Performance indicators in multiobjective optimization

Charles Audet, Jean Bigeon, Dominique Cartier, Sébastien Le Digabel, Ludovic Salomon

European Journal of Operational Research 292 (2021) 397–422

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Invited Review

Performance indicators in multiobjective optimization

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

[a] GERAD and Département de mathématiques et génie industriel, Polytechnique Montréal, C.P. 6079, Succ. Centre-ville, Montréal, Québec H3C 3A7, Canada
[b] Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, Grenoble 38000, France

ARTICLE INFO

Article history:
Received 3 December 2018
Accepted 11 November 2020
Available online 18 November 2020

Keywords:
Multiobjective optimization
Quality indicators
Performance indicators

ABSTRACT

In 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 front approximations produced by these algorithms. In this work, we propose a review of a total of 63 performance indicators partitioned into four groups according to their properties: cardinality, convergence, distribution and spread. Applications of these indicators are presented as well.

1. Introduction

Since the eighties, a large number of methods has been developed to treat multiobjective optimization problems (e.g Branke, Deb, Miettinen, & Słowiński, 2008; Collette & Siarry, 2011; Custódio, Emmerich, & Madeira, 2012; Deb, 2001; Slowinski & Teghem, 1990). Given that conflicting objectives are provided, the set of solutions, the *Pareto set*, is described as the set of best decision vectors corresponding to the best trade-off points in the objective space. Knowledge of the Pareto set 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^1 is said to *(Pareto)-dominate* another feasible vector x^2 if x^1 is at least as good as x^2 for all the objectives, and strictly better than x^2 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 set is the set of *Pareto solutions*, whose images (by the objective functions) constitute the *Pareto front*.

In single-objective minimization problems, the quality of a given solution is trivial to quantify: the smaller the corresponding 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 multiobjective optimization methods. According to Zitzler, Deb, and Thiele (2000), a Pareto set approximation should satisfy the following:

- The distance between the Pareto front and its representation in the objective space should be minimized.
- A good (according to some metric) distribution of the points of the corresponding approximated front in the objective space is desirable.
- The extent of the corresponding 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* Okabe, Jin, and Sendhoff (2003); Zitzler, Knowles, and Thiele (2008) have been introduced. Performance indicators can be considered as mappings that assign scores to Pareto front approximations.

Surveys of performance indicators already exist but they focus only on some specific properties. In (Collette & Siarry, 2011, chapter 7), the authors list some performance indicators to measure the quality of a Pareto front approximation. In Okabe et al. (2003), 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 Knowles and Corne (2002) and Zitzler, Thiele, Laumanns, Fonseca, and da Fonseca (2003) and additional measures and algorithms are listed in Faulkenberg and Wiecek (2009). In Cheng, Shi, and Qin (2012), the authors review some performance indicators and analyze their drawbacks.

* Corresponding author.
E-mail addresses: Charles.Audet@gerad.ca, charles.audet@polymtl.ca (C. Audet), jean.bigeon@grenoble-inp.fr (J. Bigeon), Dominique.cartier@csdhr.qc.ca (D. Cartier), Sebastien.Le.Digabel@gerad.ca (S. Le Digabel), ludovic.salomon@gerad.ca (L. Salomon).

https://doi.org/10.1016/j.ejor.2020.11.016
0377-2217/

In Jiang, Ong, Zhang, and Feng (2014), an empirical study focuses on the correlations between different indicators with their computational complexity on concave and convex Pareto fronts. Finally, the usage of indicators proposed by the multiobjective evolutionary optimization community prior to 2013 is analyzed in Riquelme, Lücken, and Baran (2015).

Another survey Li and Yao (2019) on the quality evaluation of solution sets was recently published. It complements the present study by categorizing more performance indicators, but presents them at a higher level of description. Most of the indicators in the present work are listed in Li and Yao (2019), but not all of them (e.g., the ones proposed in Custódio, Madeira, Vaz, & Vicente, 2011). While Li and Yao (2019) is mainly oriented toward the evolutionary algorithms community, the present work addresses the whole operational research community, and also addresses some issues in more detail, such as complexity computational costs, as well as data and performance profiles. The reader is invited to consult their survey as a useful complement.

The present work is an attempt to propose a survey offering a panorama on all important aspects of performance indicators contrary to the previous surveys, addressed to the whole multiobjective optimization community. This work systematically analyzes 63 performance indicators by partitioning them into these four categories: Cardinality, Convergence, Distribution and spread, Convergence and distribution. The use of performance metrics targets four cases: comparison of algorithms, embedding of performance indicators into multiobjective methods, suggestion of stopping criteria for multiobjective optimization and identification of promising distribution-based performance indicators. Table 1 lists these indicators and their category, classifies them based on their properties, and indicates the section in which they are discussed. This work is organized as follows. Section 2 introduces the notations and definitions related to multiobjective optimization and performance 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.

2. Notations and definitions

To apprehend performance 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 performance indicators.

2.1. Multiobjective optimization and Pareto dominance

We consider the following continuous multiobjective optimization problem:

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

where $\mathcal{X} \subseteq \mathbb{R}^n \neq \emptyset$ 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 $\mathcal{Y} = \{F(x) \in \mathbb{R}^m : x \in \mathcal{X}\}$ is called the *(feasible) objective set*. The sets $\mathbb{R}^n$ and $\mathbb{R}^m$ are respectively denoted as the *decision space* and the *objective space*.

To compare functions objective values, the following cone order relation is adopted (Ehrgott, 2005).

Definition 1 (Dominance relations between objective vectors Dächert, Klamroth, Lacour, & Vanderpooten, 2017)**.** Given two objective vectors y^1 and y^2 in the objective space $\mathbb{R}^m$, we write:

- $y^1 \leqq y^2$ (y^1 *weakly dominates* y^2) if and only if $y^1_i \leq y^2_i$ for all $i = 1, 2, \dots, m$.
- $y^1 \leq y^2$ (y^1 *dominates* y^2) if and only if $y^1 \leqq y^2$ and $y^1 \neq y^2$.
- $y^1 < y^2$ (y^1 *strictly dominates* y^2) if and only if $y^1_i < y^2_i$ for all $i = 1, 2, \dots, m$.

In the case when neither $y^1 \nleqq y^2$ nor $y^2 \nleqq y^1$, y^1 and y^2 are said to be *incomparable*.

We can now present the concept of dominance relations for the decision vectors.

Definition 2 (Dominance relations for decision vectors Audet, Savard, & Zghal, 2008)**.** Given two decision vectors x^1 and x^2 in the feasible set $\mathcal{X} \subseteq \mathbb{R}^n$, we write:

- $x^1 \preceq x^2$ (x^1 weakly dominates x^2) if and only if $F(x^1) \leqq F(x^2)$.
- $x^1 \prec x^2$ (x^1 dominates x^2) if and only if $F(x^1) \leq F(x^2)$.
- $x^1 \prec\prec x^2$ (x^1 strictly dominates x^2) if and only if $F(x^1) < F(x^2)$.
- $x^1 \parallel x^2$ (x^1 and x^2 are incomparable) if neither x^1 weakly dominates x^2 nor x^2 weakly dominates x^1.

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

Definition 3 (Pareto optimality and Pareto solutions Ehrgott, 2005)**.** The vector $x \in \mathcal{X}$ is a *Pareto-optimal* solution if there is no other vector in $\mathcal{X}$ 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 $\mathcal{Y}_P$.

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 Ehrgott, 2005). For a problem with m objectives, the Pareto front $\mathcal{Y}_P$ is at most of dimension $m - 1$. For three objectives, the Pareto front is a surface, a curve, a point, or combinations of surfaces, curves and/or points, or the empty set. For two objectives, the Pareto front can be a curve or a point or a combination of curves and/or points, or the empty set. Also, it is interesting to define some bounds on this set.

Definition 4 (Ideal and nadir objective vectors)**.** The *ideal objective vector* y^I Ehrgott (2005) is defined as the objective vector whose components are the solutions of each single-objective problem $\min\limits_{x \in \mathcal{X}} f_i(x)$, $i = 1, 2, \dots, m$. The *nadir objective vector* y^N is defined as the objective vector whose components are the solutions of the single-objective problems $\max\limits_{x \in \mathcal{X}_P} f_i(x), i = 1, 2, \dots, m$.

For computational reasons, the nadir objective vector is often approximated by $\tilde{y}^N$ for which the coordinates are defined the following way: let $x^{k,\star}$ be the solution of the single-objective problem $\min\limits_{x \in \mathcal{X}} f_k(x)$ for $k = 1, 2, \dots, m$. The ith coordinate of $\tilde{y}^N$ is given by:

$$\tilde{y}^N_i = \max_{k=1,2,\dots,m} f_i(x^{k,\star}).$$

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

An illustration is given in Fig. 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. 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.

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

Table 1
A summary of performance indicators. The 9 rightmost columns indicate references where the indicators are presented.

Category	Performance indicators	Sect.	Cheng et al. (2012)	Collette and Siarry (2011)	Knowles and Corne (2002)	Zitzler et al. (2003)	Okabe et al. (2003)	Faulkenberg and Wiecek (2009)	Jiang et al. (2014)	Riquelme et al. (2015)	Li and Yao (2019)
Cardinality 3.1	C-metric/Two sets Coverage Zitzler and Thiele (1998)	3.1.5	√	√	√		√		√	√	√
	Error ratio Van Veldhuizen (1999)	3.1.4		√	√	√	√		√		√
	Generational non dominated vector generation Van Veldhuizen and Lamont (2000)	3.1.3		√			√		√	√	
	Generational non dominated vector generation ratio Van Veldhuizen and Lamont (2000)	3.1.3					√		√	√	
	Mutual domination rate Martí et al. (2016)	3.1.6								√	
	Nondominated vector additional Van Veldhuizen and Lamont (2000)	3.1.3		√			√		√		
	Overall nondominated vector generation Van Veldhuizen (1999)	3.1.1		√	√	√	√	√	√	√	√
	Overall nondominated vector generation ratio Van Veldhuizen (1999)	3.1.2		√	√	√	√		√	√	√
	Ratio of non-dominated points by the reference set Hansen and Jaszkiewicz (1998)	3.1.5					√		√		√
	Ratio of the reference points Hansen and Jaszkiewicz (1998)	3.1.4					√		√		√
Convergence 3.2	Degree of Approximation Dilettoso et al. (2017)	3.2.7									√
	ϵ-family Zitzler et al. (2003)	3.2.6				√			√	√	√
	Generational distance Van Veldhuizen (1999)	3.2.1		√	√	√	√	√	√	√	√
	γ-metric Deb et al. (2000)	3.2.1	√	√			√		√	√	
	Maximum Pareto front error Van Veldhuizen (1999)	3.2.4		√	√	√	√	√	√		√
	$M_1^\star$-metric Zitzler et al. (2000)	3.2.1		√		√	√		√	√	√
	Progression metric Van Veldhuizen (1999)	3.2.5		√							
	Seven points average distance Schott (1995)	3.2.3					√		√		√
	Standard deviation from the Generational distance Van Veldhuizen (1999)	3.2.2		√							
Distribution and spread 3.3	Cluster Wu and Azarm (2000)	3.3.18				√	√	√	√	√	√
	Δ-index Deb et al. (2000)	3.3.2	√				√	√		√	√
	Δ'-index Deb et al. (2000)	3.3.2					√		√	√	√
	$\Delta^\star$ spread metric Zhou et al. (2006)	3.3.2						√	√	√	√
	Distribution metric Zheng et al. (2017)	3.3.12									
	Diversity comparison indicator Li et al. (2014)	3.3.18									√
	Diversity indicator Cai et al. (2018)	3.3.15									√
	Entropy metric Farhang-Mehr and Azarm (2004)	3.3.18					√	√	√	√	√

(*continued on next page*)

Table 1 (*continued*)

Category	Performance indicators	Sect.	Cheng et al. (2012)	Collette and Siarry (2011)	Knowles and Corne (2002)	Zitzler et al. (2003)	Okabe et al. (2003)	Faulkenberg and Wiecek (2009)	Jiang et al. (2014)	Riquelme et al. (2015)	Li and Yao (2019)
	Evenness Messac and Mattson (2004)	3.3.7						√			√
	Extension Meng et al. (2005)	3.3.14						√			√
	Γ-metric Custódio et al. (2011)	3.3.3				√					
	Hole Relative Size Collette and Siarry (2005)	3.3.4		√		√		√			√
	Laumanns metric Laumanns et al. (2000)	3.3.17		√							
	Modified Diversity indicator Asafuddoula et al. (2015)	3.3.18									√
	$M_2^\star$-metric Zitzler et al. (2000)	3.3.5		√		√	√	√	√		√
	$M_3^\star$-metric Zitzler et al. (2000)	3.3.5	√	√		√	√	√	√	√	√
	Number of distinct choices Wu and Azarm (2000)	3.3.18				√	√	√	√		√
	Outer diameter Zitzler et al. (2008)	3.3.11				√					√
	Overall Pareto Spread Wu and Azarm (2000)	3.3.10				√	√	√	√	√	√
	Riesz S-Energy Hardin and Saff (2004)	3.3.16									
	Sigma diversity metric Mostaghim and Teich (2005)	3.3.18							√		√
	Spacing Schott (1995)	3.3.1		√	√	√	√	√	√	√	√
	U-measure Leung and Wang (2003)	3.3.9						√			√
	Uniform assessment metric Li et al. (2008)	3.3.13									√
	Uniform distribution Tan et al. (2002)	3.3.5					√		√		√
	Uniformity Sayın (2000)	3.3.6						√			√
	Uniformity Meng et al. (2005)	3.3.8						√			√
Convergence and distribution 3.4	Averaged Hausdorff distance Schutze et al. (2012)	3.4.2							√		√
	Cone-based hypervolume Emmerich et al. (2013b)	3.4.7									
	Dominance move Li and Yao (2017)	3.4.6									√
	D-metric/Difference coverage of two sets Zitzler (1999)	3.4.7	√			√	√		√	√	√
	D_R-metric Czyzzak and Jaszkiewicz (1998)	3.4.3			√	√	√			√	√
	Hyperarea difference Wu and Azarm (2000)	3.4.7					√	√	√	√	√
	Hypervolume indicator (or S-metric) Zitzler et al. (2000)	3.4.7	√	√	√	√	√	√	√	√	√
	Hypervolume Sharpe-ratio indicator Yevseyeva et al. (2014)	3.4.8									
	Inverted generational distance Coello and Cortés (2005)	3.4.1	√						√	√	√
	Inverted generation distance with non contributed solutions detection Tian et al. (2016)	3.4.1									√
	G-metric Lizarraga-Lizarraga et al. (2008)	3.4.5									√
	Logarithmic hypervolume indicator Friedrich et al. (2011)	3.4.7									
	Modified inverted generational distance Ishibuchi et al. (2015)	3.4.3									√
	Performance comparison indicator Li et al. (2015)	3.4.6									√
	p, q-averaged distance Vargas and Bogoya (2018)	3.4.2									√
	R-metric Hansen and Jaszkiewicz (1998)	3.4.4				√	√	√	√	√	√

Table 2
Comparison relations between Pareto front approximations Zitzler et al. (2003). Note that $Y_N^1 \prec\prec Y_N^2 \Longrightarrow Y_N^1 \prec Y_N^2 \Longrightarrow Y_N^1 \lhd Y_N^2 \Longrightarrow Y_N^1 \preceq Y_N^2$.

Relation	Objective vectors y^1 and y^2		Pareto front approximations Y_N^1 and Y_N^2	
Strictly dominates	$y^1 < y^2$	y^1 is better than y^2 in all objectives	$Y_N^1 \prec\prec Y_N^2$	Every $y^2 \in Y_N^2$ is strictly dominated by at least one $y^1 \in Y_N^1$
Dominates	$y^1 \leq y^2$	y^1 is not worse than y^2 in all objectives and better in at least one objective	$Y_N^1 \prec Y_N^2$	Every $y^2 \in Y_N^2$ is dominated by at least one $y^1 \in Y_N^1$
Weakly dominates	$y^1 \leqq y^2$	y^1 is not worse than y^2 in all objectives	$Y_N^1 \preceq Y_N^2$	Every $y^2 \in Y_N^2$ is weakly dominated by at least one $y^1 \in Y_N^1$
Is better			$Y_N^1 \lhd Y_N^2$	Every $y^2 \in Y_N^2$ is weakly dominated by at least one $y^1 \in Y_N^1$ and $Y_N^1 \neq Y_N^2$
Is incomparable		Neither y^1 weakly dominates y^2 nor y^2 weakly dominates y^1	$Y_N^1 \parallel Y_N^2$	Neither Y_N^1 weakly dominates Y_N^2 nor Y_N^1 weakly dominates Y_N^2

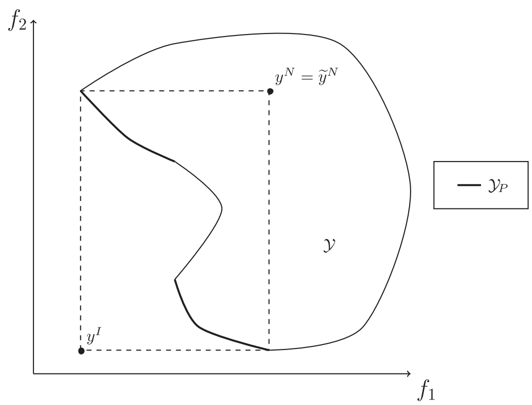

Fig. 1. Objective space, Pareto front $\mathcal{Y}_P$, ideal objective vector y^I and nadir objective vector y^N (inspired by Ehrgott, 2005).

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. Indicators should capture these factors. To compare multiobjective optimization algorithms, the choice of a good performance indicator is crucial (Knowles & Corne, 2002). Hansen and Jaszkiewicz (1998) are the first to introduce a mathematical framework to evaluate the performance of indicators 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 Knowles and Corne (2002), Zitzler et al. (2008, 2003). We next define the notion of an approximation.

Definition 5 (Pareto set approximation Zitzler et al. (2008))**.** A set of decision vectors X_N in the feasible set is called a *Pareto set approximation* if no element of this set is weakly dominated by any other. The image of such a set in the objective space is called a *Pareto front approximation* denoted $Y_N = F(X_N) \subseteq \mathbb{R}^m$. The set of all Pareto set approximations is denoted Ψ and the set of all Pareto front approximations is denoted Ω.

By definition, the Pareto front approximation corresponding to a given Pareto set approximation possesses only distinct elements, i.e. two different elements of the Pareto set approximation cannot map to the same point in the objective space. Consequently, for all $X_N \in \Psi,\ |X_N| = |Y_N|$.

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

Zitzler et al. (2003) propose an extension of the relation order for objective vectors to Pareto front approximations. They are summarized in Table 2.

These relations orders can be naturally extended to Pareto set approximations.

Measures are defined on Pareto front approximations. They are designed as quality indicators or performance indicators Zitzler et al. (2003).

Definition 6 (Performance indicator Zitzler et al., 2003)**.** A *k-ary* performance indicator is a function $I:\ \Omega^k \to \mathbb{R}$ which assigns to each collection $Y_N^1, Y_N^2, \ldots, Y_N^k$ of k Pareto front approximations a real value $I(Y_N^1, Y_N^2, \ldots, Y_N^k)$.

A performance indicator may consider several Pareto front approximations. The most common ones are mappings that take only one or two Pareto front 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 front approximations. The interesting indicators are those that capture the Pareto dominance in the objective space.

