Performance Assessment of Multiobjective Optimizers: An Analysis and Review

Eckart Zitzler, *Member, IEEE*, Lothar Thiele, Marco Laumanns, Carlos M. Fonseca, and Viviane Grunert da Fonseca

***Abstract*—An important issue in multiobjective optimization is the quantitative comparison of the performance of different algorithms. In the case of multiobjective evolutionary algorithms, the outcome is usually an approximation of the Pareto-optimal set, which is denoted as an approximation set, and therefore the question arises of how to evaluate the quality of approximation sets. Most popular are methods that assign each approximation set a vector of real numbers that reflect different aspects of the quality. Sometimes, pairs of approximation sets are considered too. In this study, we provide a rigorous analysis of the limitations underlying this type of quality assessment. To this end, a mathematical framework is developed which allows to classify and discuss existing techniques.**

***Index Terms*—Evolutionary algorithms, multiobjective optimization, performance assessment, quality indicator.**

I. Introduction

The main subject of this paper can be best understood if considering a setting as depicted in Fig. 1. We assume that a solution to the optimization problem at hand can be described in terms of a *decision vector* in the *decision space* X. The function $\boldsymbol{f}: X \rightarrow Z$ evaluates the quality of a specific solution by assigning it an *objective vector* in the *objective space* Z.

Now, let us suppose that the objective space is a subset of the real numbers, i.e., $Z \subseteq \mathbb{R}$, and that the goal of the optimization is to minimize the single objective. In such a single-objective optimization problem, a solution $\boldsymbol{x}^1 \in X$ is better than another solution $\boldsymbol{x}^2 \in X$, if $z^1 < z^2$ where $z^1 = f(\boldsymbol{x}^1)$ and $z^2 = f(\boldsymbol{x}^2)$. Although several optimal solutions may exist in decision space, they are all mapped to the same objective vector, i.e., there exists only a single optimum in objective space.

In the case of a vector-valued evaluation function $\boldsymbol{f}$ with $Z \subseteq \mathbb{R}^n$, where $n > 1$, the situation of comparing two solutions $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ is more complex. Following the well-known concept of Pareto dominance, we can say that $\boldsymbol{x}^1$ *dominates* $\boldsymbol{x}^2$ if no component of $\boldsymbol{z}^1$ is larger than the corresponding component of $\boldsymbol{z}^2$ and at least one component is smaller. Here, optimal solutions, i.e., solutions not dominated by any other solution, may be mapped to different objective vectors. In other words: there may exist several optimal objective vectors representing different tradeoffs between the objectives.

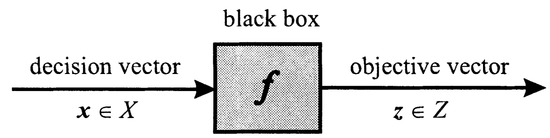


Fig. 1. Typical black box optimization problem where elements of the decision space need to be determined such that the components of the corresponding objective vector are optimal under the mapping $\boldsymbol{f}$.

The set of optimal solutions in the decision space X is in general denoted as *Pareto-optimal set*. With many multiobjective optimization problems, knowledge about this set helps the decision maker in choosing the best compromise solution. For instance, when designing computer systems, engineers often perform a so-called design space exploration to learn more about the Pareto-optimal set. Thereby, the design space is reduced to the set of optimal trade-offs: a first step in selecting an appropriate implementation.

However, generating the Pareto-optimal set can be computationally expensive and is often infeasible, because the complexity of the underlying application prevents exact methods from being applicable. Evolutionary algorithms (EAs) are an alternative: they usually do not guarantee to identify optimal tradeoffs but try to find a good approximation, i.e., a set of solutions whose objective vectors are (hopefully) not too far away from the optimal objective vectors. Various multiobjective EAs are available, and certainly we are interested in the technique that provides the best approximation for a given problem. However, in order to reveal strengths and weaknesses of certain approaches and to identify the most promising techniques, existing algorithms have to be compared—either empirically, e.g., [1]–[4], or theoretically, e.g., [5]. This, in turn, directly leads to the issue of assessing the performance of multiobjective optimizers.

The notion of performance includes both the quality of the outcome as well as the computational resources needed to generate this outcome. Concerning the latter aspect, it is common practice to monitor either the number of fitness evaluations or the overall run-time on a particular computer—in this sense, there is no difference between single and multiobjective optimization. As to the quality aspect, however, there is a difference. In single-objective optimization, we can define quality

Manuscript received June 14, 2002; revised September 30, 2002 and October 31, 2002. This work was supported by the Swiss National Science Foundation (SNF) under the ArOMA Project 2100-057156.99/1 and by the Portuguese Foundation for Science and Technology under the POCTI programme (Project POCTI/MAT/10135/98), co-financed by the European Regional Development Fund.

E. Zitzler, L. Thiele, and M. Laumanns are with the Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH), Zurich, Switzerland (e-mail: zitzler@tik.ee.ethz.ch; thiele@tik.ee.ethz.ch; laumanns@tik.ee.ethz.ch).

C. M. Fonseca is with the ADEEC and ISR (Coimbra), Faculty of Science and Technology, University of Algarve, Faro, Portugal (e-mail: cmfonsec@ualg.pt).

V. Grunert da Fonseca is with the CSI, Faculty of Science and Technology, University of Algarve, Faro, Portugal, and also with the INUAF, Loulé, Portugal (e-mail: vgrunert@csi.fct.ualg.pt).

Digital Object Identifier 10.1109/TEVC.2003.810758

1089-778X/03$17.00 © 2003 IEEE

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

by means of the objective function: the smaller (or larger) the value, the better the solution. If we compare two solutions in the presence of multiple optimization criteria, the concept of Pareto dominance can be used, although the possibility of two solutions being incomparable, i.e., neither dominates the other, complicates the situation. However, it gets even more complicated when we compare two sets of solutions because some solutions in either set may be dominated by solutions in the other set, while others may be incomparable. Accordingly, it is not clear what quality means with respect to approximations of the Pareto-optimal set: closeness to the optimal solutions in objective space, coverage of a wide range of diverse solutions, or other properties? It is difficult to define appropriate quality measures for approximations of the Pareto-optimal set, and as a consequence graphical plots have been used to compare the outcomes of multiobjective EAs until recently, as Van Veldhuizen and Lamont point out [2].

Progress, though, has been made, and several studies can be found in the literature that address the problem of comparing approximations of the Pareto-optimal set in a quantitative manner. Most popular are unary quality measures, i.e., the measure assigns each approximation set a number that reflects a certain quality aspect, and usually a combination of them is used, e.g., [2], [6]. Other methods are based on binary quality measures, which assign numbers to pairs of approximation sets, e.g., [7] and [8]. A third and conceptually different method is the attainment function approach [9], which consists of estimating the probability of attaining arbitrary goals in objective space from multiple approximation sets. Despite of this variety, it has remained unclear up to now how the different measures are related to each other and what their advantages and disadvantages are. Accordingly, there is no common agreement on which measure(s) should be used.

Recently, a few studies have been carried out to clarify this situation. Hansen and Jaszkiewicz [8] studied and proposed some quality measures that allow to incorporate knowledge about the decision maker's preferences. They first introduced three different "outperformance" relations for multiobjective optimizers and then investigated whether the measures under consideration are compliant with these relations. The question they considered was: whenever an approximation is better than another according to an "outperformance" relation, does the comparison method also evaluate the former as being better (or at least not worse) than the latter? More from a practical point of view, Knowles *et al.* [10] compared the information provided by different assessment techniques on two database management applications. More recently, Knowles [11] and Knowles and Corne [12] discussed and contrasted several commonly used quality measures in light of Hansen and Jaszkiewicz's approach, as well as according to other criteria such as, e.g., sensitivity to scaling. They showed that about one third of the investigated quality measures are not compliant with any of the "outperformance" relations introduced by Hansen and Jaszkiewicz.

This paper takes a different perspective that allows a more rigorous analysis and classification of comparison methods. In contrast to [8], [11], and [12], we focus on the statements that can be made on the basis of the information provided by quality measures. Is it, for instance, possible to conclude from the quality "measurements" that an approximation A is undoubtedly better than approximation B in the sense that A, loosely speaking, entirely dominates B? This is a crucial issue in any comparative study, and implicitly most papers in this area rely on the assumption that this property is satisfied for the measures used. To investigate quality measures from this perspective, a formal framework will be introduced that substantially goes beyond Hansen and Jaszkiewicz's approach, as well as that of Knowles and Corne; e.g., it will enable us to consider combinations of quality measures and to prove theoretical limitations of unary quality measures, both issues not addressed in [8], [11], and [12]. In detail, we will show that:

- there exists no unary quality measure that is able to indicate *whether* an approximation A is better than an approximation B;
- the above statement even holds if we consider a finite combination of unary measures;
- most quality measures that have been proposed to indicate *that* A is better than B at best allow to infer that A is not worse than B, i.e., A is better than or incomparable to B;
- unary measures being able to detect *that* A is better than B exist, but their use is in general restricted;
- binary quality measures overcome the limitations of unary measures and, if properly designed, are capable of indicating *whether* A is better than B.

Furthermore, we will review existing quality measures in light of this framework and discuss them from a practical point of view also. Note that we focus on the comparison of approximation sets rather than on algorithms, i.e., we assume that for each multiobjective EA only one run is performed. In the case of multiple runs, the distribution of the indicator values would have to be considered instead of the values themselves; this important issue will not be addressed in the present paper.

II. Theoretical Framework

Before analyzing and classifying quality measures, we must clarify the concepts we will be dealing with: what is the outcome of a multiobjective optimizer, when is an outcome considered to be better than another, what is a quality measure, what is a comparison method, etc.? These terms will be formally defined in this section.

A. Approximation Sets

The scenario considered in this paper involves an arbitrary optimization problem with n objectives $f_1, f_2, \ldots, f_n$, which are, without loss of generality, all to be minimized and all equally important, i.e., no additional knowledge about the problem is available. The only assumption we make is that a solution $\boldsymbol{x}^1$ is preferable to another solution $\boldsymbol{x}^2$ if $\boldsymbol{x}^1$ dominates $\boldsymbol{x}^2$. Furthermore, for the purpose of this paper it is sufficient to deal with the objective vector $\boldsymbol{z} \in Z$ corresponding to a particular solution $\boldsymbol{x} \in X$ in decision space. Therefore, we will use the aforementioned concepts, such as Pareto dominance, Pareto-optimal set, and approximation set solely in terms of the objective space in the following. For reasons of consistency and simplicity, we will also assume that for each objective vector

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

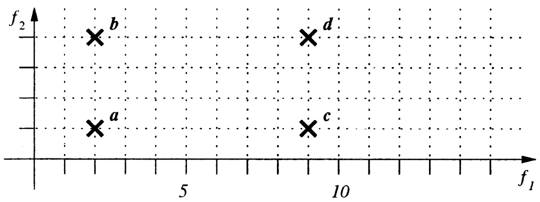


Fig. 2. Examples of dominance relations on objective vectors. Assuming that two objectives are to be minimized, it holds that $\boldsymbol{a} \succ \boldsymbol{b}$, $\boldsymbol{a} \succ \boldsymbol{c}$, $\boldsymbol{a} \succ \boldsymbol{d}$, $\boldsymbol{b} \succ \boldsymbol{d}$, $\boldsymbol{c} \succ \boldsymbol{d}$, $\boldsymbol{a} \succ\succ \boldsymbol{d}$, $\boldsymbol{a} \succeq \boldsymbol{a}$, $\boldsymbol{a} \succeq \boldsymbol{b}$, $\boldsymbol{a} \succeq \boldsymbol{c}$, $\boldsymbol{a} \succeq \boldsymbol{d}$, $\boldsymbol{b} \succeq \boldsymbol{b}$, $\boldsymbol{b} \succeq \boldsymbol{d}$, $\boldsymbol{c} \succeq \boldsymbol{c}$, $\boldsymbol{c} \succeq \boldsymbol{d}$, $\boldsymbol{d} \succeq \boldsymbol{d}$, and $\boldsymbol{b} \parallel \boldsymbol{c}$.

$\boldsymbol{z} = (z_1, z_2, \ldots, z_n) \in Z$, there is a decision vector $\boldsymbol{x} \in X$, with $\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_n(\boldsymbol{x}))$.

Analogously to solutions, we say an objective vector $\boldsymbol{z}^1$ dominates another objective vector $\boldsymbol{z}^2$ if $\boldsymbol{z}^1$ is not greater than $\boldsymbol{z}^2$ in all components *and* has a smaller value in at least one component. An objective vector is denoted as Pareto-optimal if it is not dominated by any other objective vector, and the entirety of all Pareto-optimal objective vectors forms the Pareto-optimal set in objective space. Fig. 2 visualizes the concept of Pareto dominance and also gives some examples for other common relations on pairs of objective vectors. Table I comprises a summary of the relations used in this paper. Note that there exists a natural ordering of these relations as $\boldsymbol{z}^1 \succ\succ \boldsymbol{z}^2 \Rightarrow \boldsymbol{z}^1 \succ \boldsymbol{z}^2 \Rightarrow \boldsymbol{z}^1 \succeq \boldsymbol{z}^2$. In addition, note that if $\boldsymbol{z}^1$ is not incomparable to $\boldsymbol{z}^2$, then either $\boldsymbol{z}^1 \succeq \boldsymbol{z}^2$ or $\boldsymbol{z}^1 \preceq \boldsymbol{z}^2$, i.e., $\boldsymbol{z}^1 \not\parallel \boldsymbol{z}^2 \Rightarrow \boldsymbol{z}^1 \succeq \boldsymbol{z}^2 \wedge \boldsymbol{z}^1 \preceq \boldsymbol{z}^2$.

