

# The Relationship Between the Covered Fraction, Completeness and Hypervolume Indicators

Viviane Grunert da Fonseca<sup>1,3</sup> and Carlos M. Fonseca<sup>2,3</sup>

<sup>1</sup> INUAF – Instituto Superior D. Afonso III, Loulé, Portugal  
viviane.grunert@sapo.pt

<sup>2</sup> CISUC, Department of Informatics Engineering  
University of Coimbra, Coimbra, Portugal  
cmfonsec@dei.uc.pt

<sup>3</sup> CEG-IST – Center for Management Studies  
Instituto Superior Técnico, Lisbon, Portugal

**Abstract.** This paper investigates the relationship between the covered fraction, completeness, and (weighted) hypervolume indicators for assessing the quality of the Pareto-front approximations produced by multiobjective optimizers. It is shown that these unary quality indicators are all, by definition, weighted Hausdorff measures of the intersection of the region attained by such an optimizer outcome in objective space with some reference set. Moreover, when the optimizer is stochastic, the indicators considered lead to real-valued random variables following particular probability distributions. Expressions for the expected value of these distributions are derived, and shown to be directly related to the first-order attainment function.

**Keywords:** stochastic multiobjective optimizer, performance assessment, covered fraction indicator, completeness indicator, (weighted) hypervolume indicator, attainment function, expected value, Hausdorff measure.

## 1 Introduction

The performance assessment of stochastic multiobjective optimizers (MOs) has become an emerging area of research, enabling the comparison of existing optimizers and supporting the development of new ones. In general, stochastic MO performance can be associated with the distributional behavior of the random outcomes produced in one optimization run, seen either as random non-dominated point (RNP) sets in objective space or, alternatively, as the corresponding random unbounded *attained* sets [7].

Typically, realizations of such random closed sets in  $\mathbb{R}^d$  can be observed arbitrarily often through *multiple* optimization runs. This allows stochastic MO performance assessment and comparison to be carried out with frequency-based statistical inference methodology using (simple) random samples of independent and identically distributed MO outcome sets.

## 2 The Covered Fraction, Completeness and Hypervolume Indicators

However, employing the frequency argument for MO outcome sets is not an easy task. Often, the overall outcome-set distribution is very complex – too complex to be considered as a whole in most practical situations. Therefore, in order to formulate suitable estimators and/or hypothesis tests for MO performance assessment and comparison, it needs to be agreed upon what *partial* aspect of this set distribution one is interested in.

Currently, two main stochastic MO performance assessment approaches are in use, but without really exploiting the relationship between them: the *attainment function approach* [8, 5, 7] and the *quality indicator approach* [17, 15].

The first approach describes the optimizer outcome distribution directly via a hierarchy of nested, increasingly informative attainment functions, where all functions beyond a certain order lead to a full performance description. It is known that the first-order attainment function relates to the *location* and *spread* of an outcome-set distribution, while second and higher-order versions address the corresponding *inter-point dependence structures* [7]. However, to the authors' knowledge, third- and higher-order attainment functions have not yet been considered in practice, due to computational difficulties [5, 6].

The quality indicator approach addresses the complexity of the outcome-set distribution in a different way, through the definition of so-called (unary) quality indicators which somehow transform each realized outcome set into a scalar. The corresponding, much simpler, univariate distributions of indicator values are usually studied via an estimate of their expected value (average indicator value), although other statistics could be considered as well. Even though this approach clearly implies a loss of information about the overall MO performance, there is hardly any discussion in the literature about what aspect of the outcome-set distribution is, or is not, addressed by each indicator.

On the whole, it is generally difficult to combine or compare results of different MO performance studies, unless exactly the same assessment methodology is used. Little is known about which indicators can complement each other and which, when used together, supply redundant information. Hence, there is a need to classify quality indicators with respect to the information they provide. Acknowledging the fact that unary quality indicators are transformations of the original optimizer outcome set, it seems natural to explore this link and attempt to relate the resulting indicator-value distributions with the attainment function hierarchy. As a first step in this direction, the present paper discusses the covered fraction, the completeness and the (weighted) hypervolume indicators with respect to their definition and to the mean (or expected value) of the corresponding distributions.

In Section 2, the attainment function and the quality indicator approaches are briefly outlined. Subsequently, the covered fraction indicator, the completeness indicator and the (weighted) hypervolume indicator are considered in Sections 3 to 5, respectively. The paper ends with a discussion of the results in Section 6 and some concluding remarks in Section 7.