Definition 7 (Monotonicity Zitzler et al., 2008)**.** Assuming a greater indicator value is preferable, a performance indicator $I:\ \Omega \to \mathbb{R}$ is *monotonic* if and only if

$$\text{for all } Y_N^1, Y_N^2 \in \Omega,\ \ Y_N^1 \preceq Y_N^2 \implies I(Y_N^1) \geq I(Y_N^2).$$

Similarly, assuming a greater indicator value is preferable, a performance indicator $I\ :\ \Omega \to \mathbb{R}$ is *strictly monotonic* if and only if

$$\text{for all } Y_N^1, Y_N^2 \in \Omega,\ \ Y_N^1 \lhd Y_N^2 \implies I\left(Y_N^1\right) > I(Y_N^2).$$

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

Definition 8 (Comparison method Zitzler et al., 2003)**.** Let $Y_N^1, Y_N^2 \in \Omega$ be two Pareto front approximations, $I = (I_1, I_2, \ldots, I_k)$ a combination of performance 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* $C_{I,E}(Y_N^1, Y_N^2)$ is defined as a Boolean function by the following formula:

$$C_{I,E}(Y_N^1, Y_N^2) = E\left(I(Y_N^1),\ I(Y_N^2)\right)$$

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

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

$$C_{I,E}(Y_N^1, Y_N^2) = E\left(I(Y_N^1, Y_N^2), I(Y_N^2, Y_N^1)\right)$$

where for all $Y_N^1,\ \ Y_N^2 \in \Omega, I(Y_N^1, Y_N^2) = (I_1(Y_N^1, Y_N^2), I_2(Y_N^1, Y_N^2), \ldots, I_k(Y_N^1, Y_N^2))$.

If I is composed of a single indicator I_0, we adopt the notation $C_{I_0,E}\left(Y_N^1, Y_N^2\right)$ instead of $C_{I,E}(Y_N^1, Y_N^2)$.

Informally, a comparison method is a true/false answer to:

"Is a Pareto front approximation better than another one according to the combination of performance indicators I?" A simple comparison method is the following: given an unary performance indicator I and two Pareto front approximations $Y_N^1, Y_N^2 \in \Omega$,

if the proposition $\left(C_{I,E}(Y_N^1,\ Y_N^2) = \left(I(Y_N^1) > I(Y_N^2)\right)\right)$ is true, then Y_N^1 is said to be better than Y_N^2 according to the indicator I, assuming a greater indicator scalar value is preferable.

To compare several Pareto front approximations, one can be interested in defining comparison methods that capture the Pareto dominance in the objective space, i.e. given two Pareto front approximations $Y_N^1, Y_N^2 \in \Omega$ provided by Algorithms 1 and 2,

Y_N^1 weakly dominates/strictly dominates/is better than $Y_N^2 \Rightarrow \left(C_{I,E}(Y_N^1, Y_N^2) \text{ is true}\right)$.

More precisely, good comparison methods should always be compliant with the $\triangleleft$- relation (Zitzler et al., 2003). On a given problem, Algorithm 1 should not be considered as less performant than Algorithm 2 if Y_N^1 is better than Y_N^2. The following definition summarizes these points:

Definition 9 (Compatibility and completeness Zitzler et al., 2003)**.** Let $\mathcal{R}$ be an arbitrary binary relation on Pareto front approximations (typically, $\mathcal{R} \in \{\prec, \prec\prec, \preceq, \triangleleft\}$). The comparison method $C_{I,E}$ is denoted as *$\mathcal{R}$-compatible* if either for any Y_N^1, Y_N^2 Pareto front approximations, we have:

$$C_{I,E}(Y_N^1, Y_N^2) \Rightarrow Y_N^1\ \mathcal{R}\ Y_N^2$$

or for any Y_N^1, Y_N^2 Pareto front approximations, we have:

$$C_{I,E}(Y_N^1, Y_N^2) \Rightarrow Y_N^2\ \mathcal{R}\ Y_N^1.$$

The comparison method is denoted as *$\mathcal{R}$-complete* if either for any Y_N^1, Y_N^2 Pareto front approximations, we have:

$$Y_N^1\ \mathcal{R}\ Y_N^2 \Rightarrow\ C_{I,E}(Y_N^1, Y_N^2)$$

or for any Y_N^1, Y_N^2 Pareto front approximations, we have:

$$Y_N^2\ \mathcal{R}\ Y_N^1 \Rightarrow C_{I,E}(Y_N^1, Y_N^2).$$

For any Pareto front approximations $Y_N^1, Y_N^2 \in \Omega$, there are no combination I of unary performance indicators such that $Y_N^1 \triangleleft Y_N^2 \ \Leftrightarrow\ C_{I,E}(Y_N^1, Y_N^2)$ (Zitzler et al., 2003).

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

Remark. Throughout the rest of the paper, the following notations will be used. A discrete representation of the Pareto front is denoted by $Y_P \subseteq \mathcal{Y}_P$, called the *Pareto optimal solution set* (Okabe et al., 2003).

The Pareto front approximation at iteration k will be denoted by $Y_N(k)$. In many cases, the Pareto front is unknown. The user needs to specify a set of objective vectors in the objective space, called a *reference set* and denoted by $Y_R\ \subseteq \mathbb{R}^m$. Note that a Pareto front (approximated or not) contains only feasible objective vectors, i.e. each element of a Pareto front approximation belongs to $\mathcal{Y}$. It implies that if an algorithm does not find any feasible points then $Y_N(k)$ 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 (Jiang et al., 2014; Okabe et al., 2003; Riquelme et al., 2015):

- *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: Quantify the distribution of a Pareto front approximation. Coverage measures how well every region of the objective space is represented. Spread focuses on the aspect that points should be far away from each other (typically this drives them to the boundary). The difference is discussed in Emmerich, Deutz, and Kruisselbrink (2013a) and Kidd, Lusby, and Larsen (2020). Extent refers to a more precise property, i.e. if the Pareto front approximation contains the extreme points of the Pareto front. Uniformity only considers how well the points are equally spaced (Faulkenberg & Wiecek, 2009; Sayın, 2000). Spread and uniformity properties constitute the diversity of a Pareto front approximation (Jiang et al., 2014).
- *Convergence and distribution indicators* 3.4: Capture both the properties of convergence and distribution.

3.1. Cardinality indicators

These indicators 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) (Van Veldhuizen, 1999)

ONVG returns the number of elements in the Pareto front approximation generated by the algorithm:

$$\text{for all } Y_N \in \Omega,\ \ ONVG(Y_N) = |Y_N|.$$

This indicator is to be maximized. Nonetheless, this is not a pertinent measure. For example, consider a Pareto front approximation Y_N^1 composed of one million non-dominated points and a Pareto front approximation Y_N^2 with only one point, such as this point dominates all the other points of Y_N^1. Y_N^1 outperforms Y_N^2 for this indicator but Y_N^2 is clearly better than Y_N^1 (Knowles & Corne, 2002).

3.1.2. Overall Non-dominated vector generation ratio (ONVGR) (Van Veldhuizen, 1999)

ONVGR represents the ratio of a number of elements in the Pareto front approximation with respect to a Pareto optimal solution set. Formally,

$$ONVGR(Y_N; Y_P) = \frac{|Y_N|}{|Y_P|}$$

where $|Y_P|$ is the cardinality of a Pareto optimal solution set and $|Y_N|$ the number of points of the Pareto front approximation. A higher value is considered to be better. Note that this indicator is just *ONVG* (3.1.1) divided by a scalar. Consequently, it suffers from the same drawbacks as the previous indicator.

3.1.3. Generational indicators (GNVG, GNVGR and NVA) (Van Veldhuizen, 1999)

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

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

These indicators 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) (Van Veldhuizen, 1999)

This indicator considers the number of non-dominated objective vectors which belong to the Pareto front. It is given by the following formula:

$$E(Y_N) = \frac{1}{|Y_N|} \sum_{y \in Y_N} \mathbb{1}_{\mathbb{R}^m \setminus \mathcal{Y}_P}(y)$$

where for all $y \in \mathbb{R}^m$,

$$\mathbb{1}_{\mathbb{R}^m \setminus \mathcal{Y}_P}(y) = \begin{cases} 0 & \text{if } y \text{ belongs to the Pareto front.} \\ 1 & \text{otherwise.} \end{cases}$$

The lower the indicator value, the better it is considered.

In Van Veldhuizen (1999), the author does 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 front 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 front approximation, which can misguide interpretations. Knowles and Corne (2002) illustrate this drawback with the following example. 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, Hansen and Jaszkiewicz (1998) defines the C_{1R} metric (called also *ratio of the reference points*). Given a reference set Y_R (chosen by the user) in the objective space, it is the ratio of the number of objectives vectors found in Y_R over the cardinality of the reference set Y_R. A higher value is desirable.

3.1.5. C-metric or coverage of two sets (C) (Zitzler, 1999)

Let Y_N^1 and Y_N^2 be two Pareto front approximations. The C-metric captures the proportion of points in a Pareto front approximation Y_N^2 weakly dominated by the Pareto front approximation Y_N^1. This binary performance indicator maps the ordered pair (Y_N^1, Y_N^2) to the interval $[0,1]$ and is defined by:

$$C(Y_N^1, Y_N^2) = \frac{\left|\left\{y^2 \in Y_N^2,\ \text{there exists } y^1 \in Y_N^1 \text{ such that } y^1 \leqq y^2\right\}\right|}{\left|Y_N^2\right|}.$$

If $C(Y_N^1, Y_N^2) = 1$, all the elements of Y_N^2 are dominated by (or equal to) the elements of Y_N^1. If $C(Y_N^1, Y_N^2) = 0$, no element of Y_N^2 is weakly dominated by an element of Y_N^1. Both orderings have to be computed, as $C(Y_N^1, Y_N^2)$ is not always equal to $1 - C(Y_N^2, Y_N^1)$.

Knowles and Corne (2002) point out the limits of this measure. If $C(Y_N^1, Y_N^2) \neq 1$ and if $C(Y_N^2, Y_N^1) \neq 1$, 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 indication of "how much" a Pareto front approximation strictly dominates another.

Similarly to the C-metric, given a reference set Y_R in the objective space, the C_{2R} metric (*Ratio of non-dominated points by the reference set*) introduced in Hansen and Jaszkiewicz (1998) is given by:

$$C_{2R}(Y_N; Y_R) = \frac{|\{y \in Y_N :\ \text{there does not exist } r \in Y_R \text{ such that } r \leq y\}|}{|Y_N|}.$$

A higher C_{2R} value is considered to be better. This indicator has the same drawbacks as the C-metric.

3.1.6. Mutual domination rate (MDR) (Martí, García, Berlanga, & Molina, 2016)

The authors of Martí et al. (2016) use this performance indicator in combination with a Kalman filter to monitor the progress of evolutionary algorithms along iterations and thus providing a stopping criterion. At a given iteration k, MDR captures how many non-dominated points at iteration $k-1$ are dominated by non-dominated points generated at iteration k and reciprocally. Given two Pareto front approximations Y_N^1 and Y_N^2, the function $\Delta(Y_N^1,\ Y_N^2)$ returns the set of elements of Y_N^1 that are dominated by at least one element of Y_N^2. Formally,

$$MDR(Y_N;k) = \frac{|\Delta(Y_N(k-1),\ Y_N(k))|}{|Y_N(k-1)|} - \frac{|\Delta(Y_N(k),\ Y_N(k-1))|}{|Y_N(k)|}$$

where $Y_N(k)$ is the Pareto front approximation generated at iteration k. If $MDR(Y_N;k) = 1$, the set of non-dominated points at iteration k totally dominates its predecessor at iteration $k-1$. If $MDR(Y_N;k) = 0$, no significant progress has been observed. $MDR(Y_N;k) = -1$ 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

Most of these measures require the knowledge of the Pareto Front to be evaluated. They capture the degree of proximity between a Pareto front and its approximation.

3.2.1. Generational distance (GD) (Van Veldhuizen, 1999)

The GD indicator captures the average distance between each element of a Pareto front approximation and its closest neighbor in a discrete representation of the Pareto front. This indicator is given by the following formula:

$$GD(Y_N; Y_P) = \frac{1}{|Y_N|}\left(\sum_{y^1 \in Y_N} \min_{y^2 \in Y_P} \|y^1 - y^2\|^p\right)^{\frac{1}{p}}$$

where $|Y_N|$ is the number of points in a Pareto front approximation and $Y_P \subseteq \mathcal{Y}_P$ a discrete representation of the Pareto front. Generally, $p = 2$. In this case, it is similar to the $M_1^\star$-measure defined in Zitzler et al. (2000). When $p = 1$, it is equivalent to the γ-metric defined in Deb, Agrawal, Pratap, and Meyarivan (2000). For all these indicators, a lower value is considered to be better.

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 Collette and

Siarry (2011), it could be used as a stopping criteria. A slight variation of the generational distance $GD(Y_N(k), Y_N(k+1))$ 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) (Van Veldhuizen, 1999)

It measures the deformation of the Pareto front approximation according to a discrete representation of the Pareto front. It is given by the following formula:

$$STDGD(Y_N; Y_P) = \frac{1}{|Y_N|} \sum_{y^1 \in Y_N} \left(\min_{y^2 \in Y_P} \|y^1 - y^2\| - GD(Y_N; Y_P) \right)^2.$$

The same critics as with the generational distance apply.

3.2.3. Seven points average distance (SPAD) (Schott, 1995)

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

$$Y_R = \left\{ \left(\frac{h}{3} \max_{x \in \mathcal{X}} f_1(x), \ \frac{l}{3} \max_{x \in \mathcal{X}} f_2(x) \right)_{0 \le h, l \le 3} \right\}.$$

SPAD captures the average distance of the elements of the reference set Y_R to their closest neighbor in the Pareto front approximation. Formally,

$$SPAD(Y_N; Y_R) = \frac{1}{|Y_R|} \sum_{r \in Y_R} \min_{y \in Y_N} \|y - r\|.$$

A lower value is considered to be better.

This indicator raises same critics as above. Note that the computational cost to solve the single-objective problems $\max_{x \in \mathcal{X}} f_i(x)$ 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) (Van Veldhuizen, 1999)

This indicator defined in Van Veldhuizen (1999) is another measure that evaluates the distance between a discrete representation of the Pareto front and the Pareto front approximation obtained by a given algorithm. It corresponds to the largest minimal distance between elements of the Pareto front approximation and their closest neighbors belonging to the Pareto front. This indicator is to be minimized. It is expressed with the following formula (generally, $p = 2$):

$$MPFE(Y_N; Y_P) = \max_{y^1 \in Y_N} \left(\min_{y^2 \in Y_P} \sum_{i=1}^{m} \left| y_i^1 - y_i^2 \right|^p \right)^{\frac{1}{p}}.$$

When $Y_N \subseteq Y_P$, the value $MPFE(Y_N; Y_P)$ is zero. The indicator is not relevant, as pointed out in Knowles and Corne (2002). Consider two Pareto fronts approximations in which the first possesses only one element in the Pareto front and the second has ten elements: nine of them belong to the Pareto front and one is at some positive distance from it. As *MPFE* considers only largest minimal distances, it favors the first Pareto front approximation. But the second is clearly better. However, it is straightforward and cheap to compute and may be used on continuous and discontinuous problems.

3.2.5. Progress metric (P_g) (Van Veldhuizen, 1999)

This indicator introduced in Bäck (1996) measures the progression of the Pareto front approximation given by an algorithm towards the Pareto front in function of the number of iterations for a given objective function i. It is defined by:

$$P_g(Y_N; i, k) = \ln \sqrt{\frac{\min_{y \in Y_N(0)} y_i}{\min_{y \in Y_N(k)} y_i}}.$$

In Van Veldhuizen (1999), the author modifies this metric to take into account whole Pareto front approximations:

$$RP_g(Y_N; Y_P, k) = \ln \sqrt{\frac{GD(Y_N(0); Y_P)}{GD(Y_N(k); Y_P)}}$$

where $GD(Y_N(k); Y_P)$ is the generational distance (3.2.1) of the Pareto front approximation Y_N at iteration k.

P_g is not always defined, for example when values of $\min_{y \in Y_N(0)} y_i$ or $\min_{y \in Y_N(k)} y_i$ are negative or null. As GD is always 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.

3.2.6. ϵ-indicator (I_ϵ) (Zitzler et al., 2003)

An objective vector $y^1 \in \mathbb{R}^m$ is ϵ-dominating, for $\epsilon > 0$, an objective vector $y^2 \in \mathbb{R}^m$ if:

$$\text{for all } i = 1, 2, \ldots, m, \quad y_i^1 \le \epsilon \ y_i^2.$$

The multiplicative ϵ-indicator for two Pareto front approximations Y_N^1 and Y_N^2 is defined as the minimum factor one has to multiply a Pareto front approximation to make it weakly dominate another one. It is given by

$$\begin{aligned} & I_{\epsilon_\times}(Y_N^1, Y_N^2) \\ & = \inf_{\epsilon > 0} \left\{ y^2 \in Y_N^2 : \ \exists y^1 \in Y_N^1 \text{ such that } y^1 \text{ is } \epsilon\text{-dominating } y^2 \right\}. \end{aligned}$$

It can be calculated the following way:

$$I_{\epsilon_\times}(Y_N^1, Y_N^2) = \max_{y^2 \in Y_N^2} \min_{y^1 \in Y_N^1} \max_{1 \le i \le m} \frac{y_i^1}{y_i^2}.$$

Given a discrete representation of the Pareto front Y_P, the unary metric $I_{\epsilon_\times}(Y_N; Y_P)$ (with a semicolon) is defined as $I_{\epsilon_\times}(Y_P, Y_N)$.

Similarly, Zitzler et al. (2003) define an additive ϵ-indicator based on the following additive ϵ-domination. It is said that an objective vector y^1 is additively ϵ-dominating an objective vector y^2 for $\epsilon > 0$ if for all $i = 1, 2, \ldots, m$, $y_i^1 \le \epsilon + y_i^2$. This indicator is then calculated by:

$$I_{\epsilon_+}(Y_N^1, Y_N^2) = \max_{y^2 \in Y_N^2} \min_{y^1 \in Y_N^1} \max_{1 \le i \le m} \ y_i^1 - y_i^2.$$

A value inferior to 1 (respectively 0) for the binary multiplicative ϵ-indicator (respectively additive ϵ-indicator) indicates that Y_N^1 weakly dominates Y_N^2.

Binary additive and multiplicative ϵ-indicators possess desirable properties. They are Pareto compliant and compatible (Zitzler et al., 2003), do not require the knowledge of the Pareto front, and represent natural extensions for approximation schemes in optimization theory. However, the main problem with the ϵ-indicator is that its evaluation value involves only one particular element in each

Pareto front approximation, which can misguide quality comparison between different Pareto front approximations. Furthermore, the ϵ-indicator focuses only on one objective when comparing different objective vectors, as noticed in Li and Yao (2017). For example, consider $y^1 = (0, 1, 1)$ and $y^2 = (1, 0, 0)$ in a tri-objective maximization problem. Although y^1 performs better than y^2 in two different objectives, the additive ϵ-indicators are identical:

$$I_{\epsilon_+}\left(\{y^1\}, \{y^2\}\right) = I_{\epsilon_+}\left(\{y^2\}, \{y^1\}\right).$$

On the contrary, it is straightforward to compute. It can be used for continuous and discontinuous approximations of Pareto fronts.

3.2.7. Degree of approximation (DOA) (Dilettoso, Rizzo, & Salerno, 2017)

By taking into account the dominance relation in the objective space, DOA captures an average degree of proximity from a discrete representation of the Pareto front to a Pareto front approximation. A lower value is considered to be better. This indicator is proved to be $\prec$-complete (see Definition 9). It aims to compare algorithms when the Pareto fronts are known.

Given y^2 an objective vector belonging to Y_P, the set $\mathcal{D}(y^2, Y_N)$ in the objective space is defined as the subset of points belonging to the Pareto front approximation Y_N dominated by the objective vector y^2. The minimum Euclidean distance between $y^2 \in Y_P$ and $\mathcal{D}(y^2, Y_N)$ (which may be empty) is computed with

$$d(y^2, Y_N) = \begin{cases} \min\limits_{y^1 \in \mathcal{D}(y^2, Y_N)} \|y^2 - y^1\| & \text{if } |\mathcal{D}(y^2, Y_N)| > 0 \\ \infty & \text{if } |\mathcal{D}(y^2, Y_N)| = 0. \end{cases}$$

Similarly, $d^+\left(y^2, Y_N \setminus \mathcal{D}(y^2, Y_N)\right)$ is defined for $y^2 \in Y_P$ by considering the set of points of Y_N that do not belong to $\mathcal{D}(y^2, Y_N)$ as:

$$d^+\left(y^2, Y_N \setminus \mathcal{D}(y^2, Y_N)\right) = \begin{cases} \min\limits_{y^1 \in Y_N \setminus \mathcal{D}(y^2, Y_N)} \|(y^1 - y^2)_+\| & \text{if } |Y_N \setminus \mathcal{D}(y^2, Y_N)| > 0 \\ \infty & \text{if } |Y_N \setminus \mathcal{D}(y^2, Y_N)| = 0 \end{cases}$$

where $(y^1 - y^2)_+ = (\max(0, y^1_i - y^2_i))_{i=1,2,\ldots,m}$.