The vast majority of papers in the area of evolutionary multiobjective optimization is concerned with the problem of how to identify the Pareto-optimal solutions or, if this is infeasible, to generate good approximations of them. Taking this as the basis of our study, we here consider the outcome of a multiobjective EA (or other heuristic) as a set of incomparable solutions, which will be denoted as approximation set [8]. In terms of the objective space, this can be formalized as follows.

Definition 1 (Approximation Set): Let $A \subseteq Z$ be a set of objective vectors. A is called an *approximation set* if any element of A does not weakly dominate any other objective vector in A. The set of all approximation sets is denoted as Ω.

The motivation behind this definition is that all solutions dominated by any other solution outputted by the optimization algorithm are of no interest, and therefore can be discarded. In objective space, this means we can neglect dominated objective vectors, which will simplify the considerations in the following sections.

Note that the above definition does not comprise any notion of quality. We are certainly not interested in *any* approximation set, but we want the EA to generate a *good* approximation set. The ultimate goal is to identify the Pareto-optimal set. This aim, however, is usually not achievable. Moreover, it is impossible to exactly describe what a good approximation is in terms of a number of criteria such as closeness to the Pareto-optimal set, diversity, etc.—this will be shown in Section III-A. However, we can make statements about the quality of approximation sets in comparison to other approximation sets.

Consider, e.g., the outcomes of three hypothetical algorithms as depicted in Fig. 3. Solely on the basis of Pareto dominance, one can state that A_1 and A_2 both dominate A_3 as any objective vector in A_3 is dominated by at least one vector in A_1 and A_2. Furthermore, A_1 can be considered better than A_2 as it contains all objective vectors in A_2 and another vector not included in A_2, although this statement is weaker than the previous one. Accordingly, we will distinguish four relations in this paper as defined in Table I: A strictly dominates B ($A \succ\succ B$), A dominates B ($A \succ B$), A is better than B ($A \rhd B$), and A weakly dominates B ($A \succeq B$). Note that there is a natural ordering again among the relations as $A \succ\succ B \Rightarrow A \succ B \Rightarrow A \rhd B \Rightarrow A \succeq B$.

Weak dominance ($A \succeq B$) means that any objective vector in B is weakly dominated by a vector in A. However, this does not rule out equality, because $A \succeq A$ for all approximation sets $A \in \Omega$. In this case, one cannot say that A is better than B. Instead, the relation $\rhd$ can be used. It requires that an approximation set is at least as good as another approximation set ($A \succeq B$), while the latter is not as good as the former ($B \not\succeq A$), roughly speaking. We can also conclude from the definition of the relation $\rhd$ that $A \succeq B \Rightarrow A \rhd B \vee A = B$. In other words, if A weakly dominates B, then either A is better than B or they are equal.

In the example, A_1 is better than A_2 and A_3, and A_2 is better than A_3. This definition of superiority is the one implicitly used in most papers in the field. The next level of superiority, the $\succ$ relation, is a straightforward extension of Pareto dominance to approximation sets. It does not allow that two objective vectors in A and B are equal, and therefore is stricter than what we usually require. As mentioned above, A_1 and A_2 dominate A_3, but A_1 does not dominate A_2. Strict dominance stands for the highest level of superiority and means an approximation set is superior to another approximation set in the sense that for any objective vector in the latter there exists a vector in the former that is better in all objectives. In Fig. 3, A_1 strictly dominates A_3, but A_2 does not as the objective vector (10, 4) is not strictly dominated by any objective vector in A_2.

These relations (cf. Table I) and their ordering can also be visualized using a diagram as depicted in Fig. 4. Each pair $(A, B) \in \Omega^2$ can be associated uniquely to one of the regions shown.

B. Comparison Methods

Quality measures have been introduced to compare the outcomes of multiobjective optimizers in a quantitative manner. Certainly, the simplest comparison method would be to check whether an outcome is better than another with respect to the three dominance relations $\rhd$, $\succ$, and $\succ\succ$. We have demonstrated this in the context of the discussion of Fig. 3. The reason, however, why quality measures have been used is to be able to make more precise statements in addition to that, which are inevitably based on certain assumptions about the decision maker's preferences:

- If one algorithm is better than another, can we express how much better it is?
- If no algorithm can be said to be better than the other, are there certain aspects in which respect we can say the former is better than the latter?

Hence, the key question when designing quality measures is how to best summarize approximation sets by means of a few

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

TABLE I
RELATIONS ON OBJECTIVE VECTORS AND APPROXIMATION SETS CONSIDERED IN THIS PAPER. THE RELATIONS $\prec$, $\prec\prec$, $\lhd$, AND $\preceq$ ARE DEFINED ACCORDINGLY, E.G., $z^1 \prec z^2$ IS EQUIVALENT TO $z^2 \succ z^1$ AND $A \lhd B$ IS DEFINED AS $B \rhd A$

relation	objective vectors		approximation sets	
strictly dominates	$\boldsymbol{z}^1 \succ\succ \boldsymbol{z}^2$	$\boldsymbol{z}^1$ is better than $\boldsymbol{z}^2$ in all objectives	$A \succ\succ B$	every $\boldsymbol{z}^2 \in B$ is strictly dominated by at least one $\boldsymbol{z}^1 \in A$
dominates	$\boldsymbol{z}^1 \succ \boldsymbol{z}^2$	$\boldsymbol{z}^1$ is not worse than $\boldsymbol{z}^2$ in all objectives and better in at least one objective	$A \succ B$	every $\boldsymbol{z}^2 \in B$ is dominated by at least one $\boldsymbol{z}^1 \in A$
better			$A \rhd B$	every $\boldsymbol{z}^2 \in B$ is weakly dominated by at least one $\boldsymbol{z}^1 \in A$ and $A \neq B$
weakly dominates	$\boldsymbol{z}^1 \succeq \boldsymbol{z}^2$	$\boldsymbol{z}^1$ is not worse than $\boldsymbol{z}^2$ in all objectives	$A \succeq B$	every $\boldsymbol{z}^2 \in B$ is weakly dominated by at least one $\boldsymbol{z}^1 \in A$
incomparable	$\boldsymbol{z}^1 \parallel \boldsymbol{z}^2$	neither $\boldsymbol{z}^1$ weakly dominates $\boldsymbol{z}^2$ nor $\boldsymbol{z}^2$ weakly dominates $\boldsymbol{z}^1$	$A \parallel B$	neither A weakly dominates B nor B weakly dominates A

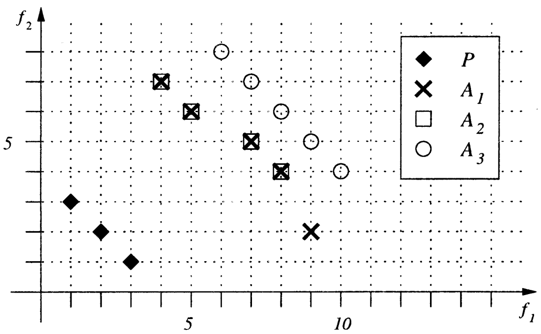


Fig. 3. Outcomes of three hypothetical algorithms for a 2-D minimization problem. The corresponding approximation sets are denoted as A_1, A_2, and A_3; the Pareto-optimal set P comprises three objective vectors. Between A_1, A_2, and A_3, the following dominance relations hold: $A_1 \succ A_3$, $A_2 \succ A_3$, $A_1 \succ\succ A_3$, $A_1 \succeq A_1$, $A_1 \succeq A_2$, $A_1 \succeq A_3$, $A_2 \succeq A_2$, $A_2 \succeq A_3$, $A_3 \succeq A_3$, $A_1 \rhd A_2$, $A_1 \rhd A_3$, and $A_2 \rhd A_3$.

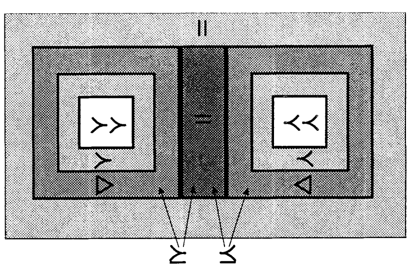


Fig. 4. Partitioning of the set of ordered pairs $(A, B) \in \Omega^2$ of approximation sets into (overlapping) subsets induced by the different dominance relations; each subset labeled with a certain relation ► contains those pairs (A, B) for which A ► B. Note that the set of all pairs (A, B) with $A \succeq B$ is the union of those with $A = B$ and $A \rhd B$.

characteristic numbers—similarly to statistics where the mean, the standard deviation, etc. are used to describe a probability distribution in a compact way. It is unavoidable to lose information by such a reduction, and the crucial point is not to lose the information one is interested in.

There are many examples of quality measures in the literature. Some aim at measuring the distance of an approximation set to the Pareto-optimal set: Van Veldhuizen [13], e.g., calculated for each solution in the approximation set under consideration the Euclidean distance to the closest Pareto-optimal objective vector and then took the average over all of these distances. Other measures try to capture the diversity of an approximation set, e.g., the chi-square-like deviation measure used by Srinivas and Deb [14]. A further example is the hypervolume measure, which considers the volume of the objective space dominated by an approximation set [7]. In these three cases, an approximation set is assigned a real number which is meant to reflect (certain aspects of) the quality of an approximation set. Alternatively, one can assign numbers to pairs of approximation sets. Zitzler and Thiele [7], e.g., introduced the coverage function which gives for a pair (A, B) of approximation sets the fraction of solutions in B that are weakly dominated by one or more solutions in A.

In summary, we can state that quality measures map approximation sets to the set of real numbers. The underlying idea is to quantify quality differences between approximation sets by applying common metrics (in the mathematical sense) to the resulting real numbers. This observation enables us to formally define what a quality measure is; however, we will use the term "quality indicator" in the following as "measure" is often used with different meanings.

Definition 2 (Quality Indicator): An m-ary *quality indicator* I is a function $I\colon \Omega^m \rightarrow \mathbb{R}$, which assigns each vector $(A_1, A_2, \ldots, A_m)$ of m approximation sets a real value $I(A_1, \ldots, A_m)$.

The measures discussed above are examples for unary and binary quality indicators; however, in principle, a quality indicator can take an arbitrary number of arguments. Thereby, other comparison methods that explicitly account for multiple runs and involve statistical testing procedures [15], [16], [9] can also be expressed within this framework. Furthermore, not a single indicator but rather a combination of different quality indicators is often used in order to assess approximation sets. Van Veldhuizen and Lamont [2], for instance, applied a combination $\boldsymbol{I} = (I_{\mathrm{GD}}, I_S, I_{\mathrm{ONVG}})$ of three indicators, where $I_{\mathrm{GD}}(A)$ denotes the average distance of objective vectors in A to the Pareto-optimal set, $I_S(A)$ measures the variance of distances between neighboring objective vectors in A, and $I_{\mathrm{ONVG}}(A)$ gives the number of elements in A. Accordingly, the combination (or quality indicator vector) $\boldsymbol{I}$ can be regarded as a function that assigns each approximation set a triple of real numbers.

Quality indicators, though, need interpretation. In particular, we would like to formally describe statements such as "if and only if $I_{\mathrm{GD}}(A) = 0$, then all objective vectors in A have zero distance to the Pareto-optimal set P, and therefore $A \subseteq P$ and also $B \not\succeq A$ for any approximation set $B \not\subseteq P$." To this end, we introduce two concepts. An interpretation function E

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

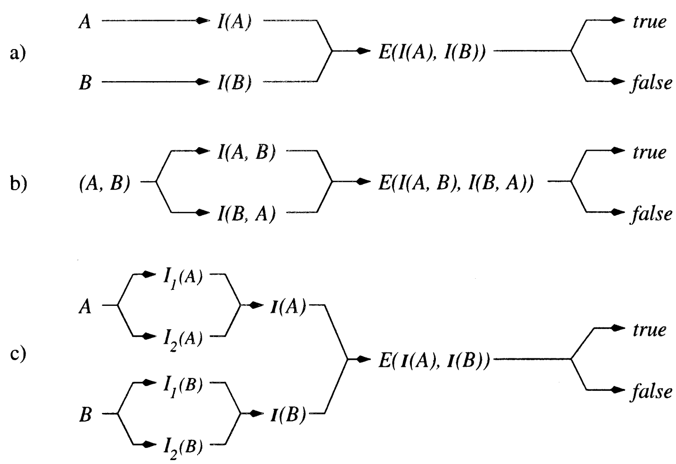


Fig. 5. Illustration of the concept of a comparison method for (a) a single unary quality indicator, (b) a single binary quality indicator, and (c) a combination of two unary quality indicators. In cases (a) and (b), first the indicator I is applied to the two approximation sets A, B. The resulting two real values are passed to the interpretation function E, which defines the outcome of the comparison. In case (c), each of the two indicators is applied to A and B, and the resulting two indicator values are combined in a vector $\boldsymbol{I}(A)$ and $\boldsymbol{I}(B)$ respectively. Afterwards, the interpretation function E decides the outcome of the comparison on the basis of these two real vectors.

maps vectors of real numbers to Booleans. In the above example, we would define $E(I_{\mathrm{GD}}(A), I_{\mathrm{GD}}(B)) := (I_{\mathrm{GD}}(A) = 0 \wedge I_{\mathrm{GD}}(B) > 0)$, i.e., E is true if and only if $I_{\mathrm{GD}}(A) = 0$ and at the same time $I_{\mathrm{GD}}(B) > 0$. Such a combination of one or more quality indicators $\boldsymbol{I}$ and an interpretation function E is also called a comparison method $C_{\boldsymbol{I}, E}$. In the example, the comparison method $C_{I_{\mathrm{GD}}, E}$ based on I_{GD} and E would be defined as $C_{I_{\mathrm{GD}}, E}(A, B) = E(I_{\mathrm{GD}}(A), I_{\mathrm{GD}}(B))$, and the conclusion is that $C_{I_{\mathrm{GD}}, E}(A, B) \Leftrightarrow A \subseteq P \wedge B \not\subseteq P \wedge B \not\succeq A$. In the following, we will focus on comparison methods that: 1) consider two approximation sets only and 2) use either only unary or only binary indicators (cf. Fig. 5).