## 2 Approaches to MO Performance Assessment

In the context of performance assessment, the outcome set of a multiobjective optimizer is considered to be the image in objective space of the non-dominated solutions produced in one optimization run (obeying some stopping criterion). When the optimizer is stochastic, such a set of objective vectors is random, and its probability distribution reflects the performance of the optimizer on a given optimization problem instance.

Mathematically, the outcome set of a  $d$ -objective stochastic optimizer can be interpreted as a random closed set [11]. More specifically, it is a random non-dominated point set (RNP set) [7]

$$\mathcal{X} = \{X_1, \dots, X_M \in \mathbb{R}^d : P(X_i \leq X_j) = 0, i \neq j\}, \quad (1)$$

where both the cardinality  $M$  and the elements  $X_i, i = 1, \dots, M$ , are random, and  $P(0 \leq M < \infty) = 1$ . In other words,  $\mathcal{X}$  has a finite, but random, number of elements which do not weakly dominate one another in the Pareto sense [17]. Therefore, stochastic MO performance is related to

1. the (identical) multivariate distribution of the random vectors  $X_1, \dots, X_M$ ,
2. the way in which these random vectors depend on each other (in pairs, triples, quadruples, etc.) and, finally,
3. the univariate distribution of the discrete random variable  $M$ .

For simplicity, it has been common practice to condition on  $M$  with realizations up to a certain value  $m^*$  and/or to study only *partial* aspects of this set distribution. The two main approaches in this context are outlined below.

### 2.1 Attainment Function Approach

Let  $m^*$  be the maximum number of non-dominated objective vectors in  $\mathbb{R}^d$  that may be generated by a  $d$ -objective optimizer. Then, increasing amounts of distributional information about the corresponding outcome set  $\mathcal{X}$  with growing  $k = 1, \dots, m^*$  are comprised in the attainment functions  $\alpha_{\mathcal{X}}^{(k)} : \mathbb{R}^{d \times k} \rightarrow [0, 1]$ , where

$$\alpha_{\mathcal{X}}^{(k)}(z_1, \dots, z_k) = P(\mathcal{X} \preceq z_1 \wedge \dots \wedge \mathcal{X} \preceq z_k), \quad (2)$$

and the event

$$[\mathcal{X} \preceq z] \iff [X_1 \leq z \vee X_2 \leq z \vee \dots \vee X_M \leq z] \quad (3)$$

denotes the attainment of a goal  $z \in \mathbb{R}^d$  by  $\mathcal{X}$ , assuming minimization without loss of generality. Thus, (2) gives the probability that the outcome set  $\mathcal{X}$  weakly dominates the set of goals  $\{z_1, \dots, z_k\}$ . The random unbounded but closed set  $\mathcal{Y} = \{z \in \mathbb{R}^d : \mathcal{X} \preceq z\}$ , which contains all goals  $z \in \mathbb{R}^d$  dominated by at least one element of  $\mathcal{X}$ , is known as the *attained set* [7].

A *complete* distributional performance description (given that  $M \leq m^*$ ) is provided by the attainment function of order  $k = m^*$ , while the first-order attainment function  $\alpha_{\mathcal{X}}^{(1)}(\cdot) = \alpha_{\mathcal{X}}(\cdot)$  is sufficient to characterize where in objective

## 4 The Covered Fraction, Completeness and Hypervolume Indicators

space goals tend to be attained (*location*) and with what degree of variability this happens across multiple runs (*spread*).

The above (theoretical) attainment functions can be estimated from a random sample of independent and identically distributed RNP sets  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n$  via the cumulative frequencies of the corresponding (combinations of) events of the type  $[\mathcal{X} \leq z]$ . For  $k = 1$ , for example, the non-parametric estimator is the *empirical first-order attainment function*

$$\alpha_n(z) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}\{z \in \mathcal{Y}_i\} = \frac{1}{n} \sum_{i=1}^n \mathbf{I}\{\mathcal{X}_i \leq z\}, \quad (4)$$

where  $\mathbf{I}_A(\cdot) = \mathbf{I}\{\cdot \in A\}$  denotes the indicator function of the set  $A$  defined over  $\mathbb{R}^d$ . Estimators for higher-order attainment functions can be constructed in a similar way [7].

## 2.2 Quality Indicator Approach

Here, the outcome RNP set  $\mathcal{X}$  is transformed into a single real-valued random variable  $I(\mathcal{X})$ , usually with respect to some non-empty, deterministic, closed reference set  $Z_{ref} \subset \mathbb{R}^d$ . A variety of such set-transformations are currently in use, each of which defining a particular unary quality indicator that reflects stochastic MO performance, in a restricted sense, via the associated *univariate* probability distribution of indicator values.