The *DOA* indicator is finally given by

$$DOA(Y_N; Y_P) = \frac{1}{|Y_P|} \sum_{y^2 \in Y_P} \min\left\{d(y^2, Y_N), d^+\left(y^2, Y_N \setminus \mathcal{D}(y^2, Y_N)\right)\right\}.$$

The value of *DOA* does not depend on the number of points of Y_P, i.e. if $|Y_P| \gg |Y_N|$ (Dilettoso et al., 2017). In fact, this indicator partitions Y_N into subsets in which each element is dominated by a point $y \in Y_P$. Its computational cost is quite low (in $\mathcal{O}(m\ |Y_N| \times |Y_P|)$). It can be used for discontinuous and continuous approximations of Pareto fronts.

3.3. Distribution and spread indicators

According to Custódio et al. (2011), *"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 corresponding to distinct objective vectors.

3.3.1. Spacing (SP) (Schott, 1995)

The *SP* indicator captures the variation of the distance between elements of a Pareto front approximation. A lower value is considered to be better. This indicator is computed with

$$SP(Y_N) = \sqrt{\frac{1}{|Y_N| - 1} \sum_{j=1}^{|Y_N|} \left(\bar{d} - d^1\left(y^j,\ Y_N \setminus \{y^j\}\right)\right)^2}$$

where $d^1\left(y^j,\ Y_N \setminus \{y^j\}\right) = \min_{y \in Y_N \setminus \{y^j\}} \|y - y^j\|_1$ is the l_1 distance of $y^j \in Y_N$ to the set $Y_N \setminus \{y^j\}$ and $\bar{d}$ is the mean of all $d^1(y^j, Y_N \setminus \{y^j\})$ for $j = 1, 2, \ldots, |Y_N|$.

This method cannot account for holes in the Pareto front approximation as it takes into account the distance between an objective vector and its closest neighbor. The major issue with this indicator 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 $\Delta^\star$) (Deb et al., 2000; Zhou, Jin, Zhang, Sendhoff, & Tsang, 2006)

Deb et al. (2000) introduce the Δ' index for biobjective problems, which captures the variation of distance between consecutive elements of the Pareto front approximation into the biobjective space. Formally,

$$\Delta'(Y_N) = \sum_{j=1}^{|Y_N|-1} \frac{\left|d^c\left(y^j, Y_N \setminus \{y^j\}\right) - \bar{d}^c\right|}{|Y_N| - 1}$$

where $d^c\left(y^j, Y_N \setminus \{y^j\}\right)$ is the Euclidean distance between consecutive elements of the Pareto front approximation Y_N, and $\bar{d}^c$ the mean of the $d^c\left(y^j, Y_N \setminus \{y^j\}\right)$ for $j = 1, 2, \ldots, |Y_N| - 1$. As this indicator considers Euclidean distances between consecutive objective vectors, 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 $d^c\left(y^j, Y_N \setminus \{y^j\}\right)$. 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:

$$\Delta(Y_N; Y_P) = \frac{\sum_{i=1}^{2} \min\limits_{y \in Y_N} \|y^{i,\star} - y\| + \sum_{j=1}^{|Y_N|-1} \left|d^c\left(y^j, Y_N \setminus \{y^j\}\right) - \bar{d}^c\right|}{\sum_{k=1}^{2} \min\limits_{y \in Y_N} \|y^{k,\star} - y\| + (|Y_N| - 1)\bar{d}^c}$$

where $\min_{y \in Y_N} \|y^{i,\star} - y\|$ for $i = 1, 2$ are the Euclidean distances between the extreme solutions of the Pareto front (i.e. $y^{i,\star} = F(x^{i,\star}) \in Y_P$ where $x^{i,\star}$ is solution to the ith single-objective problem) and the boundary solutions of the Pareto front approximation. The other notations remain the same as before. This indicator requires the resolution of each single-objective optimization problem. This indicator is extended to Pareto fronts with more than two objectives by Zhou et al. (2006) to the generalized $\Delta^\star$-index:

$$\Delta^\star(Y_N; Y_P) = \frac{\sum_{i=1}^{m} d^2(y^{i,\star}, Y_N) + \sum_{j=1}^{|Y_N|} \left|d^2(y^j, Y_N \setminus \{y^j\}) - \bar{d}^2\right|}{\sum_{i=1}^{m} d^2(y^{i,\star}, Y_N) + |Y_N|\ \bar{d}^2}$$

where $d^2(y^{i,\star}, Y_N) = \min_{y \in Y_N} \|y^{i,\star} - y\|$ with $y^{i,\star} = F(x^{i,\star}) \in Y_P$ the extreme objective vector corresponding to $x^{i,\star}$ solution to the ith single-objective problem and $d^2(y^j, Y_N \setminus \{y^j\}) = \min_{y \in Y_N \setminus \{y^j\}} \|y^j - y\|$ the minimal Euclidean distance between two points of the Pareto front approximation. $\bar{d}^2$ is the mean of all $d^2(y^j, Y_N \setminus \{y^j\})$ for $j = 1, 2, \ldots, |Y_N|$. As it considers 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.

For these three indicators, a lower value is considered to be better.

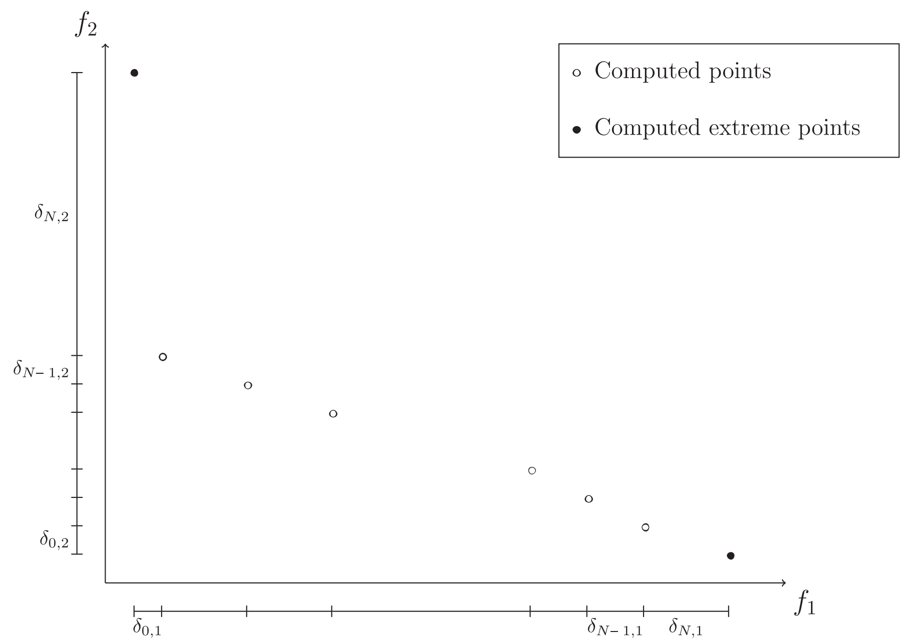

Fig. 2. Illustration of the Γ metric for a biobjective problem (inspired by Custódio et al., 2011).

3.3.3. Two measures proposed by Custódio et al. (2011) (Γ and Δ)

Given a Pareto set approximation $X_N = \{x_1, x_2, \dots, x_N\}$ to which two additional "extreme" points indexed by 0 and $N+1$ are added, for each objective i for $i = 1, 2, \dots, m$, elements x^j for $j = 0, 1, \dots, N+1$ of the Pareto set approximation are sorted so that,

$$f_i(x^0) \leq f_i(x^1) \leq f_i(x^2) \leq \ldots \leq f_i(x^{N+1}).$$

Custódio et al. (2011) introduce the following metric $\Gamma > 0$ defined by:

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

where $\delta_{i,j} = f_i(x^{j+1}) - f_i(x^j)$ and $Y_N = F(X_N)$. 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 Fig. 2.

To take into account the extent of the Pareto front approximation, the authors of Custódio et al. (2011) define the following indicator by

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

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

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. For these two indicators, a lower value is desirable. These two metrics are more adapted to continuous Pareto front approximations.

Remark. The authors of Custódio et al. (2011) suggest two ways to compute extreme points x_0 and x_{N+1}. For benchmark tests, the Pareto front is known and extreme points correspond to the ones of the Pareto set. Otherwise, the Γ and Δ indicators use the extreme points of the Pareto set approximation X_N.

3.3.4. Hole relative size (HRS) (Collette & Siarry, 2005)

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

$$HRS(Y_N) = (1/\bar{d}) \max_{j=1,2,\dots,|Y_N|-1} d^j$$

where Y_N is a Pareto front approximation whose elements have been sorted in ascendant order according to the first objective, $d^j = \|y^j - y^{j+1}\|_2$ is the l_2 distance between the two adjacent objective vectors $y^j \in Y_N$ and $y^{j+1} \in Y_N$ and $\bar{d}$ the mean of all d^j for $j = 1, 2, \dots, |Y_N| - 1$.

A lower indicator value is desirable. As it takes into account holes in the objective space, this indicator is more adapted to continuous Pareto front approximations.

3.3.5. Zitzler's metrics $M_2^\star$ and $M_3^\star$ (Zitzler et al., 2000)

The $M_2^\star$ metric returns a value in the interval $[0; |Y_N|]$ where Y_N is the Pareto front approximation. It reflects the number of subsets of the Pareto front approximation Y_N of a certain radius (σ). A higher value is considered to be better. Its expression is given by

$$M_2^\star(Y_N; \sigma) = \frac{1}{|Y_N| - 1} \sum_{y^2 \in Y_N} |\{y^1 \in Y_N, \|y^2 - y^1\| > \sigma\}|.$$

If $M_2^\star(Y_N; \sigma) = |Y_N|$, 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 Tan, Lee, and Khor (2002) introduce the Uniform distribution indicator, based too on the search of niches of size σ, given by

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

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

$$D_{nc}(Y_N, \sigma) = \sqrt{\frac{1}{|Y_N| - 1} \left(\sum_{j=1}^{|Y_N|} \left(nc(y^j, \sigma) - \frac{1}{|Y_N|} \sum_{l=1}^{|Y_N|} nc(y^l, \sigma) \right)^2 \right)}$$

with $nc(y^j, \sigma) = \left|\{y \in Y_N, \|y - y^j\| < \sigma\}\right| - 1$. The UD indicator is to be minimized.

Finally, the $M_3^\star$ metric defined by Zitzler et al. (2000), considers the extent of the front:

$$M_3^\star(Y_N) = \sqrt{\sum_{i=1}^{m} \max_{j \in \{1,2,\ldots,|Y_N|\}} \max_{y \in Y_N \setminus \{y^j\}} \|y^j - y\|}.$$

A higher value is considered to be better. 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) (Sayın, 2000)

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 distribution of the points along the Pareto front approximation.

3.3.7. Evenness (ξ) (Messac & Mattson, 2004)

The ξ-evenness indicator captures the uniformity quality of a Pareto front approximation by integrating distance values between its elements into a coefficient of variation. More specifically, two scalar values are associated to each element $y \in Y_N$ of the Pareto front approximation. $d^l(y, Y_N \setminus \{y\})$ is the minimum Euclidean distance between objective vector y and its closest neighbor in the objective space. $d^u(y, Y_N \setminus \{y\})$ is the maximal Euclidean distance between an element $y \in Y_N$ and another element of Y_N such that no other point of Y_N is within the (hyper)sphere of diameter $d^u(y, Y_N \setminus \{y\})$. ξ is then defined as

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

where σ_D and $\widehat{D}$ are respectively the standard deviation and the mean of the set of minimum distances $d^l(y, Y_N \setminus \{y\})$ and maximum diameters $d^u(y, Y_N \setminus \{y\})$ for each element y of Y_N. 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 can account for holes in the Pareto front approximation.

Reference Ghosh and Chakraborty (2015) also defines the evenness as

$$E(Y_N) = \frac{\max\limits_{y^1 \in Y_N} \min\limits_{y^2 \in Y_N \setminus \{y^1\}} \|y^1 - y^2\|}{\min\limits_{y^1 \in Y_N} \min\limits_{y^2 \in Y_N \setminus \{y^1\}} \|y^1 - y^2\|}.$$

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

3.3.8. Binary uniformity (SP_l) (Meng, Zhang, & Liu, 2005)

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

Let Y_N^1 and Y_N^2 be two Pareto front approximations. The algorithm is decomposed in several steps:

Set $l = 1$.

1. For each element $y^1 \in Y_N^1$ and $y^2 \in Y_N^2$, compute the respective distances to their closest neighbor $d_l^2(y^1, Y_N^1 \setminus \{y^1\})$ and $d_l^2(y^2, Y_N^2 \setminus \{y^2\})$ given by
$$d_l^2(y, Y_N \setminus \{y\}) = \min_{y^v \in Y_N \setminus \{y\}} \|y - y^v\|.$$
2. For both sets, compute the average distances $\overline{d_l^2}(Y_N^1)$ and $\overline{d_l^2}(Y_N^2)$ between neighbor points given by
$$\overline{d_l^2}(Y_N) = \frac{1}{|Y_N|} \sum_{y \in Y_N} d_l^2(y, Y_N \setminus \{y\}).$$
3. For each set, compute the spacing measures $SP^l(Y_N^1)$ and $SP^l(Y_N^2)$ given by
$$SP^l(Y_N) = \sqrt{\sum_{y \in Y_N} \frac{\left(1 - \psi\left(d_l^2(y, Y_N \setminus \{y\}), \overline{d_l^2}(Y_N)\right)\right)^2}{|Y_N| - 1}}$$
with $\psi(a, b) = \begin{cases} \frac{a}{b} & \text{if } a > b \\ \frac{b}{a} & \text{otherwise.} \end{cases}$
4. If $SP^l(Y_N^1) < SP^l(Y_N^2)$, then Y_N^1 has better uniformity than Y_N^2 and reciprocally. If $SP^l(Y_N^1) = SP^l(Y_N^2)$ and $l \geq \min(|Y_N^1| - 1, |Y_N^2| - 1)$ then Y_N^1 has the same uniformity as Y_N^2. Else if $SP^l(Y_N^1) = SP^l(Y_N^2)$ and $l < \min(|Y_N^1| - 1, |Y_N^2| - 1)$, then increment l by 1, and recompute the previous steps by removing the smallest distance $d_l^2(y^1, Y_N^1 \setminus \{y^1\})$ and $d_l^2(y^2, Y_N^2 \setminus \{y^2\})$ 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) (Leung & Wang, 2003)

This indicator measures the uniformity of a Pareto front approximation based on distance between its elements according to each objective. For each objective vector in the Pareto front approximation Y_N, Euclidean distance to their nearest neighbors with respect to each objective axis is determined, as well as distances of reference objectives (objective vectors corresponding to solutions of the single-objective optimization problems or determined by the user) to their nearest neighbors. Let χ be the set of distances between nearest neighbors and $\bar{\chi}^0$ the set of distances between reference objective vectors and their closest neighbor. Small variability in the set χ would reflect uniform distribution, and values close to 0 for the set $\bar{\chi}^0$ would reflect good coverage properties. For ease of computation, each distance in $\bar{\chi}^0$ is summed up with the average value of the distances in χ, resulting in the new set $\bar{\chi}$. The U-measure captures the discrepancy among the scalar elements of the set $\bar{\chi}$ and is given by

$$U(Y_N) = \frac{1}{|\bar{\chi}|} \sum_{d' \in \bar{\chi}} \left| \frac{d'}{d_{\text{ideal}}} - 1 \right|$$

where $d_{\text{ideal}} = \frac{1}{|\bar{\chi}|} \sum_{d' \in \bar{\chi}} d'$.

$\frac{d'}{d_{\text{ideal}}} - 1$ 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 elements of the set $\bar{\chi}$. A small U indicator value 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 indicators can be raised. Especially, the formula works only if there are several points. Moreover, this indicator can take time to compute when computing the minimal distances. As for the spacing metric (3.3.1), 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.

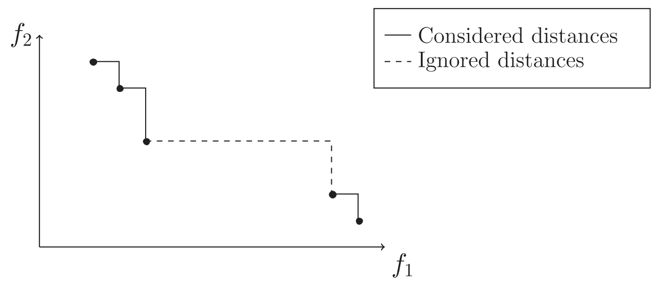

Fig. 3. An example showing the weaknesses of the spacing metric (inspired by Zheng et al., 2017): the spacing metric ignores the gap drawn in dashed lines.

3.3.10. Overall Pareto spread (OS) (Wu & Azarm, 2000)

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

$$OS(Y_N) = \prod_{i=1}^{m} \frac{\left| \max_{y \in Y_N} y_i - \min_{y \in Y_N} y_i \right|}{\left| \tilde{y}_i^I - \tilde{y}_i^M \right|}$$

where $\tilde{y}^M$ is an approximation of the maximum objective vector (objective vector composed of the maxima of each single-objective optimization problem assuming they exist) or the maximum objective vector and $\tilde{y}^I$ an approximation of the ideal objective vector or the ideal objective vector.

This is an indicator for which the values are among the values 0 and 1. The maximum and ideal objective vectors need to be computed for more precise results, resulting in 2 m single-objective problems to solve. The indicator does not take into account the distribution of points along the Pareto front approximation.

3.3.11. Outer diameter (I_{OD}) (Zitzler et al., 2008)

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. A higher indicator value is desirable. It is given by:

$$I_{OD}(Y_N) = \max_{1 \le i \le m} w_i \left(\max_{y \in Y_N} y_i - \min_{y \in Y_N} y_i \right).$$

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 extent of the Pareto front approximation. By the way, it can result in an information loss of the extent of the Pareto front approximation, as it focuses only on the largest distance along a single dimension.

3.3.12. Distribution metric (DM) (Zheng, Yang, Xu, & Hu, 2017)

This indicator introduced by Zheng et al. (2017) aims to correct several drawbacks of the spacing measure (Schott, 1995) 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 drawbacks of the spacing metric (3.3.1) is given in Fig. 3.

The DM indicator is given by

$$DM(Y_N) = \frac{1}{|Y_N|} \sum_{i=1}^{m} \left(\frac{\sigma_i}{\mu_i} \right) \left(\frac{\left| y_i^I - y_i^N \right|}{\max_{y \in Y_N} y_i - \min_{y \in Y_N} y_i} \right)$$

with $\sigma_i = \frac{1}{|Y_N|-2} \sum_{j=1}^{|Y_N|-1} \left(d_i^j - \overline{d_i} \right)^2$, $\mu_i = \frac{1}{|Y_N|-1} \sum_{j=1}^{|Y_N|-1} d_i^j$ where $|Y_N|$ is the number of non-dominated objective vectors, y^I and y^N are respectively the ideal and nadir objective vectors. d_i^j is the distance of the jth 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 $\frac{\sigma_i}{\mu_i}$ is the coefficient of variance relative to the ith objective.