Definition 3 (Comparison Method): Let $A, B \in \Omega$ be two approximation sets, $\boldsymbol{I} = (I_1, I_2, \ldots, I_k)$ a combination of quality indicators, and $E: \mathbb{R}^k \times \mathbb{R}^k \rightarrow \{false, true\}$ an interpretation function which maps two real vectors of length k to a Boolean value. If all indicators in $\boldsymbol{I}$ are unary, the *comparison method* $C_{\boldsymbol{I}, E}$ defined by $\boldsymbol{I}$ and E is a function of the form

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

where $\boldsymbol{I}(A') = (I_1(A'), I_2(A'), \ldots, I_k(A'))$ for $A' \in \Omega$. If $\boldsymbol{I}$ contains only binary indicators, the *comparison method* $C_{\boldsymbol{I}, E}$ is defined as

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

where $\boldsymbol{I}(A', B') = (I_1(A', B'), I_2(A', B'), \ldots, I_k(A', B'))$ for $A', B' \in \Omega$.

Whenever we specify a particular comparison method $C_{\boldsymbol{I}, E}$, we will write $E := \langle expression \rangle$ instead of $E(\cdots) \Leftrightarrow \langle expression \rangle$ in order to improve readability. For instance, $E := (I_1(A) > I_1(B))$ means that $E(\boldsymbol{I}(A), \boldsymbol{I}(B))$ is true if and only if $I_1(A) > I_1(B)$, given a combination of k unary indicators.

Definition 3 may appear overly formal for describing what a comparison method basically is, and furthermore it does not specify the actual conclusion (what does it mean if $C_{\boldsymbol{I}, E}(A, B)$ is true?). As we will see in the following, however, it provides a sound basis for studying the power of quality indicators—the power of indicating relationships (better, incomparable, etc.) between approximation sets.

C. Linking Comparison Methods and Dominance Relations

The goal of a comparative study is to reveal differences in performance between multiobjective optimizers, and the strongest statement we can make in this context is that an algorithm outperforms another one. Independently of what definition of "outperformance" we use, it always should be compliant with the most general notion in terms of the $\triangleright$-relation, i.e., the statement "algorithm a outperforms algorithm b" should also imply that the outcome A of the first method is better than the outcome B of the second method ($A \triangleright B$).[1]

In this paper, we are interested in the question what conclusions can be drawn with respect to the dominance relations listed in Table I on the basis of a comparison method $C_{\boldsymbol{I}, E}$. If $C_{\boldsymbol{I}, E}(A, B)$ is a sufficient condition for, e.g., $A \triangleright B$, then this comparison method is capable of indicating *that* A is better than B, i.e., $C_{\boldsymbol{I}, E}(A, B) \Rightarrow A \triangleright B$. If $C_{\boldsymbol{I}, E}(A, B)$ is, in addition, a necessary condition for $A \triangleright B$, then the comparison method even indicates *whether* A is better than B, i.e., $C_{\boldsymbol{I}, E}(A, B) \Leftrightarrow A \triangleright B$. In the following, we will use the terms compatibility and completeness in order to characterize a comparison method in the above manner.

Definition 4 (Compatibility and Completeness): Let $\blacktriangleright$ be an arbitrary binary relation on approximation sets, cf. Table I. The comparison method $C_{\boldsymbol{I}, E}$ is denoted as $\blacktriangleright$-*compatible* if either for any $A, B \in \Omega$

$$C_{\boldsymbol{I}, E}(A, B) \Rightarrow A \blacktriangleright B$$

or for any $A, B \in \Omega$

$$C_{\boldsymbol{I}, E}(A, B) \Rightarrow B \blacktriangleright A.$$

The comparison method $C_{\boldsymbol{I}, E}$ is denoted as $\blacktriangleright$-*complete* if either for any $A, B \in \Omega$

$$A \blacktriangleright B \Rightarrow C_{\boldsymbol{I}, E}(A, B)$$

or for any $A, B \in \Omega$

$$B \blacktriangleright A \Rightarrow C_{\boldsymbol{I}, E}(A, B).$$

For instance, suppose we have a comparison method that is $\triangleright$-*complete* but not compatible with respect to the $\triangleright$ relation. If we use this comparison method to compare two sets A and B with $A \triangleright B$, i.e., A is better than B, than our comparison method indicates that correctly. However, there are also sets A and B with $A \not\triangleright B$, i.e., A is not better than B, for which the comparison method returns true. A comparison method that is complete with regard to any relation is the one that always yields

[1]Recall that we assume that only a single optimization run is performed per algorithm.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

true; it is useless, though, as it does not provide any compatibility.

On the other hand, if we have a comparison method that is $\triangleright$*-compatible*, then the above situation is safe: whenever our comparison method yields true, we can be sure that A is better than B. However, we may miss opportunities, if the comparison method is not $\triangleright$-complete. In particular, there may be sets A and B where A is better than B, but our comparison method returns false. A comparison method that always yields false is compatible and not complete regarding any relation.

To further illustrate this terminology, let us go back to the example depicted in Fig. 3 and consider the following binary indicator I_ϵ, which is inspired by concepts presented in [17].

Definition 5 (Binary ϵ-Indicator): Suppose, without loss of generality, a minimization problem with n positive objectives, i.e., $Z \subseteq \mathbb{R}^{+^n}$. An objective vector $\boldsymbol{z}^1 = (z_1^1, z_2^1, \ldots, z_n^1) \in Z$ is said to *ϵ-dominate* another objective vector $\boldsymbol{z}^2 = (z_1^2, z_2^2, \ldots, z_n^2) \in Z$, written as $\boldsymbol{z}^1 \succeq_\epsilon \boldsymbol{z}^2$, if and only if

$$\forall 1 \leq i \leq n\colon z_i^1 \leq \epsilon \cdot z_i^2$$

for a given $\epsilon > 0$. We define the *binary ϵ-indicator I_ϵ* as

$$I_\epsilon(A, B) = \inf_{\epsilon \in \mathbb{R}} \{\forall \boldsymbol{z}^2 \in B\, \exists \boldsymbol{z}^1 \in A\colon \boldsymbol{z}^1 \succeq_\epsilon \boldsymbol{z}^2\}$$

for any two approximation sets $A, B \in \Omega$.[2]

Loosely speaking, a vector $\boldsymbol{z}^1$ is said to ϵ-dominate another vector $\boldsymbol{z}^2$, if we can multiply each objective value in $\boldsymbol{z}^2$ by a factor of ϵ and the resulting objective vector is still weakly dominated by $\boldsymbol{z}^1$. Therefore, $\boldsymbol{z}^1 \succ\succ \boldsymbol{z}^2$ implies that there exists $\epsilon < 1$ such that $\boldsymbol{z}^1$ ϵ-dominates $\boldsymbol{z}^2$. Accordingly, the ϵ-indicator gives the factor by which an approximation set is worse than another with respect to all objectives, or to be more precise: $I_\epsilon(A, B)$ equals the minimum factor ϵ such that any objective vector in B is ϵ-dominated by at least one objective vector in A. In the single-objective case, $I_\epsilon(A, B)$ is simply the ratio between the two objective values represented by A and B.

In practice, the binary ϵ-indicator $I_\epsilon(A, B)$ can be calculated in time $O(n \cdot |A| \cdot |B|)$ as follows:

$$\begin{aligned}
\epsilon_{\boldsymbol{z}^1, \boldsymbol{z}^2} &= \max_{1 \leq i \leq n} \frac{z_i^1}{z_i^2} \qquad \forall \boldsymbol{z}^1 \in A, \boldsymbol{z}^2 \in B \\
\epsilon_{\boldsymbol{z}^2} &= \min_{\boldsymbol{z}^1 \in A} \epsilon_{\boldsymbol{z}^1, \boldsymbol{z}^2} \qquad \forall \boldsymbol{z}^2 \in B \\
I_\epsilon(A, B) &= \max_{\boldsymbol{z}^2 \in B} \epsilon_{\boldsymbol{z}^2}
\end{aligned}$$

or equivalently

$$I_\epsilon(A, B) = \max_{\boldsymbol{z}^2 \in B} \min_{\boldsymbol{z}^1 \in A} \max_{1 \leq i \leq n} \frac{z_i^1}{z_i^2}.$$

For instance, $I_\epsilon(A_1, A_2) = 1$, $I_\epsilon(A_1, A_3) = 9/10$, and $I_\epsilon(A_1, P) = 4$ in our previous example (cf. Fig. 6). The complete table for all indicators is given in Table II.

[2]In the same manner, an additive ϵ-indicator $I_{\epsilon+}$ can be defined

$$I_{\epsilon+}(A, B) = \inf_{\epsilon \in \mathbb{R}} \{\forall z^2 \in B\, \exists z^1 \in A\colon z^1 \succeq_{\epsilon+} z^2\}$$

where $z^1 \succeq_{\epsilon+} z^2$ if and only if

$$\forall 1 \leq i \leq n\colon z_i^1 \leq \epsilon + z_i^2.$$

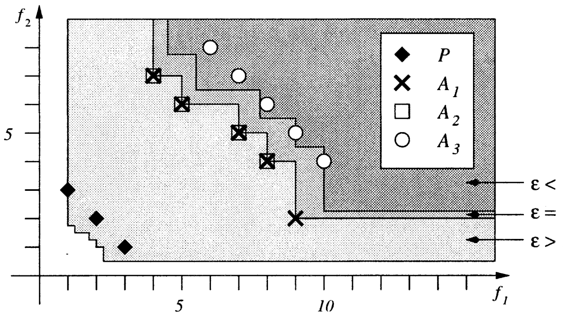


Fig. 6. The dark-shaded area depicts the subspace that is ϵ-dominated by the solutions in A_1 for $\epsilon = 9/10$; the medium-shaded area represents the subspace weakly dominated by A_1 (equivalent to $\epsilon = 1$). The light-shaded area refers to the subspace ϵ-dominated by the solutions in A_1 for $\epsilon = 4$. Note that the areas are overlapping, i.e., the medium-shaded area, includes the dark-shaded one, and the light-shaded area includes both of the other areas.

TABLE II
THE BINARY ϵ-INDICATOR VALUES $I_\epsilon(A, B)$ FOR ALL COMBINATIONS OF THE SETS $A1$, $A2$, $A3$ AND P AS GIVEN IN FIG. 6

B	A			
	A_1	A_2	A_3	P
A_1	1	2	2	1/2
A_2	1	1	3/2	3/7
A_3	9/10	1	1	1/3
P	4	4	6	1

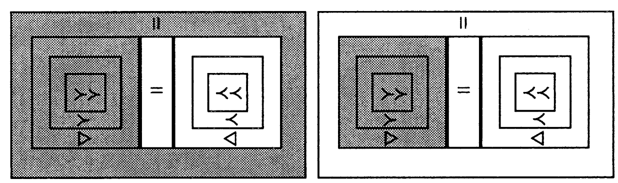


Fig. 7. The shaded area stands for those ordered pairs (A, B) for which $I_\epsilon(B, A) > 1$ (left) or $I_\epsilon(A, B) \leq 1 \wedge I_\epsilon(B, A) > 1$ (right), respectively. Note that the right comparison method is both $\triangleright$-compatible and $\triangleright$-complete.

What comparison methods can be constructed using the ϵ-indicator? Consider, e.g., the interpretation function $E := (I_\epsilon(B, A) > 1)$. The corresponding comparison method $C_{I_\epsilon, E}$ is $\triangleright$-complete as $A \triangleright B$ implies that $I_\epsilon(B, A) > 1$. On the other hand, $C_{I_\epsilon, E}$ is not $\triangleright$-compatible as $A \parallel B$ also implies that $I_\epsilon(B, A) > 1$. This is visualized in Fig. 7, where we see on the left-hand side the area for which $I_\epsilon(B, A) > 1$. It becomes obvious that the indicator value is greater than one even if the two sets are incomparable.