Like the distribution of  $\mathcal{X}$ , these indicator (value) distributions are unknown for a given MO application. Arising from the transformation applied to  $\mathcal{X}$ , they depend on the distribution of  $\mathcal{X}$  itself, which can be described through the attainment function hierarchy. Hence, determining the form of this dependence should reveal what aspect(s) of the optimizer outcome-set distribution each indicator actually addresses.

Indicator distributions can, in principle, be estimated as a whole using the information of  $n$  independent optimization runs, in a non-parametric way. For simplicity, however, it is common to begin by estimating their expected values  $E[I(\mathcal{X})]$  through the sample average  $\frac{1}{n} \sum_{i=1}^n I(\mathcal{X}_i)$ . Similarly, this work focuses on the expected value of the distributions produced by the quality indicators considered.

## 3 Covered Fraction Indicator

The original version of the (unary) covered fraction indicator, or coverage indicator, considers the fraction of the Pareto-optimal front  $\mathcal{X}^*$  (assumed to be a finite point set in  $\mathbb{R}^d$ ) that is attained by the outcome set  $\mathcal{X}$  [13, 15]. In this work, a more general definition will be used which, instead of referring to  $\mathcal{X}^*$ , may refer to any deterministic, non-empty, compact (i.e., closed and bounded) reference set  $Z_{ref}$  in objective space  $\mathbb{R}^d$ .

### 3.1 Definition

Given a general subset  $A \subset \mathbb{R}^d$ , denote its (normalized) *Hausdorff measure* of dimension  $j$  by  $\mathcal{H}^j(A)$ , where  $j \geq 0$ . Provided that  $A$  is non-empty, there is a single value of  $j$ , known as the *Hausdorff dimension* of  $A$ , for which  $\mathcal{H}^j(A)$  is finite and positive [1, 12, p. 105]. Then,  $\mathcal{H}^j(A)$  may be understood as the “size” of  $A$ . For example:

- $\mathcal{H}^0(\cdot)$ , also known as the *counting measure*, measures the cardinality of a point set in  $\mathbb{R}^d$  (where any such set has Hausdorff dimension zero),
- $\mathcal{H}^1(\cdot)$  measures the length of a Hausdorff one-dimensional smooth curve in  $\mathbb{R}^d$  (for example a straight line, a circle or an ellipse),
- $\mathcal{H}^2(\cdot)$  measures the area of a Hausdorff two-dimensional smooth surface<sup>4</sup> in  $\mathbb{R}^d$  (for example a plane or the surface of a sphere), and finally
- $\mathcal{H}^d(\cdot)$  measures the hypervolume of a Hausdorff  $d$ -dimensional set in  $\mathbb{R}^d$ , and, as such, corresponds to the usual Lebesgue measure on  $\mathbb{R}^d$ .

Thus, taking a Hausdorff  $j$ -dimensional reference set  $Z_{ref} \subset \mathbb{R}^d$ ,  $0 \leq j \leq d$ , the covered fraction indicator of  $\mathcal{X}$  can be generally defined as<sup>5</sup>

$$I_{CF}(\mathcal{X}, Z_{ref}) = \frac{\mathcal{H}^j(\{z \in Z_{ref} : \mathcal{X} \preceq z\})}{\mathcal{H}^j(Z_{ref})} \quad (5)$$

$$= \frac{1}{\int_{\mathbb{R}^d} \mathbf{I}\{z \in Z_{ref}\} \mathcal{H}^j(dz)} \cdot \int_{Z_{ref}} \mathbf{I}\{\mathcal{X} \preceq z\} \mathcal{H}^j(dz). \quad (6)$$

For Hausdorff  $d$ - and zero-dimensional reference sets in objective space  $\mathbb{R}^d$ , the above definition (with integrals with respect to the measure  $\mathcal{H}^j$ ) can be written in more familiar ways, as follows:

1. For a reference set  $Z_{ref} \subset \mathbb{R}^d$  of Hausdorff dimension  $d$ , the covered fraction indicator of  $\mathcal{X}$  can be defined as a quotient of Lebesgue integrals:

$$I_{CF}(\mathcal{X}, Z_{ref}) = \frac{1}{\int_{\mathbb{R}^d} \mathbf{I}\{z \in Z_{ref}\} dz} \cdot \int_{Z_{ref}} \mathbf{I}\{\mathcal{X} \preceq z\} dz. \quad (7)$$

<sup>4</sup> A  $j$ -dimensional smooth surface is the image of a continuously differentiable mapping (i.e. of class  $C^1$ ) from  $\mathbb{R}^j$  onto  $\mathbb{R}^d$ , where  $j < d$  [9, p. 355]. Note that a countable union of Hausdorff  $j$ -dimensional sets preserves the Hausdorff dimension [4, p. 112], i.e. any such union of smooth sets is also of integer Hausdorff dimension. A non-mathematical discussion of the concept of “smoothness” can be found in [3, p. 335].