A smaller DM indicates better distributed solutions. It takes into account the extent and distribution of the points along the Pareto front approximation. However, it requires the nadir and ideal objective vectors, 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) (Li, Zheng, & Xiao, 2008)

The I_D indicator measures the variation of distances between elements of a Pareto front approximation based on the construction of a minimum spanning tree. The indicator value is comprised between 0 and 1. The closest to 1, the better. Let Y_N be a Pareto front approximation such that $|Y_N| > 2$. The computation of this indicator is decomposed into several steps:

1. A minimum spanning tree T_G covering all the elements of Y_N based on the Euclidean distance in the objective space is built.
2. Each element $y \in Y_N$ has at least one neighbor in the spanning set, i.e a vertex adjacent to y. Let $N_{T_G}(y)$ be the set of adjacent vertices to y in the spanning tree T_G.
   For each $y^\nu \in N_{T_G}(y)$, we define a "neighborhood" (Li et al., 2008)
   $$N_{y^\nu}(y) = \left\{ y^2 \in Y_N, \|y^2 - y\| \le \|y^\nu - y\| \right\}$$
   which corresponds to the subset of Y_N contained in the closed ball of radius $\|y^\nu - y\|$ and centered in y. Note that $\{y, y^\nu\} \in N_{y^\nu}(y)$. The neighborhoods that contain only two elements, i.e. y and y^ν, are not considered.
3. For all $y \in Y_N$ and $y^\nu \in N_{T_G}(y)$, a distribution relation is defined by
   $$\psi(y, y^\nu) = \begin{cases} 0 & \text{if } |N_{y^\nu}(y)| = 2, \\ \prod_{y^2 \in N_{y^\nu}(y) \setminus \{y\}} \frac{\|y - y^2\|}{\|y - y^\nu\|} & \text{otherwise.} \end{cases}$$
4. There are $2|Y_N| - 2$ neighborhoods. Among them, N_r corresponds to the number of neighborhoods that only contain two elements. The uniform assessment metric is then defined by
   $$I_D(Y_N) = \frac{1}{2|Y_N| - N_r - 2} \sum_{y \in Y_N} \sum_{y^\nu \in N_{T_G}(y)} \psi(y, y^\nu)$$
   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.

3.3.14. Extension measure (EX) (Meng et al., 2005)

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

$$EX(Y_N; Y_P) = \frac{1}{m} \sqrt{\sum_{i=1}^{m} d^2(y^{i,\star}, Y_N)^2}$$

where $d^2(y^{i,\star}, Y_N)$ is the minimal Euclidean distance between the objective vector corresponding to 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) (Cai, Sun, & Fan, 2018)

As its name indicates, this indicator uses reference vectors in the objective space to measure the diversity of a Pareto front approximation. The lower this indicator is, the better. Let $Y_R = \{r^1, r^2, \ldots, r^{|Y_R|}\}$ be a set of uniformly generated reference vectors in $\mathbb{R}^m$. For each element of a Pareto front approximation $y \in Y_N$, the closeness between y and the reference vector r^j, for $j = 1, 2, \ldots, |Y_R|$, is given by

$$\text{angle}(r^j, y) = \arccos \frac{(r^j)^T (y - y^I)}{\|r^j\| \, \|y - y^I\|}.$$

If a reference vector r^j is the closest to an element y of Y_N relatively to the closeness metric, it is said that y *"covers the reference vector r^j"* (Cai et al., 2018). The coverage vector c of size $|Y_N|$ represents for each $y \in Y_N$ the number of reference vectors that y covers. *DIR* is the normalized standard deviation of the coverage vector c, defined as

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

where $\bar{c}$ is the mean of all c_i for $i = 1, 2, \ldots, |Y_N|$. It is intuitive to understand and cheap to compute (in $\mathcal{O}(m\,|Y_N| \times |Y_R|)$ Cai et al., 2018). 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 $|Y_N|$ to be more pertinent) equally plays an important role. It can be biased when the Pareto front is piecewise continuous.

3.3.16. The Riesz s-energy indicator (E_s) (Falcón-Cardona, Coello, & Emmerich, 2019; Hardin & Saff, 2004)

The Riesz s-energy indicator (Hardin & Saff, 2004) aims at quantifying a good distribution of points in d-dimensional manifolds. Given a Pareto front approximation Y_N, this indicator is defined as:

$$E_s(Y_N) = \sum_{y^1 \in Y_N} \sum_{y^2 \in Y_N \setminus \{y^1\}} \frac{1}{\|y^1 - y^2\|^s}$$

where $s > 0$ is a fixed external parameter which controls the degree of uniformity of the elements of Y_N.

An uniformly distributed Pareto front approximation must have a minimal Riesz s-energy value. In Hardin and Saff (2005), it is proved that configurations of points in a rectifiable $d-$dimensional manifold that have minimum Riesz s-energy possess asymptotically uniformly distribution properties for $s \geq d$. Moreover, s does not depend on the shape of the manifold (Hardin & Saff, 2005).

Use of the Riesz s-energy indicator to assess generation of an uniformly distributed Pareto front approximation can be found in Falcón-Cardona et al. (2019).

3.3.17. Laumanns metric (I_L) (Laumanns, Günter, & Schwefel, 1999; Laumanns, Zitzler, & Thiele, 2000)

The Laumanns metric measures the normalized volume of the objective space dominated by a Pareto front approximation and bounded above by an approximated nadir objective vector. Given a vector y in the objective space $\mathbb{R}^m$, let $\mathcal{D}(y) = \left\{ y^2 \in \mathbb{R}^m, \; y \leq y^2 \right\}$

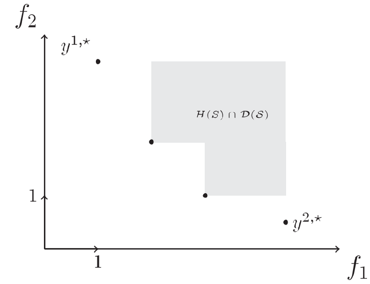

Fig. 4. The intersection of $H(S)$ and $\mathcal{D}(S)$ for a biobjective minimization problem.

be the set of objective vectors dominated by y. Given a Pareto front approximation Y_N, $\mathcal{D}(Y_N)$ is designed as the dominated space by the set Y_N and is defined as

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

Let $y^{j,\star}$ be the jth outer point of the Pareto front approximation Y_N defined by: for all $i = 1, 2, \ldots, m$,

$$(y_i^{j,\star}) = \begin{cases} \max\limits_{y \in Y_N} y_i & \text{if } i \neq j, \\ \min\limits_{y \in Y_N} y_i & \text{otherwise.} \end{cases}$$

We introduce the hypercube $H(Y_N) = \{y \in \mathbb{R}^m : y = y^I + \sum_{i=1}^m a_i (y^{i,\star} - y^I), a_i \in [0, \;\; 1]\}$ where y^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:

$$I_L(Y_N) = \frac{\lambda_m(\mathcal{D}(Y_N) \cap H(Y_N)))}{\lambda_m(H(Y_N))}$$

where $\lambda_m(A)$ is the m-dimensional Lebesgue measure of the bounded set $A \subset \mathbb{R}^m$. The metric returns a value between 0 and 1. The higher the better. An illustration is given in Fig. 4.

This indicator is biased in favor of convex and extended fronts. Moreover, its computational complexity in $\mathcal{O}(|Y_N|^{\frac{m}{3}} \text{poly} \log |Y_N|)$ Chan (2013) explodes when the objective space dimension increases: in fact, it is similar to the hypervolume indicator (3.4.7) when the reference point is chosen such as $\widetilde{y}^N$.

Similarly, the convex hull surface indicator (Zhao et al., 2018) measures the volume of the convex hull formed by the Pareto front approximation and a reference point $r \in \mathbb{R}^m$ dominated by all the elements of the Pareto front approximation. The greater the value of this indicator is, the better. Computational cost increases exponentially with the number of objectives. As it only considers points on the boundaries of the convex hull, it is more pertinent to use it on convex Pareto front approximations.

3.3.18. Other distribution indicators

Some other indicators 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 (Farhang-Mehr & Azarm, 2004): For each element of Y_N, an influential function (a Gaussian function centered in y for $y \in Y_N$) is defined, which enables the creation of a density function considered as the sum of influential functions for each element $y \in Y_N$. 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 subset of the objective space bounded by

the nadir and ideal objective vectors is firstly normalized, then divided into boxes, whose the number is decided by the user. Based on this discretization of the objective space, the indicator is computed using the values of the density function for each center of each box and the Shannon formula of entropy (Shannon, 2001). The higher the value, the better.

2. Cluster CL_μ and Number of Distinct Choices NDC_μ (Wu & Azarm, 2000): Given two respective good (ideal objective vector) and bad (maximum objective vector assuming it exists) objective vectors $\tilde{y}^I$ and $\tilde{y}^M$, the objective space (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. For this indicator, a higher value is desirable. CL_μ is then defined as $CL_\mu(Y_N) = \frac{|Y_N|}{NDC_\mu(Y_N)}$. A higher value of the CLU_μ indicator is the consequence of a more clustered distribution of the elements of a Pareto front approximation and so a lower value is considered to be better.
3. Sigma diversity metrics σ and $\overline{\sigma}$ (Mostaghim & Teich, 2005): The objective space is divided into zones delimited by uniformly distributed reference lines starting from the origin whose the number equals $|Y_N|$. The indicator value is the ratio of the number of objective vectors that are sufficiently close to the reference lines according to the Euclidean norm with a threshold d chosen by the user, over the total number of reference lines. The higher the value, the better.
4. Diversity comparison indicator DCI (Li, Yang, & Liu, 2014): 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 (Asafuddoula, Ray, & Singh, 2015) (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 indicators 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 Pareto approximations 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. Inverted generational distance (IGD) (Coello & Cortés, 2005)

IGD has a quite similar form than GD. It captures the average minimal distance from an element of a discrete representation of the Pareto front to the closest point in the Pareto front approximation. It is given by

$$IGD(Y_N; Y_P) = \frac{1}{|Y_P|}\left(\sum_{y^2\in Y_P}\left(\min_{y^1\in Y_N}\|y^1-y^2\|\right)^p\right)^{\frac{1}{p}}.$$

Generally, $p = 2$. A lower value is considered to be better. Pros and cons are the same as for the GD indicator (3.2.1).

When an element of the Pareto front approximation Y_N does not belong to the set of nearest points to the Pareto optimal solution set Y_P, it is ignored by the IGD indicator. The authors of Tian, Zhang, Cheng, and Jin (2016) propose a variant of the IGD indicator, named IGD–NS, which takes into account these non-contributed elements in an Euclidean distance-based indicator. Let

$$Y_{NC} = \left\{ y \in Y_N : \forall y^2 \in Y_P,\ \|y-y^2\| \neq \min_{y^1\in Y_N}\|y^1-y^2\| \right\}$$

be the set of non-contributed elements of Y_N. The IGD–NS indicator is defined by

$$IGD\text{–}NS(Y_N; Y_P) = \sum_{y^2\in Y_P}\min_{y^1\in Y_N}\|y^1-y^2\| + \sum_{y^2\in Y_P}\min_{y^1\in Y_{NC}}\|y^1-y^2\|.$$

A lower indicator value is desirable.

3.4.2. Averaged Hausdorff distance (Δ_p) (Schutze, Esquivel, Lara, & Coello, 2012)

In Schutze et al. (2012), the authors combine GD (3.2.1) and IGD (3.4.1) into a new indicator, called the averaged Hausdorff distance Δ_p defined by

$$\Delta_p(Y_N; Y_P) = \max\left\{GD_p(Y_N; Y_P),\ IGD_p(Y_N; Y_P)\right\}$$

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

$$GD_p(Y_N; Y_P) = \left(\frac{1}{|Y_N|}\sum_{y^1\in Y_N}\left(\min_{y^2\in Y_P}\|y^1-y^2\|\right)^p\right)^{\frac{1}{p}}$$

and

$$IGD_p(Y_N; Y_P) = \left(\frac{1}{|Y_P|}\sum_{y^2\in Y_P}\left(\min_{y^1\in Y_N}\|y^1-y^2\|\right)^p\right)^{\frac{1}{p}}.$$

This indicator is to be minimized. It is straightforward to compute and to understand. On the contrary, it requires the knowledge of the Pareto front. The authors of Schutze et al. (2012) introduce this new indicator to correct the drawbacks of the GD and IGD indicators. It can be used to compare continuous and discontinuous approximations of Pareto fronts.

In Vargas and Bogoya (2018), the authors propose an extension of the averaged Hausdorff distance indicator, called the p, q-averaged distance $\Delta_{p,q}$ for $p, q \in \mathbb{R}\setminus\{0\}$. Given two finite sets of objective vectors $Y^1 \subset \mathbb{R}^m$, $Y^2 \subset \mathbb{R}^m$, the generational distance $GD_{p,q}$ for $p, q \in \mathbb{R}\setminus\{0\}$ is defined as

$$GD_{p,q}(Y^1, Y^2) = \left(\frac{1}{|Y^1|}\sum_{y^1\in Y^1}\left(\frac{1}{|Y^2|}\sum_{y^2\in Y^2}\|y^1-y^2\|^q\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}.$$

When $p < 0$ or $q < 0$, $GD_{p,q}$ exists if and only if $Y^1 \bigcap Y^2 = \emptyset$. Given a Pareto front approximation Y_N and a discrete representation of the Pareto front $Y_P \subseteq \mathcal{Y}_P$, the p, q-averaged distance indicator is defined as

$$\Delta_{p,q}(Y_N; Y_P) = \max\left(GD_{p,q}(Y_N, Y_P\setminus Y_N), GD_{p,q}(Y_P, Y_N\setminus Y_P)\right).$$

As the averaged Hausdorff distance, this indicator is not Pareto compliant (Vargas & Bogoya, 2018). However, once that the values of p and q are selected, it is straightforward to compute and to understand.

3.4.3. Modified inverted generational distance (IGD^+) (Ishibuchi, Masuda, Tanigaki, & Nojima, 2015)

Although the GD (3.2.1) and IGD (3.4.1) indicators are commonly used due to their low computational cost (Riquelme et al., 2015), one of their major drawbacks is that they are non monotone (Ishibuchi et al., 2015). The Δ_p indicator (3.4.2) has the same problem (Schutze et al., 2012).

Also, the authors of Ishibuchi et al. (2015) propose a slightly different version of the *IGD* indicator named IGD^+ integrating the dominance relation computable in $\mathcal{O}(m\ |Y_N| \times |Y_P|)$ where Y_P is a fixed Pareto optimal solution set. It is weakly Pareto compliant, i.e. :

for all $Y_N^1, Y_N^2 \in \Omega$ such that $Y_N^1 \preceq Y_N^2$, $IGD^+(Y_N^1; Y_P) \leq IGD^+(Y_N^2; Y_P)$.

The IGD^+ indicator is defined by

$$IGD^+(Y_N; Y_P) = \frac{1}{|Y_P|} \sum_{y^2 \in Y_P} \min_{y^1 \in Y_N} \|(y^1 - y^2)_+\|$$

where $(y^1 - y^2)_+ = (\max(0, y_i^1 - y_i^2))_{i=1,2,\ldots,m}$.

As opposed to the *IGD* indicator, IGD^+ takes into account the dominance relation between an element of Y_P and an element of Y_N when computing their Euclidean distance. A reference set Y_R can also be used instead of Y_P: the authors of Ishibuchi, Masuda, Tanigaki, and Nojima (2014) analyze the choice of such reference sets. This indicator can be used with discontinuous and continuous Pareto fronts.

Similarly to IGD^+, given a reference set $Y_R \subset \mathbb{R}^m$, Dist_{1R} (Czyzzak & Jaszkiewicz, 1998) is given by

$$\text{Dist}_{1R}(Y_N; Y_R) = \frac{1}{|Y_R|} \sum_{r \in Y_R} \min_{y \in Y_N} \max_{i=1,2,\ldots,m} \{0,\ w_i(y_i - r_i)\}$$

with w_i a relative weight assigned to objective i. For all these indicators, a lower value is desirable.

3.4.4. R_1 and R_2 indicators (Hansen & Jaszkiewicz, 1998)

Let Y_N^1 and Y_N^2 be two Pareto front approximations, $\mathcal{U}$ 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 $\mathcal{U}$. To each $u \in \mathcal{U}$ are associated $u^\star(Y_N^1) = \max_{y \in Y_N^1} u(y)$ and $u^\star(Y_N^2) = \max_{y \in Y_N^2} u(y)$. The two indicators measure to which extent Y_N^1 is better than Y_N^2 over the set of utility functions $\mathcal{U}$. The R_1 indicator is given by

$$R_1(Y_N^1, Y_N^2; \mathcal{U}, p) = \int_{u \in \mathcal{U}} C(Y_N^1, Y_N^2, u)\ p(u)\ \text{d}u$$

where

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

The R_2 indicator defined as

$$\begin{aligned} R_2(Y_N^1, Y_N^2; \mathcal{U}, p) &= E\big(u^\star(Y_N^1)\big) - E\big(u^\star(Y_N^2)\big) \\ &= \int_{u \in \mathcal{U}} \big(u^\star(Y_N^1) - u^\star(Y_N^2)\big) p(u) \text{d}u. \end{aligned}$$

is the expected difference in the utility of a Pareto front approximation Y_N^1 with another one Y_N^2. In practice, these two indicators use a discrete and finite set $\mathcal{U}$ of utility functions associated with an uniform distribution over $\mathcal{U}$ (Zitzler et al., 2008). The two indicators can then be rewritten as

$$R_1(Y_N^1, Y_N^2; \mathcal{U}) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} C(Y_N^1, Y_N^2, u) \text{ and}$$

$$R_2(Y_N^1, Y_N^2; \mathcal{U}) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} u^\star\big(Y_N^1\big) - u^\star(Y_N^2).$$

If $R_2(Y_N^1, Y_N^2; \mathcal{U}) > 0$, then Y_N^1 is considered as better than Y_N^2. Else if $R_2(Y_N^1, Y_N^2; \mathcal{U}) \geq 0$, Y_N^1 is considered as not worse than Y_N^2.

The authors of Hansen and Jaszkiewicz (1998) recommend to use the utility set $\mathcal{U}_\infty = (u_w)_{w \in W}$ of weighted Tchebycheff utility functions, with

$$u_w(y) = -\max_{i=1,2,\ldots,m} (w_i |y_i - r_i|)$$

for $y \in \mathbb{R}^m$ where r is a reference objective vector chosen so that any objective vector of the feasible objective set does not dominate r (or as an approximation of the ideal point Brockhoff, Wagner, & Trautmann, 2012; Brockhoff, Wagner, & Trautmann, 2015; Zitzler et al., 2008) and $w \in W$ a weight vector such that for all $w \in W$ and $i = 1, 2, \ldots, m$,

$$w_i \geq 0 \text{ and } \sum_{i=1}^{m} w_i = 1.$$

Zitzler et al. (2008) suggest using the set of augmented weighted Tchebycheff utility functions defined by

$$u_w(y) = -\left(\max_{i=1,2,\ldots,m} w_i |y_i - r_i| + \rho \sum_{i=1}^{m} |(y_i - r_i| \right)$$

where ρ is a sufficiently small positive real number.

As given in Brockhoff et al. (2012), for $m = 2$ objectives, W can be chosen such that:

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

The I_{R_2} indicator (Brockhoff et al., 2012) is an unary indicator derived from R_2 defined as (in the case of weighted Tchebycheff utility functions)

$$I_{R_2}(Y_N; W) = \frac{1}{|W|} \sum_{w \in W} \min_{y \in Y_N} \left\{ \max_{i=1,2,\ldots,m} (w_i |y_i - r_i|) \right\}.$$

The lower this index, the better.

As Knowles and Corne (2002) remark, *"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 $\mathcal{U}$. This simply means that each utility function in $\mathcal{U}$ must be appropriately scaled with respect to the others and its relative importance"*. *R*-indicators are only monotonic, i.e. $I(Y_N^1) \geq I(Y_N^2)$ in case Y_N^1 weakly dominates Y_N^2. They do not require important computations as the number of objectives increases. The reference point has to be chosen carefully. Studies concerning the properties of the R_2 indicator can be found in Brockhoff et al. (2012), Brockhoff et al. (2015) and Wagner, Trautmann, and Brockhoff (2013).