If we choose a slightly modified interpretation function $F := (I_\epsilon(A, B) \leq 1 \wedge I_\epsilon(B, A) > 1)$, then we obtain a comparison method $C_{I_\epsilon, F}$ that is both $\triangleright$- and $\triangleright$-complete. The differences between the two comparison methods are graphically depicted in Fig. 7. On the right-hand side, we see that the new comparison method exactly characterizes all pairs (A, B) for which A is better than B, i.e., $A \triangleright B$.

In the remainder of this paper, we will theoretically study and classify quality indicators using the above framework. Given a particular quality indicator (or a combination of several indicators), we will investigate whether there exists an interpretation function such that the resulting comparison method is compatible and in addition complete with respect to the various dominance relations. That is, we determine how powerful existing

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

quality indicators are in terms of their capability of indicating *that* or *whether* $A \triangleright B$, $A \succ B$, $A \parallel B$, etc. The next section is devoted to unary quality indicators, while binary indicators will be discussed in Section IV.

III. Comparison Methods Based on Unary Quality Indicators

Unary quality indicators are most commonly used in the literature; what makes them attractive is their capability of assigning quality values to an approximation set independent of other sets under consideration. They have limitations, though, and there are differences in the power of existing indicators as will be shown in the following.

A. Limitations

Naturally, many studies have attempted to capture the multiobjective nature of approximation sets by deriving distinct indicators for the distance to the Pareto-optimal set and the diversity within the approximation set. Therefore, the question arises whether in principle there exists such a combination of, e.g., two indicators—one for distance, one for diversity—such that we can detect *whether* an approximation set is better than another. Such a combination of indicators, applicable to any type of problem, would be ideal because then any approximation set could be characterized by two real numbers that reflect the different aspects of the overall quality. The variety among the indicators proposed suggests that this goal is, at least, difficult to achieve. The following theorem shows that, in general, it cannot be achieved.

Theorem 1: Suppose an optimization problem with $n \geq 2$ objectives, where the objective space is $Z = \mathbb{R}^n$. Then, there exists no comparison method $C_{\boldsymbol{I}, E}$ based on a finite combination $\boldsymbol{I}$ of unary quality indicators that is $\triangleright$-compatible and $\triangleright$-complete at the same time, i.e,

$$C_{\boldsymbol{I}, E}(A, B) \Leftrightarrow A \triangleright B$$

for any approximation sets $A, B \in \Omega$.

That is, for any combination $\boldsymbol{I}$ of a finite number of unary quality indicators, we cannot find an interpretation function E such that the corresponding comparison method is $\triangleright$-compatible *and* $\triangleright$-complete. Or in other words: the number of criteria that determine what a good approximation set is is infinite.

We only sketch the proof here, the details can be found in the Appendix. First, we need the following fundamental results from set theory [18]:

- $\mathbb{R}$, $\mathbb{R}^k$, and any open interval (a, b) in $\mathbb{R}$ resp. hypercube $(a, b)^k$ in $\mathbb{R}^k$ have the same cardinality, denoted as $2^{\aleph_0}$, i.e., there is a bijection from any of these sets to any other;
- if a set S has cardinality $2^{\aleph_0}$, then the cardinality of the power set $\mathcal{P}(S)$ of S is $2^{2^{\aleph_0}}$, i.e., there is no injection from $\mathcal{P}(S)$ to any set of cardinality $2^{\aleph_0}$.

As we consider the most general case where $Z = \mathbb{R}^n$, we can construct a set S (cf. Fig. 8) such that any two points contained are incomparable to each other. Accordingly, any subset A of S is an approximation set and the power set of S, the cardinality of which is $2^{2^{\aleph_0}}$, is exactly the set of all approximation sets $A \subseteq S$. We will then show that any two approximation sets $A, B \subseteq S$

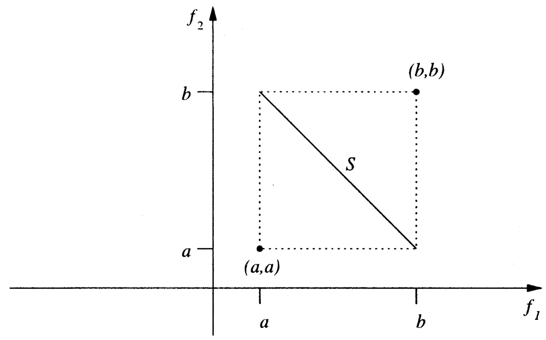


Fig. 8. Illustration of the construction used in Theorem 1 for a 2-D minimization problem. We consider an open rectangle $(a, b)^2$ and define an open line S within. For S holds that any two objective vectors contained are incomparable to each other, and therefore any subset $A \subseteq S$ is an approximation set.

with $A \neq B$ must differ in at least one of the k indicator values. Therefore, an injection from a set of cardinality $2^{2^{\aleph_0}}$ to $\mathbb{R}^k$ is required, which finally leads to a contradiction.

Note that Theorem 1 also holds: 1) if we only assume that Z contains an open hypercube in $\mathbb{R}^n$ for which $C_{\boldsymbol{I}, E}$ has the desired property and 2) if we consider any other relation from Table I (for $\parallel$ and $\succeq$ it follows directly from Theorem 1, for $\succ$ and $\succ\succ$ the proof has to be slightly modified).

Given this result, one may ask under which conditions the construction of such a comparison method is possible. For instance, such a comparison method exists if we allow an infinite number of indicators. The empirical attainment function [9], when applied to single approximation sets, can be understood as a combination of $|Z|$ unary indicators, where $|Z|$ denotes the cardinality of Z. If $Z = \mathbb{R}^n$, then this combination comprises an infinite number of unary indicators. On its basis, a $\triangleright$-compatible and $\triangleright$-complete comparison method can be constructed.

The situation also changes, if we require that each approximation set contains at maximum l objective vectors.

Corollary 1: Let $Z = \mathbb{R}^n$. It exists a unary indicator I and an interpretation function E such that

$$C_{I, E}(A, B) \Leftrightarrow A \triangleright B$$

for any $A, B \in \Omega$, with $|A|, |B| \leq l$.

Proof: Without loss of generality, we restrict ourselves to $Z = (0, 1)^n$ in the proof. The indicator I is constructed as follows:

$$I(A) = 0.d_1^1 d_2^1 \cdots d_l^1 d_1^2 d_2^2 \cdots d_l^2 d_1^3 \cdots$$

where d_j^i denotes the ith digit after the decimal point of the jth element in A. If A contains less than l elements, the first element is duplicated as many times as necessary. Accordingly, there is an injective function R that maps each real number in (0, 1) to an approximation set. If we define E as $E := (R(I(A)) \triangleright R(I(B)))$, the corresponding comparison method $C_{I, E}$ has the desired properties. $\square$

The corollary, however, is rather of theoretical than of practical use. The indicator constructed in the proof is able to indicate *whether* A is better than B, but it does not express how much better it is—this is one of the motives for using quality indicators. What we actually want is to apply a metric to the indicator values. Therefore, a reasonable requirement for a useful

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

combination of indicators may be that if A is better than or equal to B, then A is at least as good as B with respect to all k indicators, i.e.,

$$A \succeq B \Rightarrow (\forall\, 1 \leq i \leq k\colon I_i(A) \geq I_i(B)).$$

That this condition holds is an implicit assumption made in many studies. If we now restrict the size of the approximation sets to l and assume an indicator combination with the above property, can we then detect *whether* A is better than B? To answer this question, we will investigate a slightly reformulated statement, namely

$$A \succeq B \Leftrightarrow (\forall\, 1 \leq i \leq k\colon I_i(A) \geq I_i(B)$$

as this is equivalent to

$$A \rhd B \Leftrightarrow (\forall\, 1 \leq i \leq k\colon I_i(A) \geq I_i(B)) \\ \wedge (\exists\, 1 \leq j \leq k\colon I_j(A) > I_j(B)).$$

Furthermore, we will only consider the simplest case where $l = 1$, i.e., each approximation set consists of a single objective vector.

Theorem 2: Suppose an optimization problem with $n \geq 2$ objectives where the objective space is $Z = \mathbb{R}^n$. Let $\boldsymbol{I} = (I_1, I_2, \ldots, I_k)$ be a combination of k unary quality indicators and $E := (\forall\, 1 \leq i \leq k\colon I_i(\{\boldsymbol{z}^1\}) \geq I_i(\{\boldsymbol{z}^2\}))$ an interpretation function such that

$$C_{\boldsymbol{I},E}(\{\boldsymbol{z}^1\}, \{\boldsymbol{z}^2\}) \Leftrightarrow \boldsymbol{z}^1 \succeq \boldsymbol{z}^2$$

for any pair of objective vectors $\boldsymbol{z}^1, \boldsymbol{z}^2 \in Z$. Then, the number of indicators is greater than or equal to the number of objectives, i.e., $k \geq n$.

Proof: See the Appendix.

This theorem is a formalization of what is intuitively clear: we cannot reduce the dimensionality of the objective space without losing information. We need at least as many indicators as objectives to be able to detect *whether* an objective vector weakly dominates or dominates another objective vector. As a consequence, a fixed number of unary indicators is not sufficient for problems of arbitrary dimensionality even if we consider sets containing a single objective vector only.

In summary, we can state that the power of unary quality indicators is restricted. Theorem 1 proves that there does not exist any comparison method based on unary indicators that is $\rhd$-compatible and $\rhd$-complete at the same time. This rules out also other combinations, Table III shows which. It reveals that the best we can achieve is either $\succ\succ$-compatibility without any completeness, or $\not\rhd$-compatibility in combination with $\rhd$-completeness. That means we either can make strong statements ("A strongly dominates B") for only a few pairs $A \rhd B$; or we can make weaker statements ("A is not worse than B," i.e., $A \succeq B$ or $A \parallel B$) for all pairs $A \rhd B$.

TABLE III
OVERVIEW OF POSSIBLE COMPATIBILITY/COMPLETENESS COMBINATIONS WITH UNARY QUALITY INDICATORS. A MINUS MEANS THERE IS NO COMPARISON METHOD $C_{\boldsymbol{I},E}$ THAT IS COMPATIBLE REGARDING THE ROW-RELATION AND COMPLETE REGARDING THE COLUMN-RELATION. A PLUS INDICATES THAT SUCH A COMPARISON METHOD IS KNOWN, WHILE A QUESTION MARK STANDS FOR A COMBINATION FOR WHICH IT IS UNCLEAR WHETHER A CORRESPONDING COMPARISON METHOD EXISTS

compatibility	completeness						
	none	$\succ\succ$	$\succ$	$\rhd$	$\not\succ\not\succ$	$\not\succ$	$\not\rhd$
$\succ\succ$	+	-	-	-	-	-	-
$\succ$	+	?	-	-	-	-	-
$\rhd$	+	?	?	-	-	-	-
$\not\succ\not\succ$	+	+	+	+	-	?	?
$\not\succ$	+	+	+	+	-	-	?
$\not\rhd$	+	+	+	+	-	-	-

B. Classification

We now will review existing unary quality indicators according to the inferential power of the comparison methods that can be constructed on their basis: $\rhd$-compatible, $\not\rhd$-compatible, and not compatible with any relation listed in Table III. Table IV provides an overview of the various indicators discussed here. In this context, we would also like to point out the relationships between the dominance relations, e.g., $\succ\succ$-compatibility implies $\rhd$-compatibility, $\not\rhd$-compatibility implies $\not\succ$-compatibility, and $\rhd$-completeness implies $\succ\succ$-completeness. Moreover, note that in the following, we will not consider the case of identical approximation sets $A = B$ as an equality check can be easily incorporated into any comparison method. Therefore, in Table IV, the relations $\succeq$ and $\not\succeq$ are not contained.