<sup>5</sup> The Hausdorff dimension of any realization of the outcome set  $\mathcal{X}$  is equal to zero, and the Hausdorff dimension of a non-empty *attained set* realization is equal to  $d$ . Non-smooth reference sets, such as fractal sets, may have a non-integer Hausdorff dimension.

## 6 The Covered Fraction, Completeness and Hypervolume Indicators

2. For a discrete reference set  $Z_{ref}^k = \{z_1, \dots, z_k\}$  of  $k$  not necessarily non-dominated points in  $\mathbb{R}^d$ , the definition simplifies to

$$I_{CF}(\mathcal{X}, Z_{ref}^k) = \frac{1}{k} \cdot \sum_{j=1}^k \mathbf{I}\{\mathcal{X} \preceq z_j\}. \quad (8)$$

In any case, the covered fraction indicator of  $\mathcal{X}$  takes realizations in  $[0, 1]$ , where a larger observed indicator value is considered to represent a “better” optimization result (in the particular sense of the indicator).

### 3.2 Expected Value

Due to the “linearity property of expectation”, and since the binary random variable  $\mathbf{I}\{\mathcal{X} \preceq z\}$  follows a Bernoulli distribution with expected value  $P(\mathcal{X} \preceq z) = \alpha_{\mathcal{X}}(z)$ , for all  $z \in \mathbb{R}^d$ , it can be seen that the expected value of the covered fraction indicator distribution is related to the first-order attainment function. For a Hausdorff  $j$ -dimensional reference set  $Z_{ref}$ ,  $0 \leq j \leq d$ , it holds that

$$\mathbb{E}[I_{CF}(\mathcal{X}, Z_{ref})] = \frac{\mathbb{E}[\mathcal{H}^j(\{z \in Z_{ref} : \mathcal{X} \preceq z\})]}{\mathcal{H}^j(Z_{ref})} \quad (9)$$

$$= \frac{1}{\int_{\mathbb{R}^d} \mathbf{I}\{z \in Z_{ref}\} \mathcal{H}^j(dz)} \cdot \int_{Z_{ref}} \alpha_{\mathcal{X}}(z) \mathcal{H}^j(dz), \quad (10)$$

where (10) can be obtained by applying Robbins’s theorem [11, p. 59] to the random compact set  $\{z \in Z_{ref} : \mathcal{X} \preceq z\}$ , since  $\mathcal{H}^j$  is *locally finite* on a Hausdorff  $j$ -dimensional  $Z_{ref}$ .

For Hausdorff  $d$ - and zero-dimensional reference sets, the above formula can again be simplified:

1. For a reference set  $Z_{ref} \subset \mathbb{R}^d$  of Hausdorff dimension  $d$ ,

$$\mathbb{E}[I_{CF}(\mathcal{X}, Z_{ref})] = \frac{1}{\int_{\mathbb{R}^d} \mathbf{I}\{z \in Z_{ref}\} dz} \cdot \int_{Z_{ref}} \alpha_{\mathcal{X}}(z) dz. \quad (11)$$

2. For a discrete reference set  $Z_{ref}^k = \{z_1, \dots, z_k\}$  of Hausdorff dimension zero,

$$\mathbb{E}[I_{CF}(\mathcal{X}, Z_{ref}^k)] = \frac{1}{k} \cdot \sum_{j=1}^k \alpha_{\mathcal{X}}(z_j). \quad (12)$$

## 4 Completeness Indicator

In its original form, the completeness indicator is defined for a given (observed) solution set in decision space (after one optimization run), as the probability of selecting a point uniformly at random from the feasible set which is weakly dominated by that solution set [10, 15].

#### 4.1 Definition

For a random outcome set  $\mathcal{X}$  in objective space, the original completeness indicator can be defined as the conditional probability of  $\mathcal{X}$  attaining a non-uniformly distributed random reference point  $V$  from the image of the feasible set, given  $\mathcal{X}$ . However, the indicator may also be defined more generally by considering some deterministic, non-empty, closed reference set  $Z_{ref} \subset \mathbb{R}^d$  (which may or may not be the image of the feasible set), together with a random vector  $V$  taking realizations in  $Z_{ref}$  and following an explicitly defined distribution. Hence,

$$I_{CO}(\mathcal{X}, V, Z_{ref}) = P(\mathcal{X} \preceq V \mid \mathcal{X}) . \quad (13)$$

Like the covered fraction indicator of  $\mathcal{X}$ , the completeness indicator of  $\mathcal{X}$  is a random variable which takes realizations in  $[0, 1]$ , where larger values correspond to “better” optimization results.