3.4.5. G-metric (Lizarraga-Lizarraga, Hernandez-Aguirre, & Botello-Rionda, 2008)

This indicator enables to compare k Pareto front approximations based on two criteria: the repartition of their elements and the level of domination in the objective space. It is compliant with the weak dominance relation as defined above. Its computation decomposes into several steps: given k Pareto front approximations $(Y_N^1, Y_N^2, \ldots,\ Y_N^k)$:

1. Scale the values of the objective vectors corresponding to the images of the decision vectors in the k sets, i.e. extract the non-dominated objective vectors from the union $\bigcup_{j=1}^{k} Y_N^j$, then normalize all objective vectors according to the extreme values of the objective vectors of this set.

2. Group the Pareto front approximations according to their degree of dominance. In level L^1 will be put all Pareto front approximations which are not dominated by the union of the k Pareto front approximations; remove them from the considered Pareto front approximations; then in L^2, will be put the Pareto front approximations which are not dominated by the union of the remaining Pareto front approximations, and so on.
3. For each level of dominance L^q for $q = 1, 2, \dots, Q$, where Q is the number of levels, dominated points belonging in the set $\bigcup_{Y_N \in L^q} Y_N$ are removed. Each objective vector in each set of the same level possesses a zone of influence. It is a ball of radius ρ centered in this last one. The radius ρ considers distances between neighbors objective vectors (Leung & Wang, 2003) from the k Pareto front approximations. For each Pareto front approximation belonging to the same level of dominance, a measure of dispersion is computed. This last one takes into account the zone of influence that the union of non-dominated elements of the set cover in the objective space. The smaller the value, the closer the points are.
4. The G-metric associated to an Pareto front approximation is the summation of the dispersion measure of this set and the largest dispersion measure of Pareto front approximations of lower dominance degree for each level. The bigger, the better.

The computation cost is quite important (in $\mathcal{O}(k^3 \times \max_{j=1,2,\dots,k} |Y_N^j|^2)$ Lizarraga-Lizarraga et al., 2008) but the cost can be decreased when one considers a small number of Pareto front approximations. Note that this indicator highly depends on the computation of the radius ρ when defining zones of influence. This indicator can be used for continuous and discontinuous Pareto fronts, especially to compare two Pareto front approximations, in terms of dominance and distribution in the objective space.

3.4.6. Performance comparison indicator (PCI) (Li, Yang, & Liu, 2015)

The Performance Comparison indicator (Li et al., 2015) *PCI* was conceived to rectify the main drawback of the ϵ-indicator. *PCI* enables to compare k Pareto front approximations taking into account their level of dominance and their distribution in the objective space. *PCI* uses a reference set composed of all non-dominated points taken from the union of the k Pareto front approximations. Using extreme values of the reference set, all objective vectors of the k Pareto front approximations are normalized. Then *PCI* divides the set into different clusters based on a distance threshold σ. *PCI* estimates the minimum distance move of the points in the Pareto front approximations to weakly dominate all points in the clusters. A lower value reflects better closeness of the Pareto front approximation to the reference set.

This indicator possesses a quadratic computational cost and is Pareto compliant when only two Pareto front approximations are considered (Li et al., 2015). The authors propose the following choice for the threshold, i.e.

$$\sigma \approx \frac{1}{\sqrt[m-1]{|Y_R|(m-1)! - (m/2)}}$$

where $|Y_R|$ is the size of the reference set, which enables this indicator to be external-parameter-free.

The recent binary dominance move indicator (Li & Yao, 2017) *DoM* is a generalization of the *PCI* indicator, as it computes the minimum distance move from one Pareto front approximation to weakly dominate another. A polynomial algorithm is proposed in Li and Yao (2017) in the biobjective case. To the best of our knowledge, no extension of the *DoM* indicator to more objectives has been proposed yet.

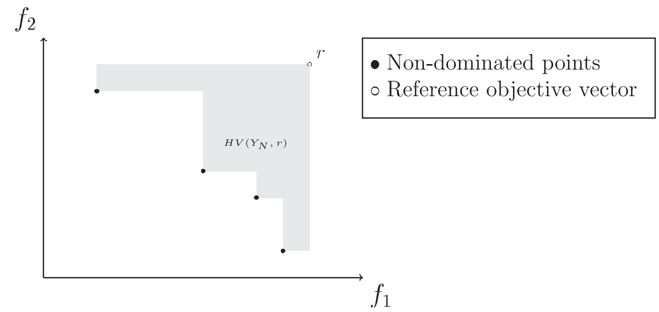

Fig. 5. Illustration of the hypervolume indicator for a biobjective problem.

3.4.7. Hyperarea/hypervolume metrics (HV) (Zitzler, 1999)

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 Y_N and delimited from above by a reference objective vector $r \in \mathbb{R}^m$ such that for all $y \in Y_N$, $y \le r$. The hypervolume indicator is given by

$$HV(Y_N; r) = \lambda_m \left(\bigcup_{y \in Y_N} [y, r] \right)$$

where λ_m is the m-dimensional Lebesgue measure. An illustration is given in Fig. 5 for the biobjective case ($m = 2$). This indicator is to be maximized.

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

$$HR(Y_N, Y_P; r) = \frac{HV(Y_N; r)}{HV(Y_P; r)}.$$

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

The hypervolume indicator and some closely related metrics are the only known unary indicators to be strictly monotonic (Falcón-Cardona, Emmerich, & Coello, 2019; Friedrich, Bringmann, Voß, & Igel, 2011; Zitzler, Brockhoff, & Thiele, 2007; Zitzler et al., 2008), i.e. if a Pareto front approximation Y_N^1 is better than another Pareto front approximation Y_N^2, $HV(Y_N^1; r) > HV(Y_N^2; r)$ assuming that all elements of the two Pareto front approximations are in the interior of the region which dominates the reference point. The best known complexity upper cost is $\mathcal{O}(|Y_N|^{\frac{m}{3}} \operatorname{poly}\log|Y_N|)$ (Chan, 2013). To the best of our knowledge, no implementation of this algorithm is available. Practically, efficient implementations of the exact hypervolume indicator can be found in Beume, Fonseca, Lopez-Ibanez, Paquete, and Vahrenhold (2009), Jaszkiewicz (2018), Lacour, Klamroth, and Fonseca (2017), Russo and Francisco (2014), Russo and Francisco (2016) and While, Bradstreet, and Barone (2012). The second drawback is the choice of the reference point, as illustrated in Fig. 6. A practical guide to specify the reference point can be found in Ishibuchi, Imada, Setoguchi, and Nojima (2018).

For the biobjective case, it was shown (Auger, Bader, Brockhoff, & Zitzler, 2009) that the optimal distribution of non-dominated points which maximizes the hypervolume indicator depends on the slope of the Pareto front. Consequently, if the Pareto front is highly nonlinear, a non-uniform Pareto front approximation may have a higher hypervolume value according to another incomparable Pareto front approximation. Other theoretical results can be found in Bringmann and Friedrich (2010a, 2010b, 2013). 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, Zitzler (1999) introduces the *Difference D of two sets* Y_N^1 and Y_N^2. $D(Y_N^1, Y_N^2)$ enables to measure the size of the area dominated by Y_N^1 but not by Y_N^2 in the objective space.

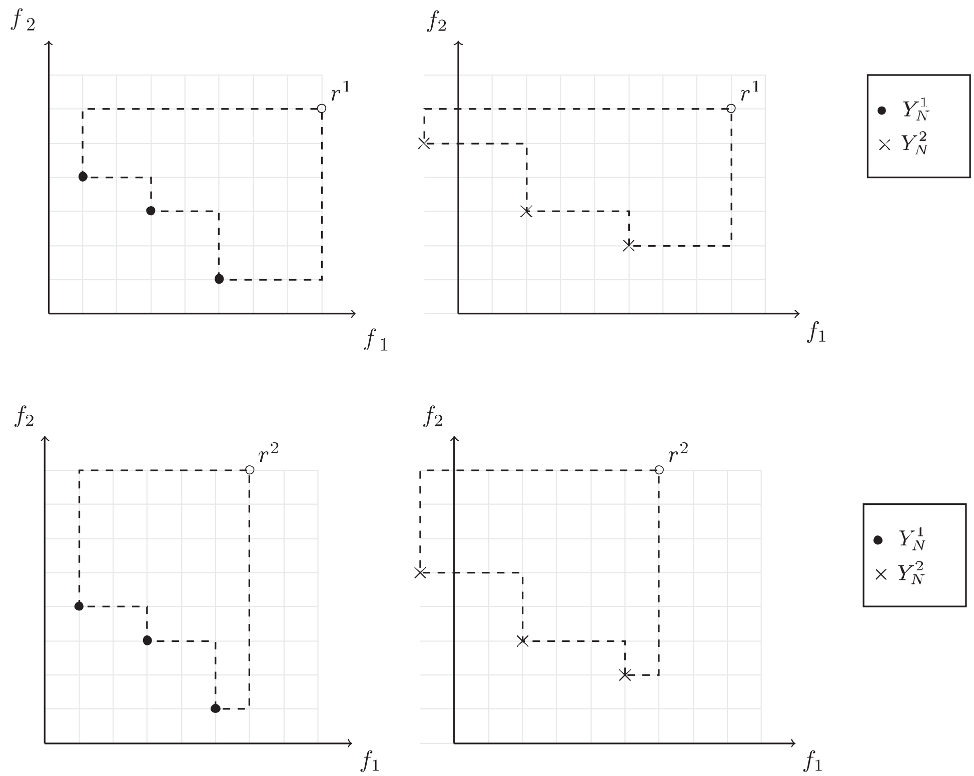

Fig. 6. The relative value of the hypervolume metric depends on the chosen reference point r^1 or r^2 (inspired by Knowles & Corne, 2002). On the top, two non-dominated Y_N^1 and Y_N^2 sets are shown, with $HV(Y_N^1; r^1) > HV(Y_N^2; r^1)$. On the bottom, $HV(Y_N^2; r^2) > HV(Y_N^1; r^2)$.

The *Hyperarea Difference* was suggested by Wu and Azarm (2000) to compensate the lack of information about the theoretical Pareto front. Given a good objective vector $\tilde{y}^I$ and a bad objective vector $\tilde{y}^M$, 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 Pareto front approximation in the objective space over the given rectangle.

More recently, a pondered hypervolume by weights indicator was introduced by Zitzler et al. (2007) to give a preference of an objective according to another. More volume indicators can be found in Wu and Azarm (2000). Some other authors (Jiang, Yang, & Li, 2016) (for biobjective optimization problems) suggest to compute the hypervolume 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 hypervolume indicator: see Auger, Bader, and Brockhoff (2010); Auger et al. (2009, 2012); Shukla, Doll, and Schmeck (2014) for more details about the properties of the hypervolume indicator). A logarithmic version of the hypervolume indicator called the logarithmic hypervolume indicator (Friedrich et al., 2011) is defined by

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

with the same notations as previously. Note that this indicator can only be used with positive vectors in $\mathbb{R}^m$. Finally, we can mention a generalization of the hypervolume indicator called the cone-based hypervolume indicator that was introduced recently by Emmerich, Deutz, Kruisselbrink, and Shukla (2013b) and an extension of the hypervolume indicator to reference point free weighted hypervolume indicators based on set monotonic and desirability functions (Emmerich, Deutz, & Yevseyeva, 2014).

3.4.8. Hypervolume Sharpe-Ratio indicator I_{HSR} (Yevseyeva, Guerreiro, Emmerich, & Fonseca, 2014)

The conception of this indicator proposed by Yevseyeva et al. (2014) lies on an analogy between the financial Portfolio Selection problem (Markowitz, 1952) and the quality of a Pareto front approximation. This analogy interprets the elements of the approximation set as assets with expected returns. An investor must allocate capital to maximize the expected return of the assembled portfolio while minimizing the expected variance (associated to risk).

To solve the financial Portfolio Selection problem, it is common to look for a strategy which maximizes the financial performance index known as reward-to-volatility ratio or Sharpe ratio (Cornuéjols, Peña, & Tütüncü, 2018). Let $A = \{a^1, a^2, \ldots, a^{|A|}\}$ be a non-empty set of assets, let $r^a \in \mathbb{R}^{|A|}$ be the expected return of these assets and $Q \in \mathbb{R}^{|A| \times |A|}$ the covariance matrix of asset returns. The Sharpe-Ratio maximization problem is defined as

$$\begin{aligned} \max_{z \in [0,1]^{|A|}} \quad & h^A(z) = \frac{(r^a)^T z - r^f}{\sqrt{z^T Q z}} \\ \text{such that} \quad & \sum_{i=1}^{|A|} z_i = 1 \end{aligned}$$

where $r^f \in \mathbb{R}$ is the return of a riskless asset. This non-linear problem which may be difficult to solve can be transformed (see Cornuéjols et al., 2018 for more details) into the following convex problem (under the condition that Q be symmetric positive)

$$\min_{w\in\mathbb{R}_+^{|A|}} \quad g^A(w) = w^T Q w$$
$$\text{such that} \quad \sum_{i=1}^{|A|}(r_i^a - r^f)w_i = 1.$$

The optimal strategy $z^\star$ of the first problem is given by $z^\star = w^\star / \left(\sum_{i=1}^{|A|} w_i^\star\right)$ where $w^\star$ is the solution of the constrained quadratic problem defined above.

The authors of Yevseyeva et al. (2014) define the Hypervolume Sharpe-Ratio indicator as

$$I_{HSR}(Y_N; y^l, y^u) = \max_{z\in\mathcal{Z}^{sr}} h^{Y_N}(z)$$

where $\mathcal{Z}^{sr} \subset [0,1]^{|Y_N|}$ is the set of feasible strategies of the Sharpe ratio maximization problem and $y^l \in \mathbb{R}^m$ and $y^u \in \mathbb{R}^m$ two reference points such that $y^l < y^u$. The definition of the covariance matrix and the asset returns inspired by the hypervolume indicator are given by

$$r_j^{|Y_N|} = \frac{\lambda_m\left([y^l, y^u] \cap [y^j, +\infty[\right)}{\lambda_m\left([y^l, y^u]\right)} = \frac{\prod_{k=1}^m \max\left(y_k^u - \max(y_k^j, y_k^l), 0\right)}{\prod_{k=1}^m (y_k^u - y_k^l)}$$

and

$$Q_{i,j} = \frac{\lambda_m\left([y^l, y^u] \cap [y^i, +\infty[\cap[y^j, +\infty[\right)}{\lambda_m\left([y^l, y^u]\right)} = \frac{\prod_{k=1}^m \max\left(y_k^u - \max(y_k^i, y_k^j, y_k^l), 0\right)}{\prod_{k=1}^m (y_k^u - y_k^l)}$$

for $i, j \in \{1, 2, \ldots, |Y_N|\}$ where λ_m is the m-dimensional Lebesgue measure. The riskless asset value is set to $r^f = 0$. The greater the indicator value is, the better.

The I_{HSR} indicator has desirable properties: providing that $y^u \leqq y < y^l$ for all $y \in Y_N$, the I_{HSR} indicator is proved to be monotonic (Guerreiro & Fonseca, 2016; 2020), but not strictly monotonic (Guerreiro & Fonseca, 2020). Other theoretical results can be found in Guerreiro and Fonseca (2016, 2020). Due to its relation with a financial model, it is interpretative. However, it is sensitive to the choice of the reference point y^u (Guerreiro & Fonseca, 2020). Its main drawback is its computational cost, directly linked to the resolution of the quadratic formulation of the Sharpe ratio maximization problem. Assuming the associated correlation matrix Q is symmetric positive, this indicator can be computed in $\mathcal{O}(|Y_N|^3)$ operations (for theoretical complexity results, the reader is invited to refer to Nesterov and Nemirovskii (1994); see also (Nocedal & Wright, 2006, chapter 16)). Practically, existing quadratic constrained solvers can efficiently compute the indicator value for a given Pareto front approximation (see Nocedal & Wright, 2006, chapter 16 for a list of quadratic solvers).

3.5. Quality assessment of Pareto set approximations in decision space

By definition, performance indicators enable to characterize properties of Pareto front approximations with respect to their diversity, spread and convergence. Moreover, it is possible to build indicators which assess the quality of Pareto set approximations in the decision space, i.e. to design mappings $I : \Psi \to \mathbb{R}$. Some research works explore the design of such indicators. For example, Zitzler et al. (2000) propose the M_1, M_2 and M_3 indicators to respectively assess convergence, diversity and extension properties of Pareto set approximations. In Sayın (2000), the author suggests using a cardinality indicator that returns the number of non-dominated points $|X_N|$, coverage indicator or uniformity indicator in the feasible set. In Ulrich, Bader, and Thiele (2010a), the authors conceive diversity indicators based on diversity preference relations in the feasible set. In Deb and Tiwari (2008), the authors define diversity crowding-distance indicator in the decision space. Indicators to take into account diversity both in the decision space and the objective space can be found in Shir, Preuss, Naujoks, and Emmerich (2009) and Ulrich, Bader, and Zitzler (2010b). Finally, in Emmerich et al. (2013a), the authors present some measures to qualify approximation sets in level set approximations, which are subsets of the feasible set.

Sayin Sayın (2000) states that the decision maker is firstly interested by the quality of the best trade-off solutions found in the objective space as long as corresponding decision variables satisfy the constraints. Furthermore, the number of objectives is usually smaller than the number of variables, which makes the Pareto front approximation easier to study/visualize.

4. Some usages of performance indicators

This section focuses on four applications of performance indicators: comparison of algorithms for multiobjective optimization, embedding performance indicators in multiobjective optimization algorithms, 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 algorithms on a multiobjective problem. In single-objective optimization, the most used graphical tools to compare algorithms include performance profiles (Dolan & Moré, 2002) and data profiles (Moré & Wild, 2009) (see also Beiranvand, Hare, & Lucet, 2017 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 et al. (2011) are the first to use data and performance profiles for multiobjective optimization. For each problem $p \in \mathcal{P}$, they build a Pareto front approximation $Y_N^p = \bigcup_{s\in\mathcal{S}} Y_N^{p,s}$ composed of the union of all Pareto front approximations $Y_N^{p,s}$ generated by each solver $s \in \mathcal{S}$ for the problem p. All dominated points are then removed. Pareto front approximations and relative Pareto optimal solution sets are then compared using cardinality and Γ and Δ metrics proposed by Custódio et al. (2011).

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 (Hansen & Jaszkiewicz, 1998; Knowles & Corne, 2002; Zitzler et al., 2008; Zitzler et al., 2003). Among them, dominance move (3.4.6), G-metric (3.4.5), binary ϵ-indicator (3.2.6), Hypervolume Sharpe-Ratio indicator (3.4.8) and volume-space metrics (3.4.7) have properties corresponding to these criteria. Mathematical proofs can be found in Auger et al. (2009), Brockhoff et al. (2012, 2015), Guerreiro and Fonseca (2016, 2020), Knowles and Corne (2002), Li and Yao (2017), Lizarraga-Lizarraga et al. (2008) and Zitzler et al. (2003)) and are synthesized in Appendices. An example of performance profile using the hypervolume indicator (3.4.7) can be found in Liuzzi, Lucidi, and Rinaldi (2016). The use of performance indicators such as *GD* (3.2.1) or *IGD* (3.4.1) as it is done in Al-Dujaili and Suresh (2016) and Brockhoff et al. (2015) is not a pertinent choice due to their inability to capture dominance relation. Instead, we suggest to use their weakly monotonic counterpart *IGD*$^+$ (3.4.3) or *DOA* (3.2.7), that can be cheaper to compute than for example the hypervolume indicator when the number of objectives is high. It is equally possible to build Pareto-compliant indicators by considering a combination of weakly Pareto compliant indicators with at least one strictly Pareto compliant indicator as it is proposed in Falcón-Cardona et al. (2019).