1) $\rhd$-Compatibility: The use of $\rhd$-compatible comparison methods based on unary indicators is restricted according to Theorem 2: in order to detect dominance between objective vectors, at least as many indicators as objectives are required. Hence, it is not surprising that, to the best of our knowledge, no $\rhd$-compatible comparison methods have been proposed in the literature. Their design, though, is possible:[3]

- Suppose a minimization problem and let

$$I_1^{HC}(A) = \sup_{a \in \mathbb{R}} \{\{(a, a, \ldots, a)\} \rhd A\}$$
$$I_2^{HC}(A) = \inf_{b \in \mathbb{R}} \{\{(b, b, \ldots, b)\} \lhd A\}.$$

  We assume that Z is bounded, i.e., $I_1^{HC}(A)$ and $I_2^{HC}(A)$ always exist. As illustrated in Fig. 9, the two indicator values characterize a hypercube that contains all objective vectors in A. If we define the indicator $\boldsymbol{I}_{HC} = (I_1^{HC}, I_2^{HC})$ and the interpretation function E as $E := (I_2^{HC}(A) < I_1^{HC}(B))$, then the comparison method $C_{\boldsymbol{I}_{HC},E}$ is $\rhd$-compatible.
- Suppose a minimization problem and let

$$I_i^O(A) = \inf_{a \in \mathbb{R}} \{\forall (z_1, \ldots, z_n) \in A\colon z_i \leq a\}$$

  for $1 \leq i \leq n$ and

$$I_{n+1}^O(A) = \begin{cases} 0, & \text{if } A \text{ contains two or more elements} \\ 1, & \text{else.} \end{cases}$$

  The idea behind these indicators is similar to the above example. We consider the smallest hyperrectangle that entirely encloses A. This hyperrectangle comprises exactly one point O that is weakly dominated by all members in A; in the case of a two-dimensional (2-D) minimization problem, it is the upper right corner of the enclosing rectangle (cf. Fig. 9). We see that $I_1^O, \ldots, I_n^O$ are the coordinates of this point O. I_{n+1}^O serves to distinguish between

[3]Note that there exists a trivial case: the comparison method that always yields false.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
OVERVIEW OF UNARY INDICATORS. EACH ENTRY CORRESPONDS TO A SPECIFIC COMPARISON METHOD DEFINED BY THE INDICATOR AND THE INTERPRETATION FUNCTION IN THAT ROW. WITH RESPECT TO COMPATIBILITY AND COMPLETENESS, NOT ALL RELATIONS ARE LISTED BUT ONLY THE STRONGEST AS, E.G., $\succ\succ$-COMPATIBILITY IMPLIES $\triangleright$-COMPATIBILITY (CF. SECTION III-B)

indicator	name / reference	Boolean function	compatibility	completeness
$\boldsymbol{I}_{HC}$	enclosing hypercube indicator / Section III-B.1	$I_2^{HC}(A) < I_1^{HC}(B)$	$\succ\succ$	-
$\boldsymbol{I}_{O}$	objective vector indicator / Section III-B.1	$I_i^{O}(A) < I_i^{O}(B)$	$\succ\succ$	-
I_H	hypervolume indicator / [7]	$I_H(A) > I_H(B)$	$\ntriangleright$	$\triangleright$
I_W	average best weight combination / [19]	$I_W(A) < I_W(B)$	$\ntriangleright$	$\succ\succ$
I_D	distance from reference set / [20]	$I_D(A) < I_D(B)$	$\ntriangleright$	$\succ\succ$
$I_{\epsilon 1}$	unary ϵ-indicator / Section III-B.2	$I_{\epsilon 1}(A) < I_{\epsilon 1}(B)$	$\ntriangleright$	$\succ\succ$
I_{PF}	fraction of Pareto-optimal front covered / [22]	$I_{PF}(A) > I_{PF}(B)$	$\ntriangleright$	-
I_P	number of Pareto points contained / Section III-B.2	$I_P(A) > I_P(B)$	$\ntriangleright$	-
I_{ER}	error ratio / [13]	$I_{ER}(A) > 0$	$\not\succ$	-
I_{CD}	chi-square-like deviation indicator / [14]	$I_{CD}(A) < I_{CD}(B)$	-	-
I_S	spacing / [23]	$I_S(A) < I_S(B)$	-	-
I_{ONVG}	overall nondominated vector generation / [13]	$I_{ONVG}(A) > I_{ONVG}(B)$	-	-
I_{GD}	generational distance / [13]	$I_{GD}(A) < I_{GD}(B)$	-	-
I_{ME}	maximum Pareto front error / [13]	$I_{ME}(A) < I_{ME}(B)$	-	-
I_{MS}	maximum spread / [21]	$I_{MS}(A) > I_{MS}(B)$	-	-
I_{MD}	minimum distance between two solutions / [24]	$I_{MD}(A) > I_{MD}(B)$	-	-
I_{CE}	coverage error / [24]	$I_{CE}(A) < I_{CE}(B)$	-	-
I_{DU}	deviation from uniform distribution / [25]	$I_{DU}(A) < I_{DU}(B)$	-	-
I_{OS}	Pareto spread / [26]	$I_{OS}(A) > I_{OS}(B)$	-	-
I_A	accuracy / [26]	$I_A(A) > I_A(B)$	-	-
I_{NDC}	number of distinct choices / [26]	$I_{NDC}(A) > I_{NDC}(B)$	-	-
I_{CL}	cluster / [26]	$I_{CL}(A) < I_{CL}(B)$	-	-

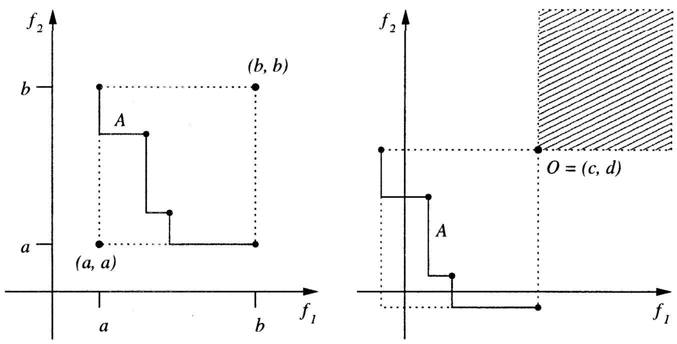


Fig. 9. Two indicators capable of indicating *that* $A \triangleright B$ for some $A, B \in \Omega$. On the left hand side, it is depicted how the $\boldsymbol{I}_{HC}$ indicator defines a hypercube around an approximation set A, where $I_1^{HC}(A) = a$ and $I_2^{HC}(A) = b$. The right picture is related to the $\boldsymbol{I}_O$ indicator: for any objective vector in the shaded area we can detect that it is dominated by the approximation set A. Here, $I_1^O(A) = c$, $I_2^O(A) = d$, and $I_3^O(A) = 0$.

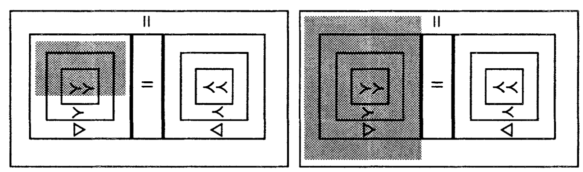


Fig. 10. The shaded area stands for those ordered pairs (A, B) for which a comparison method $C_{I,E}$ yields true. The left-hand side shows a possible pattern, if the comparison method is $\triangleright$-compatible, but not $\triangleright$-complete. The right hand side represents the hypervolume comparison method $C_{I_H,E}$ with $E := (I_H(A) > I_H(B))$ that is both $\not\succeq$-compatible and $\triangleright$-complete.

single objective vectors and larger approximation sets. Let $\boldsymbol{I}_O = (I_1^O, \ldots, I_{n+1}^O)$ and define the interpretation function E as $E := (\forall 1 \leq i \leq n+1\colon I_i^O(A) < I_i^O(B))$. Then, the comparison method $C_{\boldsymbol{I}_O,E}$ is $\triangleright$-compatible; it detects dominance between an approximation set and those objective vectors that are dominated by all members of this approximation set.

Note that both comparison methods are even $\succ\succ$-compatible, but neither is complete with regard to any dominance relation. This property is visualized in Fig. 10.

Moreover, some unary indicators can also be used to design a $\triangleright$-compatible comparison method if the Pareto-optimal set P is known. Consider, e.g., the following unary ϵ-indicator $I_{\epsilon 1}$ that is based on the binary ϵ-indicator from Definition 5:

$$I_{\epsilon 1}(A) = I_\epsilon(A, P).$$

Obviously, $I_{\epsilon 1}(A) = 1$ implies $A = P$. Thus, in combination with the interpretation function $E := (I_{\epsilon 1}(A) = 1 \wedge I_{\epsilon 1}(B) > 1)$ a comparison method can be defined that is $\triangleright$-compatible and detects *that* A is better than B for all pairs $A, B \in \Omega$ with $A = P$ and $B \neq P$. The same construction can be made for some other indicators, e.g., the hypervolume indicator, as well. Nevertheless, these comparison methods are only applicable if some of the algorithms under consideration can actually generate the Pareto-optimal set.

2) $\ntriangleright$-Compatibility: Consider the above unary ϵ-indicator $I_{\epsilon 1}$. For any pair $A, B \in \Omega$, it holds

$$A \succ\succ B \Rightarrow I_{\epsilon 1}(A) < I_{\epsilon 1}(B)$$
$$A \triangleright B \Rightarrow I_{\epsilon 1}(A) \leq I_{\epsilon 1}(B)$$

and (which follows from the latter)

$$I_{\epsilon 1}(A) < I_{\epsilon 1}(B) \Rightarrow A \not\prec\not\prec B \Rightarrow A \ntriangleleft B.$$

Therefore, the comparison method $C_{I_{\epsilon 1},E}$ with $E := (I_{\epsilon 1}(A) < I_{\epsilon 1}(B))$ is $\ntriangleright$-compatible and $\succ\succ$-complete, but neither $\triangleright$- nor $\succ$-complete. That is whenever $A \succ\succ B$, we

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

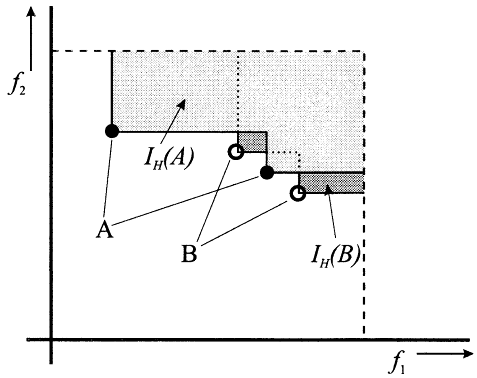


Fig. 11. Example of two incomparable sets A and B. The hypervolume comparison method that is $\not\succeq$-compatible and $\triangleright$-complete yields true as $I_H(A) > I_H(B)$.

will be able to state that A is not worse than B. On the other hand, there are cases $A \succ B$ for which this conclusion cannot be drawn, although A is actually not worse than B. The same holds for the two indicators proposed by [19] and [20]. We will not discuss these in detail and only remark that the following example can be used to show that both indicators in combination with the interpretation function $E := (I(A) < I(B))$ are not $\succ$-complete (and $\triangleright$-complete): the Pareto-optimal set is $P = \{(1, 1)\}$, and $A = \{(4, 2)\}$ and $B = \{(4, 3)\}$.

The hypervolume indicator I_H [7], [21] is the only unary indicator we are aware of that is capable of detecting that A is not worse than B for all pairs $A \triangleright B$. It gives the hypervolume of that portion of the objective space that is weakly dominated by an approximation set A.[4] We notice that from $A \triangleright B$ follows that $I_H(A) > I_H(B)$; the reason is that A must contain at least one objective vector that is not weakly dominated by B, thus, a certain portion of the objective space is dominated by A but not by B. This observation implies both $\not\triangleright$-compatibility and $\triangleright$-completeness. If we also consider the case $A = B$, this method is even $\not\succeq$-compatible (cf. Fig. 10). However, if A is incomparable to B, then $C_{I_H, E}$ may also yield true as shown in Fig. 11. Nevertheless, according to Theorem 1 and Table III $\not\triangleright$-compatibility is the best we can achieve for a $\triangleright$-complete comparison method based on unary quality indicators.

Van Veldhuizen [13] suggested an indicator, the error ratio I_{ER}, on the basis of which a $\not\succ$-compatible (but not $\not\triangleright$-compatible) comparison method can be defined. $I_{ER}(A)$ gives the ratio of Pareto-optimal objective vectors to all objective vectors in the approximation set A. Obviously, if $I_{ER}(A) > 0$, i.e., A contains at least one Pareto-optimal objective vector, then there exists no $B \in \Omega$ with $B \succ A$. On the other hand, if A consist of only a single Pareto-optimal objective vector, then $I_{ER}(A) \geq I_{ER}(B)$ for all $B \triangleright A$; if B contains not only Pareto-optimal objective vectors, then $I_{ER}(A) > I_{ER}(B)$. Therefore, $C_{(I_{ER}, E)}$ with $E := (I_{ER}(A) > I_{ER}(B))$ is not $\not\triangleright$-compatible. However, if we consider just the total number (rather than the ratio) of Pareto-optimal objective vectors in the approximation set, we obtain $\not\triangleright$-compatibility. This also holds for the indicator used in [22], which gives the ratio of the number of Pareto-optimal objective vectors in A to the number of all Pareto-optimal objective vectors. Nevertheless, the power of these comparison methods is limited because none of them is complete with respect to any dominance relation.

3) Incompatibility: Section III-A has revealed the difficulties when trying to separate the overall quality of approximation sets into distinct aspects. Nevertheless, it would be desirable if we could look at certain criteria such as diversity separately, and accordingly several authors suggested formalizations of specific aspects by means of unary indicators. However, we have to be aware that often these indicators do *in general* neither indicate *that* $A \triangleright B$ nor $A \not\triangleright B$.

One class of indicators that does not allow any conclusions to be drawn regarding the dominance relationship between approximation sets is represented by the various diversity indicators [14], [21], [23]–[26]. If we consider a pair $A, B \in \Omega$ with $A \triangleright B$, the indicator value of A can, in general, be less or greater than or even equal to the value assigned to B (for the diversity indicators referenced above). Therefore, the comparison methods based on these indicators are neither compatible nor complete with respect to any dominance relation or complement of it.