In fact, it can be seen that the definitions of the two indicators are very similar, when interpreting  $I_{CO}(\mathcal{X}, V, Z_{ref})$  as the conditional expectation of  $\mathbf{I}\{\mathcal{X} \preceq V\}$  given  $\mathcal{X}$ : with a Hausdorff  $j$ -dimensional reference set  $Z_{ref} \subset \mathbb{R}^d$ ,  $0 \leq j \leq d$ , and a random vector  $V$  supported on  $Z_{ref}$ , stochastically *independent* from the outcome set  $\mathcal{X}$ , and for which the probability density function  $f_V(\cdot)$  exists with respect to Hausdorff measure  $\mathcal{H}^j$ , it holds that

$$I_{CO}(\mathcal{X}, V, Z_{ref}) = P(\mathcal{X} \preceq V \mid \mathcal{X}) \quad (14)$$

$$= \mathbf{E}[\mathbf{I}\{\mathcal{X} \preceq V\} \mid \mathcal{X}] \quad (15)$$

$$= \int_{Z_{ref}} \mathbf{I}\{\mathcal{X} \preceq z\} \cdot f_{V|\mathcal{X}}(z) \mathcal{H}^j(dz) \quad (16)$$

$$= \int_{Z_{ref}} \mathbf{I}\{\mathcal{X} \preceq z\} \cdot f_V(z) \mathcal{H}^j(dz) . \quad (17)$$

Clearly, the particular choice of a *uniform* distribution for  $V = V_u$  over some compact reference set  $Z_{ref}$ , again stochastically independent from  $\mathcal{X}$ , leads to the identity of the covered fraction indicator and the completeness indicator. In this case,

$$f_{V_u}(z) = f_u(z) = 1/\mathcal{H}^j(Z_{ref}) \quad (18)$$

for all  $z \in Z_{ref}$ , and  $f_u(z) = 0$  otherwise.

Finally, note that for both Hausdorff  $d$ -dimensional reference sets and zero-dimensional reference sets, the formulation in (17) can be simplified, respectively in the sense of (7) and (8), by substituting the Hausdorff measure  $\mathcal{H}^j$  either by the usual Lebesgue measure on  $\mathbb{R}^d$  or by the counting measure for point sets.

#### 4.2 Expected Value

From (17) and the argumentation given in subsection 3.2, it immediately follows that

$$\mathbf{E}[I_{CO}(\mathcal{X}, V, Z_{ref})] = \int_{Z_{ref}} \alpha_{\mathcal{X}}(z) \cdot f_V(z) \mathcal{H}^j(dz) , \quad (19)$$

## 8 The Covered Fraction, Completeness and Hypervolume Indicators

while more familiar formulations can be achieved for Hausdorff  $d$ - and zero-dimensional reference sets:

1. For a reference set  $Z_{ref} \subset \mathbb{R}^d$  of Hausdorff dimension  $d$ ,

$$E[I_{CO}(\mathcal{X}, V, Z_{ref})] = \int_{Z_{ref}} \alpha_{\mathcal{X}}(z) \cdot f_V(z) dz . \quad (20)$$

2. For a discrete reference set  $Z_{ref}^k = \{z_1, \dots, z_k\}$  of Hausdorff dimension zero,

$$E[I_{CO}(\mathcal{X}, V, Z_{ref}^k)] = \sum_{j=1}^k \alpha_{\mathcal{X}}(z_j) \cdot P(V = z_j) . \quad (21)$$

## 5 Hypervolume Indicator

The hypervolume indicator (also known as dominated space), introduced in [16], considers the size of the portion of a deterministic, non-empty, closed reference set in objective space  $\mathbb{R}^d$  that is attained by the outcome set  $\mathcal{X}$ . Later, this idea was refined by including a weight function that assigns varying levels of importance to different regions of the reference set [14].

In contrast to the covered fraction and completeness indicators, the reference set for the (weighted) hypervolume indicator is usually specified through a single reference point  $z_{ref}^* \in \mathbb{R}^d$ , and can be written as

$$Z_{ref}^* = \{z \in \mathbb{R}^d : z \leq z_{ref}^*\} , \quad (22)$$

though in some cases the Pareto-optimal front  $\mathcal{X}^*$  or some other set of non-dominated points below  $z_{ref}^*$  is used to bound  $Z_{ref}^*$  from below, in such a way that it is still Hausdorff  $d$ -dimensional.