The attainment function (da Fonseca, Fonseca, & Hall, 2001) is another tool for the performance assessment of multiobjective (stochastic) solvers. Assuming a multiobjective solver has produced k Pareto front approximations Y_N^j for $j = 1, 2, \ldots, k$ on a given problem, the empirical attainment function $\alpha : \mathbb{R}^m \to [0, 1]$ is defined as

$$\alpha(y) = \frac{1}{k}\sum_{j=1}^{k} \mathbb{1}\{Y_N^j \preceq \{y\}\}.$$

For a given $y \in \mathbb{R}^m$, the attainment function estimates the probability that the multiobjective solver reaches (in term of dominance in the objective space) the objective vector y. The interested reader can refer to Brockhoff, Auger, Hansen, and Tušar (2017), da Fonseca et al. (2001), Fonseca, da Fonseca, and Paquete (2005), Fonseca, Guerreiro, López-Ibáñez, and Paquete (2011) and Zitzler et al. (2008) for additional information.

4.2. Embedding performance indicators in multiobjective optimization algorithms

Performance indicators are able to quantify properties a good Pareto front approximation should possess. It is then logical to incorporate them into multiobjective optimization methods. By optimizing directly the indicator, one can hope to obtain approximations of the Pareto front satisfying demanding properties. For these last years, the evolutionary multiobjective community has frequently adopted this approach.

For example, performance indicators such as R_2 (Trautmann, Wagner, & Brockhoff, 2013) or *HV* (Beume, Naujoks, & Emmerich, 2007), I_ϵ (Zitzler & Künzli, 2004), *IGD*$^+$ (Falcón-Cardona et al., 2019; Lopez & Coello, 2016), *IGD–NS* (Tian et al., 2016) are used in selection mechanisms in evolutionary algorithms. The reader is invited to consult the recent survey (Falcón-Cardona & Coello, 2020) for more information on indicator-based multiobjective evolutionary algorithms. Similarly, the Γ-indicator (Custódio et al., 2011) enables to identify holes in the Pareto front approximation around which the algorithm can explore to improve diversity. In global stochastic optimization, some methods integrate hypervolume indicator (Bradford, Schweidtmann, & Lapkin, 2018; Emmerich, Yang, Deutz, Wang, & Fonseca, 2016; Feliot, Bect, & Vazquez, 2017) and its variants (Feliot, Bect, & Vazquez, 2019) or R_2 (Deutz, Emmerich, & Yang, 2019) to better explore the decision space. In Akhtar and Shoemaker (2016), the authors use Radial Basis models and the hypervolume indicator to identify next promising points to evaluate. In Al-Dujaili and Suresh (2018), the authors propose a multiobjective optimistic algorithm using the additive ϵ-indicator (3.2.6) and analyse its link with the weighted Tchebysheff approach.

The transformation of a multiobjective optimization problem into a single-objective quality indicator based problem implies a loss of information. Indeed, the choice of a specific performance indicator reflects the personal preferences of the decision user. It is then important to understand the bias of this choice on the solution set found. Given a performance indicator, the concept of optimal μ-distribution (Auger et al., 2009) refers to the study of the optimal distributions of non-dominated points of size μ which belong to the Pareto front and maximize (or minimize) the performance indicator for a given multiobjective problem. Their study enables to understand bias of considered indicators and analyze the behavior of bounded size indicator-based algorithms. The first ones were done for the hypervolume indicator and some of its variants in the biobjective case (Auger et al., 2009; 2012; Emmerich et al., 2013b) then extended to more objectives in Auger, Bader, and Brockhoff (2010) and Shukla et al. (2014). Theoretical results for the R_2 indicator (Brockhoff et al., 2012), the Δ_p indicator (Rudolph, Schütze, Grimme, Domínguez-Medina, & Trautmann, 2016) or the Hypervolume Sharpe-Ratio indicator (Guerreiro & Fonseca, 2020) for the biobjective case exist too.

4.3. 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 Wiecek, Ehrgott, and Engau (2016). The second class, designed as *a posteriori articulations of preferences* (Custódio et al., 2011) methods, aims at obtaining the whole Pareto front without combining any objective function in a single-objective framework. Evolutionary algorithms, Bayesian optimization methods (Emmerich et al., 2016) or deterministic algorithms such as DMS (Custódio et al., 2011) 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 (Audet et al., 2008). The use of performance indicators also is not relevant.

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 maximum number of 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 multiobjective Bayesian optimization (Emmerich et al., 2016), 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 Martí et al. (2016) and Wagner, Trautmann, and Martí (2011). The approach is to measure the progression of the current population combining performance indicators (hypervolume, *MDR*, etc.) and statistic tools (Kalman filter Martí et al., 2016, χ^2-variance test Wagner, Trautmann, and Naujoks (2009), 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.

4.4. Distribution and spread

The choice of spread and distribution indicators has only a sense when one wants to measure the distribution 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 distribution 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 indicators in the literature use minimum Euclidean or Manhattan distance between points such as the *SP* metric (3.3.1), the Δ index (3.3.2), *HRS* (3.3.4), and so on. The *DM* (3.3.12) and Γ-metric (3.3.3) indicators use a *"sorting distance"*; I_D (3.3.13) 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 (3.3.2), *U* measure (3.3.9), *DM* (3.3.12) for example), mean square (*SP* (3.3.1), D_{nc} (3.3.5)), 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* (3.3.12), *OS* (3.3.10), I_{OD} (3.3.11), Δ (3.3.3), Γ (3.3.3)).
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 (Zheng et al., 2017). Some indicators possess this property such as *DM* (3.3.12), Γ (3.3.3), Δ (3.3.3), E_s (3.3.16) or evenness indicators (3.3.7).
3. Distribution and spread performance indicators should not be based on external parameters, such as Zitzler's metric $M_2^\star$ (3.3.5), *UD* (3.3.5), or entropy measure (3.3.18).
4. An easy interpretation: a value returned by an indicator has to be "intuitive" to understand. For example, the binary uniformity (3.3.8) 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 (Das & Dennis, Jr., 1998) controls the spread of a Pareto front approximation. This method is also used in the context of blackbox optimization (Audet, Savard, & Zghal, 2010).

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 indicator 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 (Jiang et al., 2014; Okabe et al., 2003) have confirmed its efficiency compared to other performance indicators. That is why it has been deeply used in the evolutionary community (Riquelme et al., 2015). 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, Hypervolume Sharpe-Ratio indicator, modified inverted generated distance or degree of approximation whose computational 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.

Acknowledgments

This research was financed by Le Digabel's NSERC discovery grant RGPIN-2018-05286, and also by the NSERC CRD RDCPJ 490744-15 grant and an InnovÉÉ grant, both in collaboration with Hydro-Québec and Rio Tinto. We also would like to thank the anonymous reviewers for their suggestions and comments which helped to greatly improve this work.

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 indicators are drawn from Jiang et al. (2014). $Y_P \subseteq \mathcal{Y}_P$ corresponds to the *Pareto optimal solution set* and Y_N is a Pareto front 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.

Appendix B. Compatibility and completeness

Tables 4 and 5 summarize compatibility and completeness properties.

Only the strongest relationships are kept. Some of them are drawn from Zitzler et al. (2003). All spread and distribution indicators are not compatible with Pareto front approximation relations.

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 Zitzler and Thiele (1998)	3.1.5	C	None	Binary indicator	$\mathcal{O}(m\ \|Y_N^1\| \times \|Y_N^2\|)$	-
	Error ratio Van Veldhuizen (1999)	3.1.4	ER	None	Pareto front $\mathcal{Y}_P$	Low	✗
	Generational non dominated vector generation Van Veldhuizen and Lamont (2000)	3.1.3	$GNVG$	None	None	Low	✗
	Generational non dominated vector generation ratio Van Veldhuizen and Lamont (2000)	3.1.3	$GNVGR$	None	Pareto front Y_P	Low	✗
	Mutual domination rate Martí et al. (2016)	3.1.3	MDR	None	None	Low	✗
	Nondominated vector additional Van Veldhuizen and Lamont (2000)	3.1.3	NVA	None	None	Low	✗
	Overall nondominated vector generation Van Veldhuizen (1999)	3.1.1	$ONVG$	None	None	Low	✗
	Overall nondominated vector generation ratio Van Veldhuizen (1999)	3.1.2	$ONVGR$	None	Pareto front Y_P	Low	✗
	Ratio of non-dominated points by the reference set Hansen and Jaszkiewicz (1998)	3.1.5	C_{2R}	None	Reference set Y_R	$\mathcal{O}(m\ \|Y_N\| \times \|Y_R\|)$	✗
	Ratio of the reference points Hansen and Jaszkiewicz (1998)	3.1.4	C_{1R}	None	Reference set Y_R	$\mathcal{O}(m\ \|Y_N\| \times \|Y_R\|)$	✗
Convergence 3.2	Degree of Approximation Dilettoso et al. (2017)	3.2.7	DOA	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	Not strictly
	ϵ-family Zitzler et al. (2003)	3.2.6	I_ϵ	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	Not strictly
	Generational distance Van Veldhuizen (1999)	3.2.1	GD	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	✗
	γ-metric Deb et al. (2000)	3.2.1	γ	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	✗
	Maximum Pareto front error Van Veldhuizen (1999)	3.2.4	$MPFE$	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	✗
	$M_1^\star$-metric Zitzler et al. (2000)	3.2.1	$M_1^\star$	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	✗
	Progression metric Van Veldhuizen (1999)	3.2.5	-	None	None	$\mathcal{O}(m\ \|Y_N\|)$	✗
	Seven points average distance Schott (1995)	3.2.3	$SPAD$	None	Reference set Y_R	$\mathcal{O}(m\ \|Y_N\|)$	✗
	Standard deviation from the Generational distance Van Veldhuizen (1999)	3.2.2	$STDGD$	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	✗
Distribution and spread 3.3	Cluster Wu and Azarm (2000)	3.3.18	CL_μ	A parameter μ	None	High	✗
	Δ-index Deb et al. (2000)	3.3.2	Δ	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\|^2 + m\ \|Y_N\| \times \|Y_P\|)$	✗
	Δ'-index Deb et al. (2000)	3.3.2	Δ'	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	$\Delta^\star$ spread metric Zhou et al. (2006)	3.3.2	$\Delta^\star$	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\|^2 + m\ \|Y_N\| \times \|Y_P\|)$	✗
	Distribution metric Zheng et al. (2017)	3.3.12	DM	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Diversity comparison indicator Li et al. (2014)	3.3.18	DCI	A parameter div	k-ary indicator comparing $Y_N^1, Y_N^2, \ldots, Y_N^k$ non-dominated sets	$\mathcal{O}\big(m\ (k\ \|Y_N^{\max}\|)^2\big)$	✗
	Diversity indicator Cai et al. (2018)	3.3.15	DIR	None	Reference set Y_R	$\mathcal{O}(m\ \|Y_N\| \times \|Y_R\|)$	✗
	Entropy metric Farhang-Mehr and Azarm (2004)	3.3.18	-	A parameter $grids$	None	High	✗
	Evenness Messac and Mattson (2004)	3.3.7	ξ	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Extension Meng et al. (2005)	3.3.14	EX	None	Pareto front Y_P	$\mathcal{O}(m\ \|Y_N\| \times \|Y_P\|)$	✗
	Γ-metric Custódio et al. (2011)	3.3.3	Γ	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Hole Relative Size Collette and Siarry (2005)	3.3.4	HRS	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Laumanns metric Laumanns et al. (2000)	3.3.17	-	None	None	$\mathcal{O}(\|Y_N\|^{\frac{m}{3}} \operatorname{poly}\log \|Y_N\|)$	✗
	Modified Diversity indicator Asafuddoula et al. (2015)	3.3.18	$M\text{–}DI$	A parameter δ	Reference set Y_R	$\mathcal{O}(m\ \|Y_N\|^2 \times \|Y_R\|)$	
	$M_2^\star$-metric Zitzler et al. (2000)	3.3.5	$M_2^\star$	Niche radius σ	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	$M_3^\star$-metric Zitzler et al. (2000)	3.3.5	$M_3^\star$	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Number of distinct choices Wu and Azarm (2000)	3.3.18	NDC_μ	A parameter μ	None	High	✗
	Outer diameter Zitzler et al. (2008)	3.3.11	I_{OD}	None	None	$\mathcal{O}(m\ \|Y_N\|)$	✗
	Overall Pareto Spread Wu and Azarm (2000)	3.3.10	OS	None	$\tilde{y}^I$ and $\tilde{y}^M$	$\mathcal{O}(m\ \|Y_N\|)$	✗
	Riesz S-energy Hardin and Saff (2004)	3.3.16	E_S	A parameter s	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Sigma diversity metric Mostaghim and Teich (2005)	3.3.18	σ	A parameter $lines$	None	High	✗
	Spacing Schott (1995)	3.3.1	SP	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	U-measure Leung and Wang (2003)	3.3.9	U	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Uniform assessment metric Li et al. (2008)	3.3.13	I_D	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Uniform distribution Tan et al. (2002)	3.3.5	UD	Niche radius σ	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Uniformity Sayın (2000)	3.3.6	δ	None	None	$\mathcal{O}(m\ \|Y_N\|^2)$	✗
	Uniformity Meng et al. (2005)	3.3.8	-	None	Binary	Quadratic	✗

(continued on next page)

Table 3 (*continued*)

Category	Performance indicators	Sect.	Symbol	Parameters	Comparison sets	Computational complexity	Monotone
Convergence and distribution 3.4	Averaged Hausdorff distance Schutze et al. (2012)	3.4.2	Δ_q	None	Pareto front Y_P	$\mathcal{O}(m\ \lvert Y_N\rvert \times \lvert Y_P\rvert)$	✗
	Cone-based hypervolume Emmerich et al. (2013b)	3.4.7	-	Angle γ and Reference point r	None	$\mathcal{O}(\lvert Y_N\rvert^{\frac{m}{3}} \text{poly} \log \lvert Y_N\rvert)$	Strictly
	Dominance move Li and Yao (2017)	3.4.6	*DoM*	None	Binary indicator	$\mathcal{O}(\lvert Y_N\rvert \log \lvert Y_N\rvert)$	-
	D-metric/Difference coverage of two sets Zitzler (1999)	3.4.7	-	Reference point r	Binary indicator	$\mathcal{O}(\lvert Y_N\rvert^{\frac{m}{3}} \text{poly} \log \lvert Y_N\rvert)$	-
	D_R-metric Czyzzak and Jaszkiewicz (1998)	3.4.3	-	None	Reference set Y_R	$\mathcal{O}(m\ \lvert Y_N\rvert \times \lvert Y_R\rvert)$	Not strictly
	Hyperarea difference Wu and Azarm (2000)	3.4.7	*HD*	Reference point r	None	$\mathcal{O}(\lvert Y_N\rvert^{\frac{m}{3}} \text{poly} \log \lvert Y_N\rvert)$	Strictly
	Hypervolume indicator (or S-metric) Zitzler et al. (2000)	3.4.7	*HV*	Reference point r	None	$\mathcal{O}(\lvert Y_N\rvert^{\frac{m}{3}} \text{poly} \log \lvert Y_N\rvert)$	Strictly
	Hypervolume Sharpe-ratio indicator Yevseyeva et al. (2014)	3.4.8	I_{HSR}	Reference points y^l and y^u	None	Polynomial	Not strictly
	Inverted generational distance Coello and Cortés (2005)	3.4.1	*IGD*	None	Pareto front Y_P	$\mathcal{O}(m\ \lvert Y_N\rvert \times \lvert Y_P\rvert)$	✗
	Inverted generation distance with non contributed solutions detection Tian et al. (2016)	3.4.1	*IGD–NS*	None	Pareto front Y_P	$\mathcal{O}(m\ \lvert Y_N\rvert \times \lvert Y_P\rvert)$	✗
	G-metric Lizarraga-Lizarraga et al. (2008)	3.4.5	-	None	k-ary indicator comparing $Y_N^1, Y_N^2, \ldots, Y_N^k$ non-dominated sets	$\mathcal{O}(k^3\ \lvert Y_N^{\max}\rvert^2)$	Not strictly
	Logarithmic hypervolume indicator Friedrich et al. (2011)	3.4.7	$\log HV$	Reference point r	None	$\mathcal{O}(\lvert Y_N\rvert^{\frac{m}{3}} \text{poly} \log \lvert Y_N\rvert)$	Strictly
	Modified inverted generational distance Ishibuchi et al. (2015)	3.4.3	IGD^+	None	Pareto front Y_P	$\mathcal{O}(m\ \lvert Y_N\rvert \times \lvert Y_P\rvert)$	Not strictly
	Performance comparison indicator Li et al. (2015)	3.4.6	*PCI*	σ distance	k-ary indicator comparing $Y_N^1, Y_N^2, \ldots, Y_N^k$ non-dominated sets	Quadratic	Not strictly
	p, q-averaged distance Vargas and Bogoya (2018)	3.4.2	$\Delta_{p,q}$	None	Pareto front Y_P	Quadratic	✗
	R-metric Hansen and Jaszkiewicz (1998)	3.4.4	Y_R	A set W of weights vectors	Reference set Y_R	$\mathcal{O}(m\ \lvert Y_N\rvert \times \lvert Y_R\rvert \times \lvert W\rvert)$	Not strictly

Table 4
Compatibility and completeness of unary performance indicators.