The same holds for the three indicators proposed in [13]: overall nondominated vector generation I_{ONVG}, generational distance I_{GD}, and maximum Pareto front error I_{ME}. The first just gives the number of elements in the approximation set, and it is obvious that it does not provide sufficient information to conclude $A \triangleright B$, $A \not\triangleright B$, etc. Why this also applies to the other two, both distance indicators, will only be sketched here. Assume a 2-D minimization problem for which the Pareto-optimal set P consists of the two objective vectors (1, 0) and (0, 10). Now, consider the three sets $A = \{(2, 5)\}$, $B = \{(3, 9)\}$, and $C = \{(10, 10)\}$. For both distance indicators holds $I(B) < I(A) < I(C)$, but $A \succ\succ B \succ\succ C$, provided that Euclidean distance is considered. Thus, we cannot conclude whether one set is better or worse than another by just looking at the order of the indicator values. A similar argument as for the generational distance applies to the coverage error indicator presented in [24]; the only difference is that the coverage error denotes the minimum distance to the Pareto-optimal set instead of the average distance.

Knowles and Corne [11], [12] have discussed the incompatibility of the I_S, I_{ME}, I_{DU}, and I_{ONVG} indicators, though from a different perspective, in more depth, and the interested reader is referred to [11] for a more detailed discussion of this topic.

Finally, one can ask whether it is possible to combine several indicators for which no $\not\triangleright$-compatible comparison method exists in such a way that the resulting indicator vector allows to detect that A is not worse than B. Van Veldhuizen and Lamont [2], for instance, used generational distance and overall nondominated vector generation in conjunction with the diversity indicator of [23], while Deb *et al.* [6] applied a similar combination of diversity and distance indicators. Other examples can be found in, e.g., [27] and [24]. As in all of these cases, counterexamples can be constructed that show the corresponding comparison methods to be not $\not\triangleright$-compatible, the above question remains open and is not investigated in more depth here.

[4]Note that Z has to be bounded, i.e., there must exist a hypercube in $\mathbb{R}^n$ that encloses Z. If this requirement is not fulfilled, it can be easily achieved by an appropriate transformation.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

IV. Comparison Methods Based on Binary Quality Indicators

Binary quality indicators can be used to overcome the difficulties with unary indicators. However, they also have a drawback: when we compare t algorithms using a single binary indicator, we obtain $t(t-1)$ distinct indicator values—in contrast to the t values in the case of a unary indicator. This renders the analysis and the presentation of the results more difficult. Nevertheless, Theorem 1 suggests that this is in the nature of multiobjective optimization problems.

A. Limitations

In principle, there are no such theoretical limitations of binary indicators as for unary indicators. For instance, the indicator

$$I(A, B) = \begin{cases} 4, & A \succ\succ B \\ 3, & A \succ B \\ 2, & A \rhd B \\ 1, & A = B \\ 0, & \text{else} \end{cases}$$

allows to construct compatible and complete comparison methods with regard to any of the dominance relations. However, this usually does not hold for existing practically useful binary indicators, in particular those that are, as Knowles and Corne [12] denote it, *symmetric,* i.e., $I(A, B) = c - I(B, A)$ for a constant c. Although, symmetric indicators are attractive as only half the number of indicator values has to be considered in comparison to a general binary indicator, their inferential power is restricted as we will show in the following.

Without loss of generality, suppose that $c = 0$, i.e., $I(A, B) = -I(B, A)$; otherwise consider the transformation $I'(A, B) = c/2 - I(A, B)$. The question is whether we can construct a $\rhd$-compatible and $\rhd$-complete comparison method based on this indicator; according to the discussion in Section III-A, we assume that $E =: (I(A, B) > I(B, A))$.

Theorem 3: Let I be a binary indicator with $I(A, B) = -I(B, A)$ for $A, B \in \Omega$ and E an interpretation function with $E =: I(A, B) > I(B, A)$. If the corresponding comparison method $C_{I,E}$ is $\rhd$-compatible and $\rhd$-complete, then $I(A, B) = 0$ for all $A, B \in \Omega$ with $A = B$ or $A \parallel B$.

Proof: Let $A, B \in \Omega$. From $A \rhd B \Leftrightarrow I(A, B) > I(B, A)$, it follows that $A \not\rhd B \Leftrightarrow I(A, B) \leq I(B, A)$, and therefore $A \parallel B \vee A = B \Leftrightarrow A \not\rhd B \wedge B \not\rhd A \Leftrightarrow I(A, B) = I(B, A)$. From the symmetry $I(A, B) = -I(B, A)$, it then follows that $A \parallel B \vee A = B$ is equivalent to $I(A, B) = 0$. □

A consequence of this theorem is that a symmetric binary indicator, for which $A \rhd B \Leftrightarrow I(A, B) > I(B, A)$, can detect whether A is better than B, but not whether $A \succeq B$, $A \parallel B$, or $A = B$. On the other hand, it follows from $I(A, B) \neq 0$ for a pair $A \parallel B$ that $C_{I,E}$ cannot be $\rhd$-compatible, if it is $\rhd$-complete. We will use this result in the following discussion of existing binary indicators.

B. Classification

In contrast to unary indicators, only a few binary indicators can be found in the literature. We will classify them according to the criterion of whether a corresponding comparison method exists that is $\blacktriangleright$-compatible *and* $\blacktriangleright$-complete with regard to a specific relation $\blacktriangleright$.

As mentioned in Section II-B, Zitzler and Thiele [7] suggested the coverage indicator I_C, where $I_C(A, B)$ gives the fraction of solutions in B that are weakly dominated by at least one solution in A. $I_C(A, B) = 1$ is equivalent to $A \succeq B$ (A weakly dominates B) and therefore comparison methods $C_{I_C, E}$ compatible and complete with regard to the $\rhd$, $\succeq$, $\parallel$, and $=$ relations can be constructed. Furthermore, with $E := (I_C(A, B) = 1 \wedge I_C(B, A) = 0)$, we obtain a comparison method $C_{I_C, E}$ that is $\succ$-compatible and $\succ$-complete.

Hansen and Jaszkiewicz [8] proposed three symmetric binary indicators I_{R_1}, I_{R_2}, and I_{R_3} that are based on a set of utility functions. The utility functions can be used to formalize and incorporate preference information; however, if no additional knowledge is available, Hansen and Jaszkiewicz suggest using a set of weighted Tchebycheff utility functions. In this case, the resulting comparison methods are, in general, $\rhd$-complete but not $\rhd$-compatible as Theorem 3 applies ($I(A, B)$ can be greater or less than 0 if $A \parallel B$). Accordingly, these indicators, in general, do not allow construction of a comparison method that is both compatible and complete with respect to any of the relations in Table I.

In [21], a binary version I_{H2} of the hypervolume indicator I_H [7] was proposed; the same indicator was used in [10]. $I_{H2}(A, B)$ is defined as the hypervolume of the subspace that is weakly dominated by A but not by B. From $I_{H2}(A, B) = 0$, it follows that $B \succeq A$, and therefore, as with the coverage indicator, comparison methods $C_{I_H, E}$ compatible and complete regarding the $\rhd$, $\succeq$, $\parallel$, and $=$ relations are possible. However, there exists no $\succ$-compatible and $\succ$-complete or $\succ\succ$-compatible and $\succ\succ$-complete comparison method solely based on the binary hypervolume indicator.

Knowles and Corne [16] presented a comparison method based on the study by Fonseca and Fleming [15]. Although designed for the statistical analysis of multiple optimization runs, the method can be formulated in terms of an m-ary indicator I_{LI} if only one run is performed per algorithm or the algorithms are deterministic. Here, we restrict ourselves to the case $m = 2$ as all of the following statements also hold for $m > 2$. A user-defined set of lines in the objective space, all of them passing the origin and none of them perpendicular to any of the axes, forms the scaffolding of Knowles and Corne's approach. First, for each line, the intersections with the attainment surfaces [15] defined by the approximation sets under consideration are calculated. The intersections are then sorted according to their distance to the origin, and the resulting order defines a ranking of the approximation sets with respect to this line. If only two approximation sets are considered, then $I_{LI}(A, B)$ gives the fraction of the lines for which A is ranked higher than B. Accordingly, the most significant outcome would be $I_{LI}(A, B) = 1$ and $I_{LI}(B, A) = 0$. However, this method strongly depends on the choice of the lines, and certain parts of the attainment surface are not sampled. Therefore, in the above case either A is better than B or both approximation are incomparable to each other. As a consequence, the comparison method $C_{I_{LI}, E}$ with $E := (I_{LI}(A, B) = 1 \wedge I_{LI}(B, A) = 0)$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

TABLE V

OVERVIEW OF BINARY INDICATORS. A MINUS MEANS THAT IN GENERAL THERE IS NO COMPARISON METHOD $C_{I,E}$ BASED ON THE INDICATOR I IN THE CORRESPONDING ROW THAT IS COMPATIBLE AND COMPLETE REGARDING THE RELATION IN THE CORRESPONDING COLUMN. OTHERWISE, AN EXPRESSION IS GIVEN THAT DESCRIBES AN APPROPRIATE INTERPRETATION FUNCTION E

ind.	name / reference	compatible and complete with respect to relation:					
		$\succ\succ$	$\succ$	$\triangleright$	$\succeq$	$=$	$\parallel$
I_ϵ	epsilon indicator / Section II-B	$I_\epsilon(A,B)<1$	-	$I_\epsilon(A,B)\leq 1$ $I_\epsilon(B,A)>1$	$I_\epsilon(A,B)\leq 1$	$I_\epsilon(A,B)=1$ $I_\epsilon(B,A)=1$	$I_\epsilon(A,B)>1$ $I_\epsilon(B,A)>1$
$I_{\epsilon+}$	additive epsilon indicator / Section II-B	$I_{\epsilon+}(A,B)<0$	-	$I_{\epsilon+}(A,B)\leq 0$ $I_{\epsilon+}(B,A)>0$	$I_{\epsilon+}(A,B)\leq 0$	$I_{\epsilon+}(A,B)=0$ $I_{\epsilon+}(B,A)=0$	$I_{\epsilon+}(A,B)>0$ $I_{\epsilon+}(B,A)>0$
I_C	coverage / [7]	-	$I_C(A,B)=1$ $I_C(B,A)=0$	$I_C(A,B)=1$ $I_C(B,A)<1$	$I_C(A,B)=1$	$I_C(A,B)=1$ $I_C(B,A)=1$	$0<I_C(A,B)<1$ $0<I_C(B,A)<1$
I_{H2}	binary hypervolume indicator / [21]	-	-	$I_{H2}(A,B)>0$ $I_{H2}(B,A)=0$	$I_{H2}(A,B)\geq 0$ $I_{H2}(B,A)=0$	$I_{H2}(A,B)=0$ $I_{H2}(B,A)=0$	$I_{H2}(A,B)>0$ $I_{H2}(B,A)>0$
I_{R1}	utility function indicator $R1$ / [8]	-	-	-	-	-	-
I_{R2}	utility function indicator $R2$ / [8]	-	-	-	-	-	-
I_{R3}	utility function indicator $R3$ / [8]	-	-	-	-	-	-
I_{LI}	lines of intersection / [16]	-	-	-	-	-	-

is, in the general case, not $\triangleright$-compatible; however, it is $\not\succ$-compatible and $\succ\succ$-complete.

Finally, we have already shown in Section II-C that a $\triangleright$-compatible and $\triangleright$-complete comparison method exists for the ϵ-indicator. The case $I_\epsilon(A,B)\leq 1$ is equivalent to $A\succeq B$ and the same statements as for the coverage and the binary hypervolume indicators hold. Furthermore, the comparison method $C_{I_\epsilon,E}$ with $E := (I_\epsilon(A,B) < 1)$ is $\succ\succ$-compatible and $\succ\succ$-complete.

Table V summarizes the results of this section. Note that it only contains information about comparison methods that are both compatible and complete with respect to the different dominance relations.

V. DISCUSSION

A. Summary of Results

We have proposed a mathematical framework to study quality assessment methods for multiobjective optimizers. Starting with the assumption that the outcome of a multiobjective EA can be represented by a set of incomparable objective vectors, a so-called approximation set, we have introduced several dominance relations on approximation sets. These relations represent a formal description of what we intuitively understand by one approximation set being better than another. The term *quality indicator* has been used to capture the notion of a quality measure, and a comparison method has been defined as a combination of quality indicators and an interpretation function that evaluates the indicator values. Furthermore, we have discussed two properties of comparison methods, namely compatibility and completeness, which characterize the relationship between comparison methods and dominance relations. On the basis of this framework, existing comparison methods have been analyzed and discussed. The key results are as follows.