### 5.1 Definition

For a reference set  $Z_{ref}^*$  as given in (22), the (weighted) hypervolume indicator is defined as

$$I_H(\mathcal{X}, w, Z_{ref}^*) = \int_{Z_{ref}^*} \mathbf{I}\{\mathcal{X} \preceq z\} \cdot w(z) dz , \quad (23)$$

where  $w(\cdot)$  is a non-negative valued weight function, integrable over  $Z_{ref}^*$ , i.e.

$$\int_{Z_{ref}^*} w(z) dz < \infty . \quad (24)$$

In this general form, the (weighted) hypervolume indicator of  $\mathcal{X}$  takes realizations in the interval

$$\left[0, \int_{Z_{ref}^*} w(z) dz\right] , \quad (25)$$

where larger observed indicator values are associated with “better” optimization results.

Some choices for the weight function  $w(\cdot)$  deserve special attention:



- The indicator weight function,  $w(z) = \mathbf{I}_{Z_{ref}^*}(z)$ ,  $z \in \mathbb{R}^d$ , defines the classical, non-weighted hypervolume indicator with possible values ranging from 0 to the size of  $Z_{ref}^*$ , i.e.,  $\mathcal{H}^d(Z_{ref}^*)$ , if the reference set is compact.
- When  $w(\cdot)$  is the *probability density* function [2] of a random variable  $V$ ,

$$\int_{Z_{ref}^*} w(z) dz = 1, \quad (26)$$

and all possible indicator values are contained in the interval  $[0, 1]$ . Further, if  $V$  is independent from  $\mathcal{X}$ , the (weighted) hypervolume indicator corresponds to the completeness indicator with respect to the reference set  $Z_{ref}^*$ .

- For the probability density function  $w(\cdot) = f_u(\cdot)$  of a random variable  $V_u$ , *uniformly* distributed over some compact reference set  $Z_{ref}^*$  and independent from  $\mathcal{X}$ , all three indicators considered in this paper are identical, i.e.

$$I_{CF}(\mathcal{X}, Z_{ref}^*) = I_{CO}(\mathcal{X}, V_u, Z_{ref}^*) = I_H(\mathcal{X}, f_u, Z_{ref}^*). \quad (27)$$

## 5.2 Expected value

Again, from the argumentation given in subsection 3.2, it follows that

$$\mathbb{E}[I_H(\mathcal{X}, w, Z_{ref}^*)] = \int_{Z_{ref}^*} \alpha_{\mathcal{X}}(z) \cdot w(z) dz. \quad (28)$$

## 6 Discussion

The covered fraction, completeness and weighted hypervolume indicators are all, by definition, weighted Hausdorff measures of the intersection of the attained set  $\mathcal{Y} = \{z \in \mathbb{R}^d : \mathcal{X} \sqsubseteq z\}$  with some reference set  $Z_{ref}$  in objective space  $\mathbb{R}^d$ . As such, they can be seen as special cases of a *generalized hypervolume indicator* of  $\mathcal{X}$ , defined as

$$I_{GH}(\mathcal{X}, w, Z_{ref}) = \int_{Z_{ref}} \mathbf{I}\{\mathcal{X} \sqsubseteq z\} \cdot w(z) \mathcal{H}^j(dz), \quad (29)$$

where

- $Z_{ref}$  is a Hausdorff  $j$ -dimensional, deterministic, closed reference set in  $\mathbb{R}^d$ ,  $0 \leq j \leq d$ , representing the region in objective space which is of interest for performance assessment, and
- $w(\cdot)$  is a non-negative valued weight function, integrable over  $Z_{ref}$ , which assigns different levels of importance to the points in  $Z_{ref}$ .

Note that, technically, the weight function  $w(\cdot)$  alone would be sufficient to parametrize this indicator, as its support would define the corresponding reference set  $Z_{ref}$ . In other words, the generalized hypervolume indicator measures the *mass* of the region attained by an optimizer outcome set, as determined by

## 10 The Covered Fraction, Completeness and Hypervolume Indicators

**Table 1.** Indicators as special cases of the *generalized hypervolume indicator*.

Indicator	Reference set	Weight function
weighted hypervolume	$Z_{ref}^*$ , closed	integrable over $Z_{ref}^*$
non-weighted hypervolume	$Z_{ref}^*$ , compact	indicator function $\mathbf{I}_{Z_{ref}^*}(\cdot)$
completeness	$Z_{ref}$ , closed	density function $f_V$ with support on $Z_{ref}$ , $V$ independent of $\mathcal{X}$
covered fraction	$Z_{ref}$ , compact	uniform density function with support on $Z_{ref}$

some *mass density* function  $w(\cdot)$ , which may or may not be also a probability density function. The generalized hypervolume indicator can be reduced to each of the indicators considered in this paper, as summarized in Table 1.