Category	Performance indicators	Sect.	Symbol	Boolean function	Compatible	Complete
Cardinality 3.1	Error ratio Van Veldhuizen (1999)	3.1.4	*ER*	$ER(Y_N^1) < ER(Y_N^2)$	✗	✗
	Generational non dominated vector generation Van Veldhuizen and Lamont (2000)	3.1.3	*GNVG*	-	-	-
	Generational non dominated vector generation ratio Van Veldhuizen and Lamont (2000)	3.1.3	*GNVGR*	-	-	-
	Mutual domination rate Martí et al. (2016)	3.1.6	*MDR*	-	-	-
	Nondominated vector additional Van Veldhuizen and Lamont (2000)	3.1.3	*NVA*	-	-	-
	Overall nondominated vector generation Van Veldhuizen (1999)	3.1.1	*ONVG*	$ONVG(Y_N^1) > ONVG(Y_N^2)$	✗	✗
	Overall nondominated vector generation ratio Van Veldhuizen (1999)	3.1.2	*ONVGR*	$ONVGR(Y_N^1; Y_P) > ONVGR(Y_N^2; Y_P)$	✗	✗
	Ratio of non-dominated points by the reference set Hansen and Jaszkiewicz (1998)	3.1.5	C_{2R}	$C_{2R}(Y_N^1; Y_R) > C_{2R}(Y_N^2; Y_R)$	✗	✗
	Ratio of the reference points Hansen and Jaszkiewicz (1998)	3.1.4	C_{1R}	$C_{1R}(Y_N^1; Y_R) > C_{1R}(Y_N^2; Y_R)$	✗	✗
Convergence 3.2	Degree of Approximation Dilettoso et al. (2017)	3.2.7	*DOA*	$DOA(Y_N^1; Y_P) < DOA(Y_N^2; Y_P)$	Not better than	$\prec$
	Generational distance Van Veldhuizen (1999)	3.2.1	*GD*	$GD(Y_N^1; Y_P) < GD(Y_N^2; Y_P)$	✗	✗
	γ-metric Deb et al. (2000)	3.2.1	γ	$\gamma(Y_N^1; Y_P) < \gamma(Y_N^2; Y_P)$	✗	✗
	Maximum Pareto front error Van Veldhuizen (1999)	3.2.4	*MPFE*	$MPFE(Y_N^1; Y_P) < MPFE(Y_N^2; Y_P)$	✗	✗
	$M_1^\star$-metric Zitzler et al. (2000)	3.2.1	$M_1^\star$	$M_1^\star(Y_N^1; Y_P) < M_1^\star(Y_N^2; Y_P)$	✗	✗
	Progression metric Van Veldhuizen (1999)	3.2.5	-	-	-	-
	Seven points average distance Schott (1995)	3.2.3	*SPAD*	$SPAD(Y_N^1; Y_P) < SPAD(Y_N^2; Y_P)$	✗	✗
	Standard deviation from the Generational distance Van Veldhuizen (1999)	3.2.2	*STDGD*	-	-	-
Distribution and spread	Cluster Wu and Azarm (2000)	3.3.18	CL_μ	-	-	-
	Δ-index Deb et al. (2000)	3.3.2	Δ	$\Delta(Y_N^1; Y_P) < \Delta(Y_N^2; Y_P)$	✗	✗

(*continued on next page*)

Table 4 (*continued*)

Category	Performance indicators	Sect.	Symbol	Boolean function	Compatible	Complete
3.3	Δ'-index Deb et al. (2000)	3.3.2	Δ'	$\Delta'(Y_N^1) < \Delta'(Y_N^2)$	✗	✗
	$\Delta^{\star}$ spread metric Zhou et al. (2006)	3.3.2	$\Delta^{\star}$	$\Delta^{\star}(Y_N^1; Y_P) < \Delta^{\star}(Y_N^2; Y_P)$	✗	✗
	Distribution metric Zheng et al. (2017)	3.3.12	DM	$DM(Y_N^1) < DM(Y_N^2)$	✗	✗
	Diversity indicator Cai et al. (2018)	3.3.15	DIR	$DIR(Y_N^1) < DIR(Y_N^2)$	✗	✗
	Entropy metric Farhang-Mehr and Azarm (2004)	3.3.18	-	-	-	-
	Evenness Messac and Mattson (2004)	3.3.7	ξ	$\xi(Y_N^1) < \xi(Y_N^2)$	✗	✗
	Extension Meng et al. (2005)	3.3.14	EX	$EX(Y_N^1; Y_P) < EX(Y_N^2; Y_P)$	✗	✗
	Γ-metric Custódio et al. (2011)	3.3.3	Γ	$\Gamma(Y_N^1) < \Gamma(Y_N^2)$	✗	✗
	Hole Relative Size Collette and Siarry (2005)	3.3.4	HRS	$HRS(Y_N^1) < HRS(Y_N^2)$	✗	✗
	Laumanns metric Laumanns et al. (2000)	3.3.17	-	$I_L(Y_N^1) > I_L(Y_N^2)$	✗	✗
	Modified Diversity indicator Asafuddoula et al. (2015)	3.3.18	$M\text{–}DI$	$M\text{–}DI(Y_N^1; Y_R) > M\text{–}DI(Y_N^2; Y_R)$	✗	✗
	$M_2^{\star}$-metric Zitzler et al. (2000)	3.3.5	$M_2^{\star}$	$M_2^{\star}(Y_N^1; \sigma) > M_2^{\star}(Y_N^2; \sigma)$	✗	✗
	$M_3^{\star}$-metric Zitzler et al. (2000)	3.3.5	$M_3^{\star}$	$M_3^{\star}(Y_N^1) > M_3^{\star}(Y_N^2)$	✗	✗
	Number of distinct choices Wu and Azarm (2000)	3.3.18	NDC_{μ}	$NDC_{\mu}(Y_N^1) > NDC_{\mu}(Y_N^2)$	✗	✗
	Outer diameter Zitzler et al. (2008)	3.3.11	I_{OD}	$I_{OD}(Y_N^1) > I_{OD}(Y_N^2)$	✗	✗
	Overall Pareto Spread Wu and Azarm (2000)	3.3.10	OS	$OS(Y_N^1) > OS(Y_N^2)$	✗	✗
	Riesz S-energy Hardin and Saff (2004)	3.3.16	E_S	$E_S(Y_N^1) < E_S(Y_N^2)$	✗	✗
	Sigma diversity metric Mostaghim and Teich (2005)	3.3.18	σ	$\sigma(Y_N^1; d) > \sigma(Y_N^2; d)$	✗	✗
	Spacing Schott (1995)	3.3.1	SP	$SP(Y_N^1) < SP(Y_N^2)$	✗	✗
	U-measure Leung and Wang (2003)	3.3.9	U	$U(Y_N^1) < U(Y_N^2)$	✗	✗
	Uniform assessment metric Li et al. (2008)	3.3.13	I_D	$I_D(Y_N^1) > I_D(Y_N^2)$	✗	✗
	Uniform distribution Tan et al. (2002)	3.3.5	UD	$UD(Y_N^1; \sigma) < UD(Y_N^2; \sigma)$	✗	✗
	Uniformity Sayın (2000)	3.3.6	δ	$\delta(Y_N^1) < \delta(Y_N^2)$	✗	✗
Convergence and distribution 3.4	Averaged Hausdorff distance Schutze et al. (2012)	3.4.2	Δ_q	$\Delta_q(Y_N^1; Y_P) < \Delta_q(Y_N^2; Y_P)$	✗	✗
	Cone-based hypervolume Emmerich et al. (2013b)	3.4.7	-	$\chi(Y_N^1) > \chi(Y_N^2)$	Not better than	$\lhd$
	D_R-metric Czyzzak and Jaszkiewicz (1998)	3.4.3	-	$D_R(Y_N^1; Y_R) < D_R(Y_N^2; Y_R)$	Not better than	$\prec\prec$
	Hyperarea difference Wu and Azarm (2000)	3.4.7	HD	$HD(Y_N^1) < HD(Y_N^2)$	Not better than	$\lhd$
	Hypervolume indicator (or S-metric) Zitzler et al. (2000)	3.4.7	HV	$HV(Y_N^1; r) > HV(Y_N^2; r)$	Not better than	$\lhd$
	Hypervolume Sharpe-ratio indicator Yevseyeva et al. (2014)	3.4.8	I_{HSR}	$I_{HSR}(Y_N^1; y^l, y^u) > I_{HSR}(Y_N^2; y^l, y^u)$	Not better than	$\prec$
	Inverted generational distance Coello and Cortés (2005)	3.4.1	IGD	$IGD(Y_N^1, P) < IGD(B, P)$	✗	✗
	Inverted generation distance with non contributed solutions detection Tian et al. (2016)	3.4.1	$IGD\text{–}NS$	$IGD\text{–}NS(Y_N^1; Y_P) < IGD\text{–}NS(Y_N^2; Y_P)$	✗	✗
	Logarithmic hypervolume indicator Friedrich et al. (2011)	3.4.7	$\log HV$	$\log HV(Y_N^1; r) > \log HV(Y_N^2; r)$	Not better than	$\lhd$
	Modified inverted generational distance Ishibuchi et al. (2015)	3.4.3	IGD^{+}	$IGD^{+}(Y_N^1; Y_P) < IGD^{+}(Y_N^2; Y_P)$	Not better than	$\preceq$
	p, q-averaged distance Vargas and Bogoya (2018)	3.4.2	$\Delta_{p,q}$	$\Delta_{p,q}(Y_N^1; Y_P) < \Delta_{p,q}(Y_N^1; Y_P)$	✗	✗

Table 5
Compatibility and completeness of binary performance indicators (inspired by Zitzler et al. (2003)): 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 $\lhd$	$\preceq$	$=$	$\parallel$
Cardinality 3.1	C-metric/Two sets Coverage Zitzler and Thiele (1998)	3.1.5	C	$C(Y_N^1, Y_N^2) = 1$ $C(Y_N^2, Y_N^1) < 1$	$C(Y_N^1, Y_N^2) = 1$	$C(Y_N^1, Y_N^2) = 1$ $C(Y_N^2, Y_N^1) = 1$	$C(Y_N^1, Y_N^2) > 1$ $C(Y_N^2, Y_N^1) > 1$
Convergence 3.2	Additive ϵ-indicator Zitzler et al. (2003)	3.2.6	I_{ϵ}	$I_{\epsilon}(Y_N^1, Y_N^2) \le 0$ $I_{\epsilon}(Y_N^2, Y_N^1) > 0$	$I_{\epsilon}(Y_N^1, Y_N^2) \le 0$	$I_{\epsilon}(Y_N^1, Y_N^2) = 0$ $I_{\epsilon}(Y_N^2, Y_N^1) = 0$	$I_{\epsilon}(Y_N^1, Y_N^2) > 0$ $I_{\epsilon}(Y_N^2, Y_N^1) > 0$
Distribution and spread 3.3	Diversity comparison indicator Li et al. (2014)	3.3.18	DCI	✗	✗	✗	✗
	Uniformity Meng et al. (2005)	3.3.8	–	✗	✗	✗	✗
Convergence and distribution 3.4	Dominance move Li and Yao (2017)	3.4.6	DoM	$DoM(Y_N^1, Y_N^2) = 0$ $DoM(Y_N^2, Y_N^1) > 0$	$DoM(Y_N^1, Y_N^2) = 0$ $DoM(Y_N^2, Y_N^1) \ge 0$	$DoM(Y_N^1, Y_N^2) = 0$ $DoM(Y_N^2, Y_N^1) = 0$	$DoM(Y_N^1, Y_N^2) > 0$ $DoM(Y_N^1, Y_N^2) > 0$
	D-metric/Difference coverage of two sets Zitzler (1999)	3.4.7	–	$D(Y_N^1, Y_N^2) > 0$ $D(Y_N^2, Y_N^1) = 0$	$D(Y_N^1, Y_N^2) \ge 0$ $D(Y_N^2, Y_N^1) = 0$	$D(Y_N^1, Y_N^2) = 0$ $D(Y_N^2, Y_N^1) = 0$	$D(Y_N^1, Y_N^2) > 0$ $D(Y_N^2, Y_N^1) > 0$
	G-metric Lizarraga-Lizarraga et al. (2008)	3.4.5	–	–	–	-	-
	Performance comparison indicator Li et al. (2015)	3.4.6	PCI	–	–	-	-
	R-metric Hansen and Jaszkiewicz (1998)	3.4.4	R	–	–	-	-

References

Akhtar, T., & Shoemaker, C. (2016). Multi objective optimization of computationally expensive multi-modal functions with RBF surrogates and multi-rule selection. *Journal of Global Optimization, 64*(1), 17–32. https://doi.org/10.1007/s10898-015-0270-y.

Al-Dujaili, A., & Suresh, S. (2016). BMOBench: Black-box multi-objective optimization benchmarking platform. *Technical Report*. arXiv.

Al-Dujaili, A., & Suresh, S. (2018). Revisiting norm optimization for multi-objective black-box problems: A finite-time analysis. *Journal of Global Optimization, 73*(3), 659–673. https://doi.org/10.1007/s10898-018-0709-z.

Asafuddoula, M., Ray, T., & Singh, H. (2015). Characterizing Pareto front approximations in many-objective optimization. In *Proceedings of the 2015 annual conference on genetic and evolutionary computation* (pp. 607–614). New York, NY, USA: ACM. https://doi.org/10.1145/2739480.2754701.

Audet, C., Savard, G., & Zghal, W. (2008). Multiobjective optimization through a series of single-objective formulations. *SIAM Journal on Optimization, 19*(1), 188–210. https://doi.org/10.1137/060677513.

Audet, C., Savard, G., & Zghal, W. (2010). A mesh adaptive direct search algorithm for multiobjective optimization. *European Journal of Operational Research, 204*(3), 545–556. https://doi.org/10.1016/j.ejor.2009.11.010.

Auger, A., Bader, J., & Brockhoff, D. (2010). Theoretically investigating optimal μ-distributions for the hypervolume indicator: First results for three objectives. In R. Schaefer, C. Cotta, J. Kołodziej, & G. Rudolph (Eds.), *Parallel problem solving from nature, PPSN XI* (pp. 586–596). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-15844-5_59.

Auger, A., Bader, J., Brockhoff, D., & Zitzler, E. (2009). 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* (pp. 87–102). New York, NY, USA: ACM. https://doi.org/10.1145/1527125.1527138.

Auger, A., Bader, J., Brockhoff, D., & Zitzler, E. (2012). Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications. *Theoretical Computer Science, 425*, 75–103. https://doi.org/10.1016/j.tcs.2011.03.012.

Bäck, T. (1996). *Evolutionary algorithms in theory and practice: Evolution strategies, evolutionary programming, genetic algorithms*. New York, NY, USA: Oxford University Press.

Beiranvand, V., Hare, W., & Lucet, Y. (2017). Best practices for comparing optimization algorithms. *Optimization and Engineering, 18*(4), 815–848. https://doi.org/10.1007/s11081-017-9366-1.

Beume, N., Fonseca, C., Lopez-Ibanez, M., Paquete, L., & Vahrenhold, J. (2009). On the complexity of computing the hypervolume indicator. *IEEE Transactions on Evolutionary Computation, 13*(5), 1075–1082. https://doi.org/10.1109/tevc.2009.2015575.

Beume, N., Naujoks, B., & Emmerich, M. (2007). SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research, 181*(3), 1653–1669. https://doi.org/10.1016/j.ejor.2006.08.008.

Bradford, E., Schweidtmann, A., & Lapkin, A. (2018). Efficient multiobjective optimization employing Gaussian processes, spectral sampling and a genetic algorithm. *Journal of Global Optimization, 71*(2), 407–438. https://doi.org/10.1007/s10898-018-0609-2.

Branke, J., Deb, K., Miettinen, K., & Słowiński, R. (2008). *Multiobjective optimization: Interactive and evolutionary approaches*: 5252. Springer. https://doi.org/10.1007/978-3-540-88908-3.

Bringmann, K., & Friedrich, T. (2010a). Approximating the volume of unions and intersections of high-dimensional geometric objects. *Computational Geometry, 43*(6), 601–610. https://doi.org/10.1016/j.comgeo.2010.03.004.

Bringmann, K., & Friedrich, T. (2010b). Tight bounds for the approximation ratio of the hypervolume indicator. In R. Schaefer, C. Cotta, J. Kołodziej, & G. Rudolph (Eds.), *Parallel problem solving from nature, PPSN XI* (pp. 607–616). Berlin, Heidelberg: Springer.

Bringmann, K., & Friedrich, T. (2013). Approximation quality of the hypervolume indicator. *Artificial Intelligence, 195*, 265–290. https://doi.org/10.1016/j.artint.2012.09.005.

Brockhoff, D., Auger, A., Hansen, N., & Tušar, T. (2017). Quantitative performance assessment of multiobjective optimizers: The average runtime attainment function. In H. Trautmann, G. Rudolph, K. Klamroth, O. Schütze, M. Wiecek, Y. Jin, & C. Grimme (Eds.), *Evolutionary multi-criterion optimization* (pp. 103–119). Cham: Springer. https://doi.org/10.1007/978-3-319-54157-0_8.

Brockhoff, D., Wagner, T., & Trautmann, H. (2012). On the properties of the R2 indicator. In *Proceedings of the 14th annual conference on genetic and evolutionary computation* (pp. 465–472). New York, NY, USA: ACM. https://doi.org/10.1145/2330163.2330230.

Brockhoff, D., Wagner, T., & Trautmann, H. (2015). 2 indicator-based multiobjective search. *Evolutionary Computation, 23*(3), 369–395. https://doi.org/10.1162/EVCO_a_00135.

Cai, X., Sun, H., & Fan, Z. (2018). A diversity indicator based on reference vectors for many-objective optimization. *Information Sciences, 430–431*, 467–486. https://doi.org/10.1016/j.ins.2017.11.051.

Chan, T. (2013). Klee's measure problem made easy. In *IEEE 54th annual symposium on foundations of computer science* (pp. 410–419). IEEE. https://doi.org/10.1109/FOCS.2013.51.

Cheng, S., Shi, Y., & Qin, Q. (2012). On the performance metrics of multiobjective optimization. In Y. Tan, Y. Shi, & Z. Ji (Eds.), *Advances in swarm intelligence* (pp. 504–512). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-30976-2_61.

Coello, C., & Cortés, N. (2005). Solving multiobjective optimization problems using an artificial immune system. *Genetic Programming and Evolvable Machines, 6*(2), 163–190. https://doi.org/10.1007/s10710-005-6164-x.

Collette, Y., & Siarry, P. (2005). Three new metrics to measure the convergence of metaheuristics towards the Pareto frontier and the aesthetic of a set of solutions in biobjective optimization. *Computers and Operations Research, 32*(4), 773–792. https://doi.org/10.1016/j.cor.2003.08.017.

Collette, Y., & Siarry, P. (2011). *Multiobjective optimization: Principles and case studies*. Springer. https://doi.org/10.1007/978-3-662-08883-8.

Cornuéjols, G., Peña, J., & Tütüncü, R. (2018). *Optimization methods in finance*. Cambridge University Press. https://doi.org/10.1017/9781107297340.

Custódio, A., Emmerich, M., & Madeira, J. (2012). Recent developments in derivative-free multiobjective optimization. *Computational Technology Reviews, 5*(1), 1–30. https://doi.org/10.4203/ctr.5.1.

Custódio, A., Madeira, J., Vaz, A., & Vicente, L. (2011). Direct multisearch for multiobjective optimization. *SIAM Journal on Optimization, 21*(3), 1109–1140. https://doi.org/10.1137/10079731X.

Czyzzak, P., & Jaszkiewicz, A. (1998). Pareto simulated annealing–a metaheuristic technique for multiple-objective combinatorial optimization. *Journal of Multi-Criteria Decision Analysis, 7*(1), 34–47. https://doi.org/10.1002/(sici)1099-1360(199801)7:1<34::aid-mcda161⟩3.0.co;2-6.

Dächert, K., Klamroth, K., Lacour, R., & Vanderpooten, D. (2017). Efficient computation of the search region in multi-objective optimization. *European Journal of Operational Research, 260*(3), 841–855. https://doi.org/10.1016/j.ejor.2016.05.029.

Das, I., & DennisJr. , J. (1998). Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization, 8*(3), 631–657. https://doi.org/10.1137/S1052623496307510.

Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*. New York, NY, USA: John Wiley and Sons.

Deb, K., Agrawal, S., Pratap, A., & Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J. Merelo, & H.-P. H. P. Schwefel (Eds.), *Parallel problem solving from nature PPSN VI* (pp. 849–858). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-45356-3_83.

Deb, K., & Tiwari, S. (2008). Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization. *European Journal of Operational Research, 185*(3), 1062–1087. https://doi.org/10.1016/j.ejor.2006.06.042.

Deutz, A., Emmerich, M., & Yang, K. (2019). The expected R2-indicator improvement for multi-objective Bayesian optimization. In K. Deb, E. Goodman, C. Coello, K. Klamroth, K. Miettinen, S. Mostaghim, & P. Reed (Eds.), *Evolutionary multi-criterion optimization* (pp. 359–370). Cham: Springer. https://doi.org/10.1007/978-3-030-12598-1_29.

Dilettoso, E., Rizzo, S., & Salerno, N. (2017). A weakly Pareto compliant quality indicator. *Mathematical and Computational Applications, 22*(1). https://doi.org/10.3390/mca22010025.

Dolan, E., & Moré, J. (2002). Benchmarking optimization software with performance profiles. *Mathematical Programming, 91*(2), 201–213. https://doi.org/10.1007/s101070100263.

Ehrgott, M. (2005). *Multicriteria optimization, volume 491 of lecture notes in economics and mathematical systems* (2nd). Berlin: Springer.

Emmerich, M., Deutz, A., & Kruisselbrink, J. (2013a). *On Quality Indicators for Black-Box Level Set Approximation* (pp. 157–185)). Berlin, Heidelberg: Springer.

Emmerich, M., Deutz, A., Kruisselbrink, J., & Shukla, P. (2013b). Cone-based hypervolume indicators: Construction, properties, and efficient computation. In R. Purshouse, P. Fleming, C. Fonseca, S. Greco, & J. Shaw (Eds.), *Evolutionary multi-criterion optimization* (pp. 111–127). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-37140-0_12.

Emmerich, M., Deutz, A., & Yevseyeva, I. (2014). On Reference Point Free Weighted Hypervolume Indicators based on Desirability Functions and their Probabilistic Interpretation. *Procedia Technology, 16*, 532–541. https://doi.org/10.1016/j.protcy.2014.10.001.

Emmerich, M., Yang, K., Deutz, A., Wang, H., & Fonseca, C. (2016). *A Multicriteria Generalization of Bayesian Global Optimization* (pp. 229–242)). Cham: Springer.

Falcón-Cardona, J., & Coello, C. (2020). Indicator-based multi-objective evolutionary algorithms: A comprehensive survey. *ACM Computing Surveys, 53*(2), 1–35. https://doi.org/10.1145/3376916.