- Unary quality indicators, i.e., quality measures that summarize an approximation set in terms of a real number, are *in general* not capable of indicating *whether* an approximation set is better than another—also if several of them are used. This even holds if we consider approximation sets containing a single objective vector only.
- Existing unary indicators at best allow to infer *that* an approximation set is not worse than another, e.g., the distance indicator by Czyzak and Jaszkiewicz [20], the hypervolume indicator by Zitzler and Thiele [7], or the unary ϵ-indicator presented in this paper. However, with many unary indicators and also combinations of unary indicators, no statement about the relation between the corresponding approximation sets can be made. That is, although an approximation set A may be evaluated better than an approximation set B with respect to all of the indicators, B can actually be superior to A with respect to the dominance relations. This holds especially for the various diversity measures and also for some of the distance indicators proposed in the literature.
- We have given two examples demonstrating that comparison methods based on unary indicators can be constructed, such that A can be recognized as being better than B for some approximation sets A, B. It has also been shown that the practical use of this type of indicator is naturally restricted.
- Binary indicators, which assign real numbers to ordered pairs of approximation sets, in principle do not possess the theoretical limitations of unary indicators. The binary ϵ-indicator proposed in this paper, e.g., is capable of detecting *whether* an approximation set is better than another. However, not all existing binary indicators have this property. Furthermore, it has to be mentioned that the greater inferential power comes along with additional complexity: in contrast to unary indicators, the number of indicator values to be considered is not linear but quadratic in the number of approximation sets.

B. Conclusions

The results of this paper have been obtained analytically, and naturally, the question arises: how they translate into practice,

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

i.e., what consequences do the theoretical considerations have for the researcher who is carrying out a comparative study?

The choice of the quality indicator(s) strongly depends on the type of statements we would like to make. Are we interested in general statements that hold for many scenarios and are based on as few assumptions as possible regarding the decision maker's preferences? Or do we consider a specific scenario where the preferences of the decision maker are (partially) known, e.g., in terms of a ranking of the objectives? Often, it is even desirable to be able to draw both general and specific conclusions.

The most general statements possible are of the form "approximation set A strictly dominates / dominates / is better than / weakly dominates approximation set B," which are only based on the concept of Pareto dominance (cf. Table I). In this case, the comparison method should be at least capable of detecting *whether* A is better than B, i.e., it should be compatible and complete with respect to as many of the dominance relations as possible. Since this cannot be achieved by using a (finite) combination of unary quality indicators, e.g., one for distance and another for diversity, appropriate binary quality indicators are needed. Among the ones discussed in this paper, the two binary ϵ-indicators (Section II-C) and the coverage indicator [7] are best suited as they provide compatibility and completeness to most of the dominance relations.

In practice, though, we would like to make additional, more precise statements—beyond the dominance relations. If A is better than B, the question is how much better it is. If A is incomparable to B, it would be interesting to know whether there are certain aspects in which A is better than B. To answer these questions, we ineluctably have to make assumptions about the decision maker's preferences, and the more assumptions are incorporated the more specific but also weaker the conclusions will be in the sense that they only hold if these preferences apply.

Binary quality indicators are usually designed to formalize broadly valid preferences, and in the ideal case an indicator allows to make statements about both dominance relations and preference-dependent performance differences. As to the second aspect, the coverage indicator is only of limited use because it does not say anything about how much A is better than B. In contrast, the binary ϵ-indicators and the binary hypervolume indicator [21] are able to detect performance differences, but the latter does not indicate *whether* A dominates or even strictly dominates B.

Unary quality indicators, e.g., the various diversity measures, often represent problem-dependent knowledge. Nevertheless, they can be useful to focus on specific aspects in order to design and improve algorithms. We will not discuss the issue of how to integrate preference information into the quality assessment in detail here and refer to the work of Hansen and Jaszkiewicz [8] instead. However, note that a comparison method, independent of the type of indicators used, should be at least $\not\triangleright$-compatible, i.e., there is no pair of approximation sets A and B where the comparison method says B is preferred to A, while actually A is better than B ($A \triangleright B$). If the comparison method is in addition $\triangleright$-complete, we know that whenever $A \triangleright B$ then A will also be evaluated better than B.

Finally, the above considerations concentrate on only one, but essential criterion: the inferential power. Certainly, there are many other aspects according to which comparison methods can be investigated, e.g., the computational effort, the sensitivity to scaling, the requirement to have knowledge about the Pareto-optimal set, etc. Several such aspects have been investigated in [11] and [12]. Also, in light of these aspects, the binary ϵ-indicator possesses several desirable features. It represents a natural extension to the evaluation of approximation schemes in theoretical computer science [28] and gives the factor by which an outcome is worse than another. Thus, it has a clear interpretation—in contrast to, e.g., the coverage and the hypervolume indicators. Furthermore, it is cheap to compute as opposed to the hypervolume indicator, which is computationally expensive [11].

In summary, for general statements, we recommend to use the binary ϵ-indicator (multiplicative or additive), according to the current status of knowledge. However, there may be particular scenarios where this indicator is not appropriate. In addition, more specific, usually problem-dependent indicators can be included in order to exploit knowledge about the decision maker's preferences and to draw more precise conclusions.

C. *Future Work*

An important issue, which has not been addressed in this paper, is the stochasticity of multiobjective EAs. Multiple optimization runs require the application of statistical tests, and in principle there are two ways to incorporate these tests in a comparison method: the statistical testing procedure can be included in the indicator functions or in the interpretation function. Knowles and Corne's approach [16] belongs to the first category, while Van Veldhuizen and Lamont's study [2] is an example for the second category. The attainment function method proposed by Grunert da Fonseca *et al.* [9] can be expressed in terms of an infinite number of indicators and therefore falls in the second category. However, in contrast to [2] and [16], this method is able to detect *whether* an approximation set is better than another. Investigating, in more depth, how all of these approaches are related to each other is the subject of ongoing research.

APPENDIX

Proof of Theorem 1: Let us suppose that such a comparison method $C_{\boldsymbol{I},E}$ exists where $\boldsymbol{I} = (I_1, I_2, \ldots, I_k)$ is a combination of k unary quality indicators and E a corresponding interpretation function $\mathbb{R}^{2k} \rightarrow \{false, true\}$. Furthermore, assume, without loss of generality, that the first two objectives are to be minimized (otherwise, the definition of the following set S has to be modified accordingly).

Choose $a, b \in \mathbb{R}$ with $a < b$, and consider $S = \{(z_1, z_2, \ldots, z_n) \in Z;\, a < z_i < b,\, 1 \leq i \leq n \wedge z_2 = b + a - z_1\}$. Obviously, for any $\boldsymbol{z}^1, \boldsymbol{z}^2 \in Z$, either $\boldsymbol{z}^1 = \boldsymbol{z}^2$ or $\boldsymbol{z}^1 \parallel \boldsymbol{z}^2$, because $z_1^1 > z_1^2$ implies $z_2^1 < z_2^2$. Furthermore, let $\Omega_S \subseteq \Omega$ denote the set of approximations sets $A \in \Omega$ with $A \subseteq S$.

As $S \in \Omega$ and any subset of an approximation set is again an approximation set, Ω_S is identical to the power set $\mathcal{P}(\mathcal{S})$ of S. In addition, there is an injection f from the open interval (a, b) to S with $f(r) = (r, b + a - r, (b + a)/2, (b + a)/2, \ldots, (b+$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

$a)/2)$, it follows that the cardinality of S is at least $2^{\aleph_0}$. As a consequence, the cardinality of Ω_S is at least $2^{2^{\aleph_0}}$.

Now, we will use Lemma 1 (see below): it shows that for any $A, B \in \Omega_S$ with $A \neq B$ the quality indicator values differ, i.e., $I_i(A) \neq I_i(B)$ for at least one indicator I_i, $1 \leq i \leq k$. Therefore, there must be an injection from Ω_S to $\mathbb{R}^k$, the codomain of $\boldsymbol{I}$. This means there is an injection from a set of cardinality $2^{2^{\aleph_0}}$ (or greater) to a set of cardinality $2^{\aleph_0}$. From this absurdity, it follows that such a comparison method $C_{\boldsymbol{I}, E}$ cannot exist. □

Lemma 1: Let $Z = \{(z_1, z_2, \ldots, z_n) \in \mathbb{R}^n;\ a < z_i < b,\ 1 \leq i \leq n\}$ be an open hypercube in $\mathbb{R}^n$ with $n \geq 2$, $a, b \in \mathbb{R}$, and $a < b$. Furthermore, assume there exists a combination of unary quality indicators $\boldsymbol{I} = (I_1, I_2, \ldots, I_k)$ and an interpretation function E, such that for any approximation sets $A, B \in \Omega$

$$C_{\boldsymbol{I}, E}(A, B) \Leftrightarrow A \rhd B.$$

Then, for all $A, B \in \Omega$ with $A \neq B$, there is at least one quality indicator I_i with $1 \leq i \leq k$, such that $I_i(A) \neq I_i(B)$.

Proof: Let $A, B \in \Omega$ be two arbitrary approximation sets with $A \neq B$. First, note that $C_{\boldsymbol{I}, E}(A, B)$ implies $C_{\boldsymbol{I}, E}(B, A)$ is false (and vice versa) as $A \rhd B$ implies $B \not\rhd A$. If $A \rhd B$ or $B \rhd A$, then $I_i(A) \neq I_i(B)$ for at least one $1 \leq i \leq k$, because otherwise $C_{\boldsymbol{I}, E}(A, B) = C_{\boldsymbol{I}, E}(B, A) = C_{\boldsymbol{I}, E}(A, A)$ would be false. If $A \parallel B$, there are two cases: 1) both A and B contain only a single objective vector or 2) either set consists of more than one element.

Case 1) Choose $\boldsymbol{z} \in Z$ with $A \parallel \{\boldsymbol{z}\}$ and $B \parallel \{\boldsymbol{z}\}$ (such an objective vector exists as Z is an open hypercube in $\mathbb{R}^n$). Then $A \cup \{\boldsymbol{z}\} \rhd A$ and $A \cup \{\boldsymbol{z}\} \parallel B$, and from the former follows $C_{\boldsymbol{I}, E}(A \cup \{\boldsymbol{z}\}, A)$ is true. Accordingly, $I_i(A) \neq I_i(B)$ for at least one $1 \leq i \leq k$, because otherwise $C_{\boldsymbol{I}, E}(A \cup \{\boldsymbol{z}\}, B) = C_{\boldsymbol{I}, E}(A \cup \{\boldsymbol{z}\}, A)$ would be true, which contradicts $A \cup \{\boldsymbol{z}\} \parallel B$.

Case 2) Assume, without loss of generality, that A contains more than one objective vector, and choose $\boldsymbol{z} \in A$ with $\{\boldsymbol{z}\} \parallel B$ (such an element must exist as $A \parallel B$). Then, $A \rhd \{\boldsymbol{z}\}$, which implies that $C_{\boldsymbol{I}, E}(A, \{\boldsymbol{z}\})$. Now suppose $I_i(A) = I_i(B)$ for all $1 \leq i \leq k$; it follows that $C_{\boldsymbol{I}, E}(B, \{\boldsymbol{z}\}) = C_{\boldsymbol{I}, E}(A, \{\boldsymbol{z}\})$ is true which is a contradiction to $B \parallel \{\boldsymbol{z}\}$.

In summary, all cases ($A \rhd B$, $B \rhd A$, and $A \parallel B$) imply that $I_i(A) \neq I_i(B)$ for at least one $1 \leq i \leq k$. □

Proof of Theorem 2: We will exploit the fact that in $\mathbb{R}$, the number of disjoint open intervals $(a, b) = \{z \in \mathbb{R};\ a < z < b\}$ with $a < b$ is countable [18]. In general, this means that $\mathbb{R}^k$ contains only countably many disjoint open hyperrectangles $(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_k, b_k) = \{(z_1, z_2, \ldots, z_k) \in \mathbb{R}^k;\ a_i < z_i < b_i,\ 1 \leq i \leq k\}$ with $a_i < b_i$. The basic idea is that whenever fewer indicators than objectives are available, uncountably many disjoint open hyperrectangles arise—a contradiction. Furthermore, we will show a slightly modified statement, which is more general: if Z contains an open hypercube $(u, v)^n$ with $u < v$ such that for any $\boldsymbol{z}^1, \boldsymbol{z}^2 \in (u, v)^n$

$$(\forall\, 1 \leq i \leq k\colon I_i(\{\boldsymbol{z}^1\}) \geq I_i(\{\boldsymbol{z}^2\})) \Leftrightarrow \boldsymbol{z}^1 \succeq \boldsymbol{z}^2$$

then $k \geq n$.

Without loss of generality, assume a minimization problem in the following. We will argue by induction.

$n = 2$: Let $a, b \in (u, v)$ with $a < b$ and consider the incomparable objective vectors (a, b) and (b, a). If $k = 1$, then either $I_1(\{(a, b)\}) \geq I_1(\{(b, a)\})$ or vice versa; this leads to a contradiction to $(a, b) \not\succeq (b, a)$ and $(b, a) \not\succeq (a, b)$.