As a function of the outcome set  $\mathcal{X}$ , the indicator  $I_{GH}(\mathcal{X}, w, Z_{ref})$  leads to a random variable with realizations in the interval

$$\left[0, \int_{Z_{ref}} w(z) \mathcal{H}^j(dz)\right], \quad (30)$$

where larger indicator values indicate a “better” optimization result. At this point it should be noted, however, that for the purpose of “unit-independent” comparisons of optimizer performance, it may be preferable to always use a normalizing density weight function, so that indicator values can be limited to the interval  $[0, 1]$ .

For a general weight function  $w(\cdot)$ , the expected value of the generalized hypervolume indicator distribution can be expressed as

$$\mathbb{E}[I_{GH}(\mathcal{X}, w, Z_{ref})] = \int_{Z_{ref}} \alpha_{\mathcal{X}}(z) \cdot w(z) \mathcal{H}^j(dz). \quad (31)$$

For the *probability density* function of a random vector  $V$ , supported on  $Z_{ref}$  and distributed *independently* from the outcome set  $\mathcal{X}$ , this additionally leads to the interesting relationship:<sup>6</sup>

$$\mathbb{E}[I_{GH}(\mathcal{X}, f_V, Z_{ref})] = \mathbb{E}[\alpha_{\mathcal{X}}(V)] = P(\mathcal{X} \leq V). \quad (32)$$

Hence, the expected value of the generalized hypervolume indicator, and consequently of all three indicators considered in this paper, can be obtained from the first-order attainment function.

This result is important because it sheds light on the relationship between the quality indicator approach and the attainment function approach. It can be seen that the three indicators considered, unified in the generalized hypervolume indicator, convey information related to the *location* of the optimizer

<sup>6</sup> For independent  $\mathcal{X}$  and  $V$ :  $\mathbb{E}[P(\mathcal{X} \leq V | V)] = \mathbb{E}[P(\mathcal{X} \leq V | \mathcal{X})] = P(\mathcal{X} \leq V)$ .

outcome-set distribution via their mean indicator values. Moments of the generalized hypervolume indicator distribution other than the mean may still contain information beyond that captured by the first-order attainment function.

The generalized hypervolume indicator also allows new indicators to be constructed using less usual reference sets and/or weight functions. For example, assume that the outcome set  $\mathcal{X}$  of a hypothetical two-objective optimizer, when applied to a given problem instance, is simply the set of minima [6] of a set of two stochastically independent random vectors in  $\mathbb{R}^2$ . Assume also that both vectors are distributed according to a bivariate exponential distribution with parameter  $\lambda > 0$ , with density function

$$f(t_1, t_2) = \lambda^2 \cdot e^{-\lambda \cdot t_1 - \lambda \cdot t_2} \cdot \mathbf{I}_{[0, \infty)^2}((t_1, t_2)') \quad (33)$$

and cumulative distribution function

$$F(t_1, t_2) = (1 - e^{-\lambda \cdot t_1} - e^{-\lambda \cdot t_2} + e^{-\lambda \cdot t_1 - \lambda \cdot t_2}) \cdot \mathbf{I}_{[0, \infty)^2}((t_1, t_2)') . \quad (34)$$

Then, the first-order attainment function of  $\mathcal{X}$  at a goal  $z = (t_1, t_2)' \in \mathbb{R}^2$  is

$$\alpha_{\mathcal{X}}(z) = 1 - (1 - F(t_1, t_2))^2 . \quad (35)$$

Given  $(r_1, r_2)' \in \mathbb{R}_+^2$ , the reference set  $Z_{seg} = \{(r_1 \cdot t, r_2 - r_2 \cdot t)' : t \in [0, 1]\}$  is a line segment in  $\mathbb{R}^2$ , and has Hausdorff dimension 1. In this case, the generalized hypervolume indicator with weight function  $\mathbf{I}_{Z_{seg}}(\cdot)$  measures the *length* of the intersection of this segment with the attained set. From (31), the expected indicator value may then be calculated via the line integral

$$\mathbb{E}[I_{GH}(\mathcal{X}, \mathbf{I}_{Z_{seg}}, Z_{seg})] = \sqrt{r_1^2 + r_2^2} \cdot \int_0^1 \alpha_{\mathcal{X}}((r_1 \cdot t, r_2 - r_2 \cdot t)') dt . \quad (36)$$

Clearly, the expected value can also be directly evaluated using the standard integral based on the joint density of the two exponential random vectors, but the latter leads to much more complicated expressions.

