move: A measure of comparing solution sets in multiobjective optimization, arXiv preprint arXiv:1702.00477.

[44] E. Zitzler, Evolutionary algorithms for multiobjective optimization: Methods and applications, Ph.D. thesis, Swiss Federal Institute of Technology Zurich (1999).

[45] G. Lizarraga-Lizarraga, A. Hernandez-Aguirre, S. Botello-Rionda, G-metric: an m-ary quality indicator for the evaluation of non-dominated sets, in: Proceedings of the 10th annual conference on Genetic and evolutionary computation, ACM, 2008, pp. 665–672.

[46] T. Friedrich, K. Bringmann, T. Voß, C. Igel, The logarithmic hypervolume indicator, in: Proceedings of the 11th workshop proceedings on Foundations of genetic algorithms, ACM, 2011, pp. 81–92.

[47] M. Ehrgott, Multicriteria optimization, Springer Science & Business Media, 2006.

[48] P. Czyżżak, A. Jaszkiewicz, Pareto simulated annealing—a metaheuristic technique for multiple-objective combinatorial optimization, Journal of Multi-Criteria Decision Analysis 7 (1) (1998) 34–47.

[49] H. Ishibuchi, H. Masuda, Y. Tanigaki, Y. Nojima, Difficulties in specifying reference points to calculate the inverted generational distance for many-objective optimization problems, in: Computational Intelligence in Multi-Criteria Decision-Making (MCDM), 2014 IEEE Symposium on, IEEE, 2014, pp. 170–177.

[50] T. Back, Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms, Oxford university press, 1996.

[51] D. Ghosh, D. Chakraborty, A direction based classical method to obtain complete pareto set of multi-criteria optimization problems, Opsearch 52 (2) (2015) 340–366.

[52] M. Laumanns, R. Günter, H. Schwefel, Approximating the pareto set: Concepts, diversity issues, and performance assessment, Secretary of the SFB 531, 1999.

[53] N. Beume, C. Fonseca, M. López-Ibáñez, L. Paquete, J. Vahrenhold, On the complexity of computing the hypervolume indicator, IEEE Transactions on Evolutionary Computation 13 (5) (2009) 1075–1082.

[54] C. E. Shannon, A mathematical theory of communication, ACM SIGMOBILE mobile computing and communications review 5 (1) (2001) 3–55.

[55] D. Brockhoff, T. Wagner, H. Trautmann, On the properties of the r2 indicator, in: Proceedings of the 14th annual conference on Genetic and evolutionary computation, ACM, 2012, pp. 465–472.

[56] D. Brockhoff, T. Wagner, H. Trautmann, 2 indicator-based multiobjective search, Evolutionary Computation 23 (3) (2015) 369–395.

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

[58] M. Li, S. Yang, X. Liu, A performance comparison indicator for pareto front approximations in many-objective optimization, in: Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, ACM, 2015, pp. 703–710.

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

[60] K. Bringmann, T. Friedrich, Tight bounds for the approximation ratio of the hypervolume indicator, in: International Conference on Parallel Problem Solving from Nature, Springer, 2010, pp. 607–616.

[61] K. Bringmann, T. Friedrich, Approximation quality of the hypervolume indicator, Artificial Intelligence 195 (2013) 265–290.

[62] E. Zitzler, D. Brockhoff, L. Thiele, The hypervolume indicator revisited: On the design of pareto-compliant indicators via weighted integration, in: Evolutionary multi-criterion optimization, Springer, 2007, pp. 862–876.

[63] S. Jiang, S. Yang, M. Li, On the use of hypervolume for diversity measurement of pareto front approximations, in: Computational Intelligence (SSCI), 2016 IEEE Symposium Series on, IEEE, 2016, pp. 1–8.

[64] A. Auger, J. Bader, D. Brockhoff, E. Zitzler, Theory of the hypervolume indicator: optimal μ-distributions and the choice of the reference point, in: Proceedings of the tenth ACM SIGEVO workshop on Foundations of genetic algorithms, ACM, 2009, pp. 87–102.

[65] P. K. Shukla, N. Doll, H. Schmeck, A theoretical analysis of volume based pareto front approximations, in: Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation, ACM, 2014, pp. 1415–1422.

[66] E. D. Dolan, J. J. Moré, Benchmarking optimization software with performance profiles, Mathematical programming 91 (2) (2002) 201–213.

[67] J. J. Moré, S. M. Wild, Benchmarking derivative-free optimization algorithms, SIAM Journal on Optimization 20 (1) (2009) 172–191.

[68] V. Beiranvand, W. Hare, Y. Lucet, Best practices for comparing optimization algorithms, Optimization and Engineering 18 (4) (2017) 815–848.

[69] G. Liuzzi, S. Lucidi, F. Rinaldi, A derivative-free approach to constrained multiobjective nonsmooth optimization, SIAM Journal on Optimization 26 (4) (2016) 2744–2774.

[70] S. Zapotecas-Martínez, C. A. Coello Coello, Monss: A multi-objective nonlinear simplex search approach, Engineering Optimization 48 (1) (2016) 16–38.

[71] A. Al-Dujaili, S. Suresh, Bmobench: Black-box multi-objective optimization benchmarking platform, arXiv preprint arXiv:1605.07009.

[72] D. Brockhoff, T. Tran, N. Hansen, Benchmarking numerical multiobjective optimizers revisited, in: Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, ACM, 2015, pp. 639–646.

[73] M. M. Wiecek, M. Ehrgott, A. Engau, Continuous multiobjective programming, in: Multiple Criteria Decision Analysis, Springer, 2016, pp. 739–815.

[74] M. Emmerich, K. Yang, A. Deutz, H. Wang, C. M. Fonseca, A multicriteria generalization of bayesian global optimization, in: Advances in Stochastic and Deterministic Global Optimization, Springer, 2016, pp. 229–242.

[75] C. Audet, G. Savard, W. Zghal, Multiobjective optimization through a series of single-objective formulations, SIAM Journal on Optimization 19 (1) (2008) 188–210.

[76] T. Wagner, H. Trautmann, L. Martí, A taxonomy of online stopping criteria for multi-objective evolutionary algorithms, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2011, pp. 16–30.

[77] T. Wagner, H. Trautmann, B. Naujoks, Ocd: Online convergence detection for evolutionary multi-objective algorithms based on statistical testing, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2009, pp. 198–215.

[78] I. Das, J. E. Dennis, Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems, SIAM Journal on Optimization 8 (3) (1998) 631–657.

[79] C. Audet, G. Savard, W. Zghal, A mesh adaptive direct search algorithm for multiobjective optimization, European Journal of Operational Research 204 (3) (2010) 545–556.