$n - 1 \to n$: Suppose $n > 2$, $k < n$ and that the statement holds for $n - 1$. Choose $a, b \in (u, v)$ with $a < b$, and consider the $n-1$ dimensional open hypercube $S_c = \{(z_1, z_2, \ldots, z_{n-1}, c) \in (u, v)^n;\ a < z_i < b,\ 1 \leq i \leq n - 1\}$ for an arbitrary $c \in (u, v)$. First, we will show that $I_i(\{(b, \ldots, b, c)\}) < I_i(\{(a, \ldots, a, c)\})$ for all $1 \leq i \leq k$. Assume $I_i(\{(b, \ldots, b, c)\}) \geq I_i(\{(a, \ldots, a, c)\})$ for any i. If $I_i(\{(b, \ldots, b, c)\}) > I_i(\{(a, \ldots, a, c)\})$, then $(a, \ldots, a, c) \not\succeq (b, \ldots, b, c)$, which yields a contradiction. If $I_i(\{(b, \ldots, b, c)\}) = I_i(\{(a, \ldots, a, c)\})$, then $I_i(\{\boldsymbol{z}\}) = I_i(\{(b, \ldots, b, c)\})$ for all $\boldsymbol{z} \in S_c$, because $(a, \ldots, a, c) \succeq \boldsymbol{z}$ if $\boldsymbol{z} \in S_c$. Then, for any $\boldsymbol{z}^1, \boldsymbol{z}^2 \in S_c$, it holds that

$$\forall\, 1 \leq j \leq k,\ j \neq i\colon I_j(\{\boldsymbol{z}^1\}) \geq I_j(\{\boldsymbol{z}^2\}) \Leftrightarrow \boldsymbol{z}^1 \succeq \boldsymbol{z}^2$$

which contradicts the assumption that for any $n-1$ dimensional open hypercube in $\mathbb{R}^{n-1}$, at least $n-1$ indicators are necessary. Therefore, $I_i(\{(b, \ldots, b, c)\}) < I_i(\{(a, \ldots, a, c)\})$.

Now, we consider the image of S_c in indicator space. The vectors $\boldsymbol{I}(\{(b, \ldots, b, c)\})$ and $\boldsymbol{I}(\{(a, \ldots, a, c)\})$ determine an open hyperrectangle $H_c = \{(y_1, y_2, \ldots, y_k) \in \mathbb{R}^k;\ I_i(\{(b, \ldots, b, c)\}) < y_i < I_i(\{(a, \ldots, a, c)\}),\ 1 \leq i \leq k\}$, where $\boldsymbol{I}(\boldsymbol{z}) = (I_1(\boldsymbol{z}), I_2(\boldsymbol{z}), \ldots, I_k(\boldsymbol{z}))$. H_c has the following properties:

1) Since c was arbitrarily chosen within (u, v), there are uncountably many disjoint open hyperrectangles of dimensionality k in the k-dimensional indicator space. This contradiction implies that $k \geq n$. □

2) H_c is open in all k dimensions as for all $1 \leq i \leq k$: $\inf\{y_i;\ (y_1, y_2, \ldots, y_k) \in H_c\} = I_i(\{(b, \ldots, b, c)\}) < I_i(\{(a, \ldots, a, c)\}) = \sup\{y_i;\ (y_1, y_2, \ldots, y_k) \in H_c\}$

3) H_c contains an infinite number of elements.

4) $H_c \cap H_d = \emptyset$ for any $d \in (u, v)$, $d > c$: assume $\boldsymbol{y} \in H_c \cap H_d$; then $\boldsymbol{I}(\{(a, \ldots, a, c)\}) \geq \boldsymbol{y} \geq \boldsymbol{I}(\{(b, \ldots, b, d)\})$, which yields a contradiction as $(a, \ldots, a, c) \not\succeq (b, \ldots, b, d)$.

ACKNOWLEDGMENT

The authors would like to thank W. Ramey for the tip about the proof of Theorem 2, T. Erlebach, S. Chakraborty, and B. Schwikowski for interesting discussions regarding this work, and the anonymous reviewers for their helpful comments.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, pp. 257–271, 1999.

[2] D. A. Van Veldhuizen and G. B. Lamont, "On measuring multiobjective evolutionary algorithm performance," in *Congress on Evolutionary Computation (CEC2000)*, A. Zalzala and R. Eberhart, Eds. Piscataway, NJ: IEEE Press, 2000, vol. 1, pp. 204–211.

[3] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evol. Comput.*, vol. 8, no. 2, pp. 173–195, 2000.

[4] K. C. Tan, T. H. Lee, and E. F. Khor, "Evolutionary algorithms for multiobjective optimization: Performance assessments and comparisons," in *Congress on Evolutionary Computation (CEC2001)*. Piscataway, NJ: IEEE Press, 2001, pp. 979–986.

[5] M. Laumanns, L. Thiele, E. Zitzler, E. Welzl, and K. Deb, "Running time analysis of multi-objective evolutionary algorithms on a simple discrete optimization problem," in *Parallel Problem Solving From Nature (PPSN VII)*, J. J. Merelo *et al.*, Eds. Berlin, Germany: Springer, 2002, pp. 44–53.

[6] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, "A fast elitist nondominated sorting genetic algorithm for multi-objective optimization: NSGA-II," in *Parallel Problem Solving from Nature (PPSN VI)*, M. Schoenauer *et al.*, Eds. Berlin, Germany: Springer, 2000, pp. 849–858.

[7] E. Zitzler and L. Thiele, "Multiobjective optimization using evolutionary algorithms—A comparative case study," in *Parallel Problem Solving from Nature (PPSN V)*, A. E. Eiben *et al.*, Eds. Berlin, Germany: Springer, 1998, pp. 292–301.

[8] M. P. Hansen and A. Jaszkiewicz, "Evaluating the quality of approximations of the nondominated set," Tech. Rep., Inst. of Mathematical Modeling, Tech. Univ. of Denmark, IMM Tech. Rep. IMM-REP-1998-7, 1998.

[9] V. Grunert da Fonseca, C. M. Fonseca, and A. O. Hall, "Inferential performance assessment of stochastic optimizers and the attainment function," in *Evolutionary Multi-Criterion Optimization (EMO2001)*, E. Zitzler *et al.*, Eds. Berlin, Germany: Springer, 2001, pp. 213–225.

[10] J. D. Knowles, D. W. Corne, and M. J. Oates, "On the assessment of multiobjective approaches to the adaptive distributed database management problem," in *Parallel Problem Solving from Nature (PPSN VI)*, M. Schoenauer *et al.*, Eds. Berlin, Germany: Springer, 2000, pp. 869–878.

[11] J. D. Knowles, "Local-search and hybrid evolutionary algorithms for Pareto optimization," Ph.D. dissertation, Dept. of Comput. Sci., Univ. of Reading, Reading, U.K., 2002.

[12] J. Knowles and D. Corne, "On metrics for comparing nondominated sets," in *Congress on Evolutionary Computation (CEC 2002)*. Piscataway, NJ: IEEE Press, 2002, pp. 711–716.

[13] D. A. Van Veldhuizen, "Multiobjective evolutionary algorithms: Classifications, analyzes, and new innovations," Ph.D. dissertation, Graduate School of Eng. of the Air Force Inst. of Technol., Air Univ., June 1999.

[14] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evol. Comput.*, vol. 2, no. 3, pp. 221–248, 1994.

[15] C. M. Fonseca and P. J. Fleming, "On the performance assessment and comparison of stochastic multiobjective optimizers," in *Parallel Problem Solving from Nature (PPSN-IV)*, H.-M. Voigt *et al.*, Eds. Berlin, Germany: Springer, 1996, pp. 584–593.

[16] J. D. Knowles and D. W. Corne, "Approximating the nondominated front using the pareto archived evolution strategy," *Evol. Comput.*, vol. 8, no. 2, pp. 149–172, 2000.

[17] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multi-objective optimization," *Evol. Comput.*, vol. 10, no. 3, pp. 263–282, 2002.

[18] K. Hrbacek and T. Jech, *Introduction to Set Theory*. New York: Marcel Dekker, 1999.

[19] H. Esbensen and E. S. Kuh, "Design space exploration using the genetic algorithm," in *IEEE Int. Symp. Circuits and Systems (ISCAS'96)*. Piscataway, NJ: IEEE Press, 1996, vol. 4, pp. 500–503.

[20] P. Czyzak and A. Jaszkiewicz, "Pareto simulated annealing—A metaheuristic for multiobjective combinatorial optimization," *Multi-Criteria Decision Anal.*, vol. 7, pp. 34–47, 1998.

[21] E. Zitzler, "Evolutionary algorithms for multiobjective optimization: Methods and applications," Ph.D. dissertation, Shaker Verlag, Aachen, Germany, 1999.

[22] E. L. Ulungu, J. Teghem, P. Fortemps, and D. Tuyttens, "Mosa method: A tool for solving multiobjective combinatorial optimization problems," *J. Multi-Criteria Decision Analysis*, vol. 8, pp. 221–236, 1999.

[23] J. Schott, "Fault tolerant design using single and multicriteria genetic algorithm optimization," M.S. dissertation, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technol., Cambridge, 1995.

[24] S. Sayin, "Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming," *Math. Program., Ser. A 87*, pp. 543–560, 2000.

[25] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. Chichester, U.K.: Wiley, 2001.

[26] J. Wu and S. Azarm, "Metrics for quality assessment of a multiobjective design optimization solution set," *Trans. ASME, J. Mechan. Des.*, vol. 123, pp. 18–25, Mar. 2001.

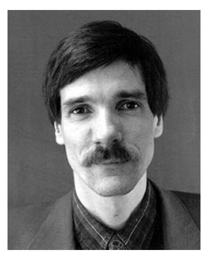
[27] P. De, J. B. Ghosh, and C. E. Wells, "Heuristic estimation of the efficient frontier for a bi-criteria scheduling problem," *Decis. Sci.*, vol. 23, pp. 596–609, 1992.

[28] T. Erlebach, H. Kellerer, and U. Pferschy, "Approximating multi-objective knapsack problems," in *Workshop on Algorithms and Data Structures (WADS2001)*, F. K. H. A. Dehne *et al.*, Eds. Berlin, Germany: Springer, 2001, pp. 210–221.

Eckart Zitzler (M'02) received the diploma degree in computer science from the University of Dortmund, Dortmund, Germany, and the Ph.D. degree in technical sciences from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland.

Since 2000, he has been a Research Associate with the Computer Engineering Group, Department of Information Technology and Electrical Engineering, ETH. His research focuses on bio-inspired computation, multiobjective optimization, computational biology, and computer engineering applications.

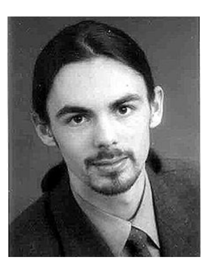
Dr. Zitzler was General Co-Chairman of the first two international conferences on evolutionary multi-criterion optimization (EMO 2001 and EMO 2003) held in Zurich, Switzerland, and Faro, Portugal, respectively.



Lothar Thiele received the Diplom-Ing. and Dr.-Ing. degrees in electrical engineering from the Technical University of Munich, Munich, Germany, in 1981 and 1985, respectively.

During 1981–1987, he was a Research Associate at the Institute of Network Theory and Circuit Design, Technical University of Munich. After finishing his Habilitation thesis, he joined the Information Systems Laboratory, Stanford University, Stanford, CA, in 1987. In 1988, he took up the Chair of Microelectronics at the Faculty of Engineering, University of Saarland, Saarbrücken, Germany. He joined the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, as a Full Professor in computer engineering at the end of 1994. His research interests include models, methods and software tools for the design of hardware/software systems and array processors, as well as the development of parallel algorithms for signal and image processing, combinatorial optimization, and cryptography. He has authored and co-authored more than 100 papers.

Dr. Thiele received the Ph.D. thesis award from the Technical University of Munich in 1986, the Outstanding Young Author Award of the IEEE Circuits and Systems Society in 1987, and the Browder J. Thompson Memorial Prize Award of the IEEE in 1988.



Marco Laumanns received the Diploma degree in computer science from the University of Dortmund, Dortmund, Germany, in 1999.

He is currently a Research Assistant at the Computer Engineering and Networks Laboratory, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. His research interests include multiobjective optimization and evolutionary computation.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.

Carlos M. Fonseca received the Licenciatura degree in electronic and telecommunications engineering from the University of Aveiro, Aveiro, Portugal, in 1991, and the Ph.D. degree from the University of Sheffield, Sheffield, U.K., in 1995, for research into multiobjective genetic algorithms.

He was a Research Associate in the Department of Automatic Control and Systems Engineering, University of Sheffield, during 1994–1997. He then joined the University of Algarve, Faro, Portugal, as an Invited Lecturer in 1997. He was appointed Lecturer in 1998, and is currently the Vice-Coordinator of the Centre for Intelligent Systems. He has been a member of the IFAC Technical Committee on Optimal Control since 1996, and a member of the EvoNet Management Board since 2001. His main research interests are evolutionary multiobjective optimization and its applications to control and systems engineering

He was awarded the Prize "Eng. José Ferreira Pinto Basto" by Alcatel, Portugal, in appreciation of the classification achieved in his Ph.D. degree.

Viviane Grunert da Fonseca received the diploma degree from the University of Dortmund, Dortmund, Germany, in 1994, and the Ph.D. degree from the University of Sheffield, Sheffield, U.K., in 1999, both in statistics.

After a postdoctoral research position at the Faculty of Science and Technology, University of the Algarve, Faro, Portugal, she joined the Instituto Superior Dom Afonso III, Loulé, Portugal, in 2002, where she currently lectures in applied statistics. Her research interests include the statistical performance assessment of stochastic optimizers, the theory of random closed sets, and methodologies in nonparametric statistics.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on June 14,2023 at 18:55:26 UTC from IEEE Xplore. Restrictions apply.