## 7 Concluding Remarks

In this paper the relationship between three popular quality indicators for multiobjective optimizer performance assessment has been studied with respect to their definitions. By considering Hausdorff measures, a *single* notation was introduced for all indicators, leading to their unification. Furthermore, it was shown that, when the optimizer is stochastic, the corresponding expected indicator values depend on the first-order attainment function of the optimizer outcomes.

Investigating how the variance and other aspects of these univariate indicator-value distributions relate to the attainment function hierarchy will be the subject of future work, as well as considering other unary quality indicators. In particular, it would be interesting to see how indicators designed to assess the dispersion or uniformity of points in observed approximation sets relate in distribution to (higher-order) attainment functions.

## References

1. Alberti, G.: Geometric measure theory. In: Françoise, J.P., et al. (eds.) *Encyclopedia of Mathematical Physics*, vol. 2, pp. 520–527. Elsevier, Oxford (2006)
2. Bader, J.M.: *Hypervolume-Based Search for Multiobjective Optimization: Theory and Methods*. Ph.D. thesis, Swiss Federal Institute of Technology, Zurich (2009)
3. Barenblatt, G.I.: *Scaling, Self-similarity, and Intermediate Asymptotics*. Cambridge University Press, Cambridge (1996)
4. DiBenedetto, E.: *Real Analysis*. Birkhäuser, Boston (2002)
5. Fonseca, C.M., Grunert da Fonseca, V., Paquete, L.F.: Exploring the performance of stochastic multiobjective optimisers with the second-order attainment function. In: Coello Coello, C.A., et al. (eds.) *EMO 2005, LNCS*, vol. 3410, pp. 250–264. Springer, Heidelberg (2005)
6. Fonseca, C.M., Guerreiro, A.P., López-Ibáñez, M., Paquete, L.: On the computation of the empirical attainment function. In: Takahashi, R.H.C., et al. (eds.) *EMO 2011, LNCS*, vol. 6576, pp. 106–120. Springer, Heidelberg (2011)
7. Grunert da Fonseca, V., Fonseca, C.M.: The attainment-function approach to stochastic multiobjective optimizer assessment and comparison. In: Bartz-Beielstein, T., et al. (eds.) *Experimental Methods for the Analysis of Optimization Algorithms*, chap. 5, pp. 103–130. Springer, Berlin (2010)
8. Grunert da Fonseca, V., Fonseca, C.M., Hall, A.O.: Inferential performance assessment of stochastic optimisers and the attainment function. In: Zitzler, E., et al. (eds.) *EMO 2001, LNCS*, vol. 1993, pp. 213–225. Springer, Heidelberg (2001)
9. Ito, K. (ed.): *Encyclopedic Dictionary of Mathematics 2*. The Mathematical Society of Japan. The MIT Press (1987)
10. Lotov, A.V., Bushenkov, V.A., Kamenev, G.K.: *Interactive Decision Maps: Approximation and Visualization of Pareto Frontier*. Kluwer Academic Publishers, Dordrecht (2004)
11. Molchanov, I.: *Theory of Random Sets*. Springer, London (2005)
12. Ott, E.: *Chaos in Dynamical Systems*. Cambridge University Press, Cambridge (2002)
13. Ulungu, E.L., Teghem, J., Fortemps, P.H., Tuyttens, D.: MOSA method: A tool for solving multiobjective combinatorial optimization problems. *Journal of Multi-Criteria Decision Analysis* 8(4), 221–236 (1999)
14. Zitzler, E., Brockhoff, D., Thiele, L.: The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. In: Obayhashi, S., et al. (eds.) *EMO 2007, LNCS*, vol. 4403, pp. 862–876. Springer, Heidelberg (2007)
15. Zitzler, E., Knowles, J., Thiele, L.: Quality assessment of Pareto set approximations. In: Branke, J., et al. (eds.) *Multiobjective Optimization. Interactive and Evolutionary Approaches, LNCS*, vol. 5252, pp. 373–404. Springer, Heidelberg (2008)
16. Zitzler, E., Thiele, L.: Multiobjective optimization using evolutionary algorithms – A comparative case study. In: Eiben, A.E., et al. (eds.) *PPSN V 1998, LNCS*, vol. 1498, pp. 292–301. Springer, Heidelberg (1998)
17. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Grunert da Fonseca, V.: Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation* 7(2), 117–132 (2003)