Falcón-Cardona, J., Coello, C., & Emmerich, M. (2019). CRI-EMOA: A Pareto-front shape invariant evolutionary multi-objective algorithm. In K. Deb, E. Goodman, C. Coello, K. Klamroth, K. Miettinen, S. Mostaghim, & P. Reed (Eds.), *Evolutionary multi-criterion optimization* (pp. 307–318). Cham: Springer.

Falcón-Cardona, J., Emmerich, M., & Coello, C. (2019). On the Construction of Pareto-Compliant Quality Indicators. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (pp. 2024–2027). ACM. https://doi.org/10.1145/3319619.3326902.

Farhang-Mehr, A., & Azarm, S. (2004). An information-theoretic entropy metric for assessing multi-objective optimization solution set quality. *Journal of Mechanical Design, 125*(4), 655–663. https://doi.org/10.1115/1.1623186.

Faulkenberg, S., & Wiecek, M. (2009). On the quality of discrete representations in multiple objective programming. *Optimization and Engineering, 11*(3), 423–440. https://doi.org/10.1007/s11081-009-9099-x.

Feliot, P., Bect, J., & Vazquez, E. (2017). A Bayesian approach to constrained single- and multi-objective optimization. *Journal of Global Optimization, 67*(1–2), 97–133. https://doi.org/10.1007/s10898-016-0427-3.

Feliot, P., Bect, J., & Vazquez, E. (2019). User preferences in Bayesian multi-objective optimization: The expected weighted hypervolume improvement criterion. In

G. Nicosia, P. Pardalos, G. Giuffrida, R. Umeton, & V. Sciacca (Eds.), *Machine learning, optimization, and data science* (pp. 533–544). Cham: Springer. https://doi.org/10.1007/978-3-030-13709-0_45.

Fonseca, C., da Fonseca, V., & Paquete, L. (2005). Exploring the performance of stochastic multiobjective optimisers with the second-order attainment function. In C. Coello, A. H. Aguirre, & E. Zitzler (Eds.), *Evolutionary multi-criterion optimization* (pp. 250–264). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-31880-4_18.

Fonseca, C., Guerreiro, A., López-Ibáñez, M., & Paquete, L. (2011). On the computation of the empirical attainment function. In R. Takahashi, K. Deb, E. Wanner, & S. Greco (Eds.), *Evolutionary multi-criterion optimization* (pp. 106–120). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-19893-9_8.

da Fonseca, V., Fonseca, C., & Hall, A. (2001). Inferential performance assessment of stochastic optimisers and the attainment function. In E. Zitzler, L. Thiele, K. C. A. C. Deb, & D. Corne (Eds.), *Evolutionary multi-criterion optimization* (pp. 213–225). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-44719-9_15.

Friedrich, T., Bringmann, K., Voß, T., & Igel, C. (2011). The logarithmic hypervolume indicator. In *Proceedings of the 11th workshop proceedings on foundations of genetic algorithms* (pp. 81–92). New York, NY, USA: ACM. https://doi.org/10.1145/1967654.1967662.

Ghosh, D., & Chakraborty, D. (2015). A direction based classical method to obtain complete Pareto set of multi-criteria optimization problems. *Opsearch, 52*(2), 340–366. https://doi.org/10.1007/s12597-014-0178-1.

Guerreiro, A., & Fonseca, C. (2016). Hypervolume sharpe-ratio indicator: Formalization and first theoretical results. In *Parallel problem solving from nature – PPSN XIV* (pp. 814–823). https://doi.org/10.1007/978-3-319-45823-6_76.

Guerreiro, A., & Fonseca, C. (2020). An analysis of the hypervolume sharpe-ratio indicator. *European Journal of Operational Research, 283*(2), 614–629. https://doi.org/10.1016/j.ejor.2019.11.023.

Hansen, M., & Jaszkiewicz, A. (1998). *Evaluating the quality of approximations to the non-dominated set*. IMM, Department of Mathematical Modelling.

Hardin, D., & Saff, E. (2004). Discretizing manifolds via minimum energy points. *Notices of the AMS, 51*(10), 1186–1194.

Hardin, D., & Saff, E. (2005). Minimal Riesz energy point configurations for rectifiable d-dimensional manifolds. *Advances in Mathematics, 193*(1), 174–204. https://doi.org/10.1016/j.aim.2004.05.006.

Ishibuchi, H., Imada, R., Setoguchi, Y., & Nojima, Y. (2018). How to specify a reference point in hypervolume calculation for fair performance comparison. *Evolutionary Computation, 26*(3), 411–440. https://doi.org/10.1162/evco_a_00226.

Ishibuchi, H., Masuda, H., Tanigaki, Y., & Nojima, Y. (2014). Difficulties in specifying reference points to calculate the inverted generational distance for many-objective optimization problems. In *IEEE symposium on computational intelligence in multi-criteria decision-making* (pp. 170–177). https://doi.org/10.1109/mcdm.2014.7007204.

Ishibuchi, H., Masuda, H., Tanigaki, Y., & Nojima, Y. (2015). Modified distance calculation in generational distance and inverted generational distance. In A. Gaspar-Cunha, C. H. Antunes, & C. Coello (Eds.), *Evolutionary multi-criterion optimization* (pp. 110–125). Cham: Springer. https://doi.org/10.1007/978-3-319-15892-1_8.

Jaszkiewicz, A. (2018). Improved quick hypervolume algorithm. *Computers and Operations Research, 90*, 72–83. https://doi.org/10.1016/j.cor.2017.09.016.

Jiang, S., Ong, Y., Zhang, J., & Feng, L. (2014). Consistencies and contradictions of performance metrics in multiobjective optimization. *IEEE Transactions on Cybernetics, 44*(12), 2391–2404. https://doi.org/10.1109/tcyb.2014.2307319.

Jiang, S., Yang, S., & Li, M. (2016). On the use of hypervolume for diversity measurement of Pareto front approximations. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1–8). IEEE. https://doi.org/10.1109/SSCI.2016.7850225.

Kidd, M., Lusby, R., & Larsen, J. (2020). Equidistant representations: Connecting coverage and uniformity in discrete biobjective optimization. *Computers and Operations Research, 117*, 104872. https://doi.org/10.1016/j.cor.2019.104872.

Knowles, J., & Corne, D. (2002). On metrics for comparing nondominated sets. In *Evolutionary computation: 1* (pp. 711–716). Honolulu, HI: IEEE. https://doi.org/10.1109/CEC.2002.1007013.

Lacour, R., Klamroth, K., & Fonseca, C. (2017). A box decomposition algorithm to compute the hyperVolume indicator. *Computers and Operations Research, 79*, 347–360. https://doi.org/10.1016/j.cor.2016.06.021.

Laumanns, M., Günter, R., & Schwefel, H. (1999). *Approximating the Pareto set: Concepts, diversity issues, and performance assessment*. Secretary of the SFB 531.

Laumanns, M., Zitzler, E., & Thiele, L. (2000). A unified model for multi-objective evolutionary algorithms with elitism. In *Congress on evolutionary computation: 1* (pp. 46–53). https://doi.org/10.1109/CEC.2000.870274. Piscataway, New Jersey, USA

Leung, Y., & Wang, Y. (2003). U-measure: a quality measure for multiobjective programming. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 33*(3), 337–343. https://doi.org/10.1109/TSMCA.2003.817059.

Li, M., Yang, S., & Liu, X. (2014). Diversity comparison of Pareto front approximations in many-objective optimization. *IEEE Transactions on Cybernetics, 44*(12), 2568–2584. https://doi.org/10.1109/TCYB.2014.2310651.

Li, M., Yang, S., & Liu, X. (2015). A performance comparison indicator for Pareto front approximations in many-objective optimization. In *Proceedings of the 2015 annual conference on genetic and evolutionary computation* (pp. 703–710). New York, NY, USA: ACM. https://doi.org/10.1145/2739480.2754687.

Li, M., & Yao, X. (2017). Dominance move: A measure of comparing solution sets in multiobjective optimization. *Technical Report*. arXiv.

Li, M., & Yao, X. (2019). Quality evaluation of solution sets in multiobjective optimisation: A survey. *ACM Computing Surveys, 52*(2), 26:1–26:38. https://doi.org/10.1145/3300148.

Li, M., Zheng, J., & Xiao, G. (2008). Uniformity assessment for evolutionary multi-objective optimization. In *IEEE congress on evolutionary computation (IEEE World Congress on Computational Intelligence)* (pp. 625–632). https://doi.org/10.1109/CEC.2008.4630861.

Liuzzi, G., Lucidi, S., & Rinaldi, F. (2016). A derivative-free approach to constrained multiobjective nonsmooth optimization. *SIAM Journal on Optimization, 26*(4), 2744–2774. https://doi.org/10.1137/15M1037810.

Lizarraga-Lizarraga, G., Hernandez-Aguirre, A., & Botello-Rionda, S. (2008). 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* (pp. 665–672). New York, NY, USA: ACM. https://doi.org/10.1145/1389095.1389227.

Lopez, E., & Coello, C. (2016). IGD+-EMOA: A multi-objective evolutionary algorithm based on IGD+. In *2016 IEEE Congress on Evolutionary Computation (CEC)* (pp. 999–1006). https://doi.org/10.1109/CEC.2016.7743898.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*(1), 77–91.

Martí, L., García, J., Berlanga, A., & Molina, J. (2016). A stopping criterion for multi-objective optimization evolutionary algorithms. *Information Sciences, 367–368*, 700–718. https://doi.org/10.1016/j.ins.2016.07.025.

Meng, H., Zhang, X., & Liu, S. (2005). New quality measures for multiobjective programming. In L. Wang, K. Chen, & Y. Ong (Eds.), *Advances in natural computation* (pp. 1044–1048). Berlin, Heidelberg: Springer. https://doi.org/10.1007/11539117_143.

Messac, A., & Mattson, C. (2004). Normal constraint method with guarantee of even representation of complete Pareto frontier. *AIAA Journal, 42*(10), 2101–2111. https://doi.org/10.2514/1.8977.

Moré, J., & Wild, S. (2009). Benchmarking derivative-free optimization algorithms. *SIAM Journal on Optimization, 20*(1), 172–191. https://doi.org/10.1137/080724083.

Mostaghim, S., & Teich, J. (2005). A new approach on many objective diversity measurement. In J. Branke, K. Deb, K. Miettinen, & R. Steuer (Eds.), *Practical approaches to multi-objective optimization*. In *Dagstuhl Seminar Proceedings*. Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

Nesterov, Y., & Nemirovskii, A. (1994). *Interior-point polynomial algorithms in convex programming*. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611970791.

Nocedal, J., & Wright, S. (2006). Numerical optimization. *Springer Series in Operations Research and Financial Engineering* (2nd). Berlin: Springer.

Okabe, T., Jin, Y., & Sendhoff, B. (2003). A critical survey of performance indices for multi-objective optimisation. In *Evolutionary computation: 2* (pp. 878–885). https://doi.org/10.1109/CEC.2003.1299759. Canberra, Australia

Riquelme, N., Lücken, C. V., & Baran, B. (2015). Performance metrics in multi-objective optimization. In *Latin American computing conference* (pp. 1–11). IEEE. https://doi.org/10.1109/CLEI.2015.7360024.

Rudolph, G., Schütze, O., Grimme, C., Domínguez-Medina, C., & Trautmann, H. (2016). Optimal averaged Hausdorff archives for bi-objective problems: theoretical and numerical results. *Computational Optimization and Applications, 64*(2), 589–618. https://doi.org/10.1007/s10589-015-9815-8.

Russo, L., & Francisco, A. (2014). Quick hypervolume. *IEEE Transactions on Evolutionary Computation, 18*(4), 481–502. https://doi.org/10.1109/TEVC.2013.2281525.

Russo, L., & Francisco, A. (2016). Extending quick hypervolume. *Journal of Heuristics, 22*(3), 245–271. https://doi.org/10.1007/s10732-016-9309-6.

Sayın, S. (2000). Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming. *Mathematical Programming, 87*(3), 543–560. https://doi.org/10.1007/s101070050011.

Schott, J. (1995). Fault tolerant design using single and multicriteria genetic algorithm optimization. *Technical Report*. Air force Institute Of Tech Wright-Patterson AFB OH.

Schutze, O., Esquivel, X., Lara, A., & Coello, C. (2012). Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization. *IEEE Transactions on Evolutionary Computation, 16*(4), 504–522. https://doi.org/10.1109/TEVC.2011.2161872.

Shannon, C. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review, 5*(1), 3–55. https://doi.org/10.1145/584091.584093.

Shir, O., Preuss, M., Naujoks, B., & Emmerich, M. (2009). Enhancing decision space diversity in evolutionary multiobjective algorithms. In M. Ehrgott, C. Fonseca, X. Gandibleux, J.-K. Hao, & M. Sevaux (Eds.), *Evolutionary multi-criterion optimization* (pp. 95–109). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-01020-0_12.

Shukla, P., Doll, N., & Schmeck, H. (2014). A theoretical analysis of volume based Pareto front approximations. In *Proceedings of the 2014 annual conference on genetic and evolutionary computation* (pp. 1415–1422). New York, NY, USA: ACM. https://doi.org/10.1145/2576768.2598348.

Slowinski, R., & Teghem, J. (1990). *Stochastic versus fuzzy approaches to multiobjective mathematical programming under uncertainty*. Springer. https://doi.org/10.1007/978-94-009-2111-5.

Tan, K., Lee, T., & Khor, E. (2002). Evolutionary algorithms for multi-objective optimization: performance assessments and comparisons. *Artificial Intelligence Review, 17*(4), 251–290. https://doi.org/10.1023/A:1015516501242.

Tian, Y., Zhang, X., Cheng, R., & Jin, Y. (2016). A multi-objective evolutionary algorithm based on an enhanced inverted generational distance metric. In *Proceed-*

ings of the 2016 IEEE congress on evolutionary computation (CEC'2016) (pp. 5222–5229). IEEE. https://doi.org/10.1109/CEC.2016.7748352.
Trautmann, H., Wagner, T., & Brockhoff, D. (2013). R2-EMOA: Focused multiobjective search using R2-indicator-based selection. In G. Nicosia, & P. Pardalos (Eds.), *Learning and intelligent optimization* (pp. 70–74). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-44973-4_8.
Ulrich, T., Bader, J., & Thiele, L. (2010a). Defining and optimizing indicator-based diversity measures in multiobjective search. In R. Schaefer, C. Cotta, J. Kołodziej, & G. Rudolph (Eds.), *Parallel problem solving from nature, PPSN XI* (pp. 707–717). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-15844-5_71.
Ulrich, T., Bader, J., & Zitzler, E. (2010b). Integrating decision space diversity into hypervolume-based multiobjective search. In *Proceedings of the 12th annual conference on genetic and evolutionary computation* (pp. 455–462). New York, NY, USA: ACM. https://doi.org/10.1145/1830483.1830569.
Van Veldhuizen, D. (1999). Multiobjective evolutionary algorithms: Classifications, analyses, and new innovations. *Technical Report*. School of Engineering of the Air Force Institute of Technology, Dayton, Ohio.
Van Veldhuizen, D., & Lamont, G. (2000). On measuring multiobjective evolutionary algorithm performance. In *Proceedings of the 2000 congress on evolutionary computation: 1* (pp. 204–211). https://doi.org/10.1109/CEC.2000.870296.
Vargas, A., & Bogoya, J. (2018). A generalization of the averaged Hausdorff distance. *Computación y Sistemas, 22*(2), 331–345. https://doi.org/10.13053/CyS-22-2-2950.
Wagner, T., Trautmann, H., & Brockhoff, D. (2013). Preference articulation by means of the R2 indicator. In R. Purshouse, J. Fleming, C. Fonseca, S. Greco, & J. Shaw (Eds.), *Evolutionary multi-criterion optimization* (pp. 81–95). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-37140-0_10.
Wagner, T., Trautmann, H., & Martí, L. (2011). A taxonomy of online stopping criteria for multi-objective evolutionary algorithms. In R. Takahashi, K. Deb, E. Wanner, & S. Greco (Eds.), *Evolutionary multi-criterion optimization* (pp. 16–30). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-19893-9_2.
Wagner, T., Trautmann, H., & Naujoks, B. (2009). OCD: Online convergence detection for evolutionary multi-objective algorithms based on statistical testing. In M. Ehrgott, C. Fonseca, X. Gandibleux, J.-K. Hao, & M. Sevaux (Eds.), *Evolutionary multi-criterion optimization* (pp. 198–215). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-01020-0_19.
While, L., Bradstreet, L., & Barone, L. (2012). A fast way of calculating exact hypervolumes. *IEEE Transactions on Evolutionary Computation, 16*(1), 86–95. https://doi.org/10.1109/TEVC.2010.2077298.
Wiecek, M., Ehrgott, M., & Engau, A. (2016). Continuous multiobjective programming. In *Multiple criteria decision analysis* (pp. 739–815). Springer New York. https://doi.org/10.1007/978-1-4939-3094-4_18.
Wu, J., & Azarm, S. (2000). Metrics for quality assessment of a multiobjective design optimization solution set. *Journal of Mechanical Design, 123*(1), 18–25. https://doi.org/10.1115/1.1329875.
Yevseyeva, I., Guerreiro, A., Emmerich, M., & Fonseca, C. (2014). A portfolio optimization approach to selection in multiobjective evolutionary algorithms. In *Parallel problem solving from nature – PPSN XIII* (pp. 672–681). https://doi.org/10.1007/978-3-319-10762-2_66.
Zhao, J., Jiao, L., Liu, F., Fernandes, V., Yevseyeva, I., Xia, S., & Emmerich, M. (2018). 3D fast convex-hull-based evolutionary multiobjective optimization algorithm. *Applied Soft Computing, 67*, 322–336. https://doi.org/10.1016/j.asoc.2018.03.005.
Zheng, K., Yang, R., Xu, H., & Hu, J. (2017). A new distribution metric for comparing Pareto optimal solutions. *Structural and Multidisciplinary Optimization, 55*(1), 53–62. https://doi.org/10.1007/s00158-016-1469-3.
Zhou, A., Jin, A., Zhang, Q., Sendhoff, B., & Tsang, E. (2006). Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion. In *IEEE international conference on evolutionary computation* (pp. 892–899). IEEE. https://doi.org/10.1109/CEC.2006.1688406.
Zitzler, E. (1999). *Evolutionary algorithms for multiobjective optimization: Methods and applications*. Zurich, Swiss Federal: Institute of Technology Ph.D. thesis..
Zitzler, E., Brockhoff, D., & Thiele, L. (2007). The hypervolume indicator revisited: on the design of Pareto-compliant indicators via weighted integration. In S. Obayashi, K. Deb, C. Poloni, T. Hiroyasu, & T. Murata (Eds.), *Evolutionary multi-criterion optimization* (pp. 862–876). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-70928-2_64.
Zitzler, E., Deb, K., & Thiele, L. (2000). Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary Computation, 8*(2), 173–195. https://doi.org/10.1162/106365600568202.
Zitzler, E., Knowles, J., & Thiele, L. (2008). Quality assessment of Pareto set approximations. In *Multiobjective optimization* (pp. 373–404). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-88908-3_14.
Zitzler, E., & Künzli, S. (2004). Indicator-based selection in multiobjective search. In X. Yao, E. Burke, J. Lozano, J. Smith, J. Merelo-Guervós, J. Bullinaria, … H.-S. Schwefel (Eds.), *Parallel problem solving from nature - PPSN VIII* (pp. 832–842). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-30217-9_84.
Zitzler, E., & Thiele, L. (1998). Multiobjective optimization using evolutionary algorithms – A comparative case study. In A. Eiben, T. Bäck, M. Schoenauer, & H. Schwefel (Eds.), *Parallel problem solving from nature – PPSN V* (pp. 292–301). Berlin, Heidelberg: Springer. https://doi.org/10.1007/BFb0056872.
Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C., & da Fonseca, V. (2003). Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation, 7*(2), 117–132. https://doi.org/10.1109/tevc.2003.810758.