Consistencies and Contradictions of Performance Metrics in Multiobjective Optimization

Siwei Jiang, Yew-Soon Ong, Jie Zhang, Liang Feng

***Abstract*—An important consideration of Multiobjective Optimization (MOO) is the quantitative metrics used for defining the optimality of different solution sets, which is also the basic principle for the design and evaluation of MOO algorithms. Although a plethora of performance metrics have been proposed in the MOO context, there has been a lack of insights on the relationships between metrics. In this paper, we first group the major MOO metrics proposed to date according to four core performance criteria considered in the literature, namely *Capacity*, *Convergence*, *Diversity*, and *Convergence–Diversity*. Then, a comprehensive study is conducted to investigate the relationships among representative group metrics, including *Generational Distance* (GD), ϵ-*indicator* ($I^1_{\epsilon+}$), *Spread* (Δ), *Generalized Spread* (Δ^*), *Inverted Generational Distance* (IGD) and *Hypervolume* (HV). Experimental results indicated that these six metrics show high consistencies when Pareto fronts (PFs) are convex, whereas they show certain contradictions on concave PFs.**

***Index Terms*—Multiobjective Optimization, Performance Metrics, Capacity, Convergence, Diversity, Hypervolume, jMetal**

I. Introduction

Multiobjective Optimization Problems (MOPs) involve several conflicting objectives to be optimized simultaneously [1–5]. A plethora of approaches, such as Multiobjective Evolutionary Algorithms (MOEAs), have been well established as efficient methods to deal with MOPs that are now prevalent in the fields of engineering, finance, logistic, etc [1–40].

To evaluate different approaches, it is critical to design appropriate performance metrics in various context. For instance, the goal of Single-objective Optimization Problems (SOPs) is to find an optimal solution with regards to minimization or maximization. It is relatively easy in SOPs to compare two solutions and regard the one with a better fitness as superior. In MOPs, however, evaluating solution superiority is much more complex due to the presence of conflicting objectives. Various approaches often obtain an *optimal solution set* comprising a number of solutions that fair equivalently according to Pareto dominance concept [1–19]. However, it is often non-trivial to provide a quantitative comparison of different optimal solution sets in Multiobjective Optimization (MOO) [1–5].

To date, a number of quantitative metrics have been proposed in MOO for defining solution set optimality and the assessment of MOEAs [4–6, 41–45]. Each metric is designed with a standpoint that takes one or more performance criteria into considerations. The typical performance criteria[1] include the capacity of the non-dominated solutions set, the convergence of solutions to true Pareto fronts (PFs), the diversity of solutions in the objective space, the dominated volume of solutions with respect to the reference sets, etc. For instance, Van Veldhuizen et al. [41] designed metrics to tally the number of non-dominated solutions. The *Generational Distance* (GD) metric [6, 42] is introduced to measure the proximity of solutions to the true PFs. Zhou et al. [43] defined the *Generalized Spread* (Δ^*) metric to measure the diversity of solutions on high dimensional MOPs. Zitzler et al. [44, 45], on the other hand, established the popular *Hypervolume* (HV) metric to calculate the dominated volume of the optimal solution sets, with respect to the reference sets.

With the plethora of performance metrics that have been proposed in the last decade, research efforts devoted to surveys on MOO metrics have also emerged alongside [46–51]. For instance, Okabe et al. [46] in their survey categorized MOO metrics in terms of cardinality, distance, volume, distribution and spread. Zitzler et al. [47] classified the MOO metrics using a mathematical framework. After analysing the classical MOO metrics, Tan et al. [48] proposed the Uniform Distribution (UD) as a parameter dependent metric. Wu et al. [51], on the other hand, summarized six variants of metrics associated with HV, and studied them on a bi-objective engineering MOP.

It is worth noting that several notable surveys, experimental studies and analyses made on MOO metrics [46–51] have been limited to a focus on individual metric. Little or no work on establishing the relationships among MOO metrics have been explicitly considered to date. This paper thus makes an attempt to fill in this gap. In particular, we begin with a categorization of existing MOO metrics into four groups according to the core performance criteria typically used in the literature, namely *Capacity*, *Convergence*, *Diversity*, and *Convergence–Diversity*. With the categorization, we are then able to systematically study the relationships among representative group metrics on the symmetric and continuous PFs. More specifically, we investigate six representative metrics, including *Generational Distance* (GD), ϵ-*indicator* ($I^1_{\epsilon+}$), *Spread* (Δ), *Generalized Spread* (Δ^*), *Inverted Generational Distance* (IGD) and *Hypervolume* (HV). Experimental studies indicated that all the six metrics show high consistencies when the Pareto Fronts (PFs) are convex. Surprisingly, contradictions among metrics are found on concave PFs, even on metrics of the same group that cater to a common performance criterion.

Siwei Jiang, Yew-Soon Ong, Jie Zhang, Liang Feng are with the School of Computer Engineering, Nanyang Technological University, Singapore 639798 (e-mail: {sjiang1, asysong, zhangj, feng0039}@ntu.edu.sg).

Manuscript received August 20, 2013; Revised January 2, 2014; Accepted February 11, 2014

[1]Performance criteria in MOO: the particular standard of the true quality of optimal solution sets, i.e., capacity, convergence, diversity.

In summary, the contributions of this paper are listed:

1) A classification of MOO metrics according to the performance criteria, including *Capacity*, *Convergence*, *Diversity*, and *Convergence–Diversity*, is presented.
2) The inadequacies of some MOO metrics in omitting partial information on the true quality of optimal solution sets are illustrated.
3) The relationships of six representative metrics are analyzed based on experimental studies. Specifically, two metrics, $I^1_{\epsilon+}$ and IGD, are consistent with HV on convex PFs, whereas, certain contradictions to HV are uncovered on concave PFs.
4) Last but not least, the presented results will serve as a guide on the appropriate use of MOO metrics.

The rest of the paper is organized as follows. Section II gives a brief background on MOO. In Section III, we survey various MOO metrics and categorize them into four groups. Then, the inadequacies of some MOO metrics are highlighted in Section IV. In Section V, experimental results on the relationships among representative group metrics are presented. The conclusions and future works are then given in Section VI.

II. Background

Without loss of generality, Multiobjective Optimization Problem (MOP) for minimization can be stated as [1–5]:

$$\begin{array}{rcl} min & F(\vec{x}) & = & (f_1(\vec{x}), \ldots, f_m(\vec{x})) \\ s.t. & G(\vec{x}) & \leq & 0,\ H(\vec{x}) = 0,\ \vec{x} \in \Omega, \end{array} \tag{1}$$

where $\vec{x} = (x_1, \ldots, x_D)$, Ω is the *variable space*, R^m is the *objective space*, and $F : \Omega \rightarrow R^m$ consists of m real-valued objective functions with constraints $G(\vec{x}) \leq 0$ and $H(\vec{x}) = 0$, the feasible solution space is $\Omega = \Pi_{i=1}^{D}[L_i, U_i]$, and L_i, U_i are the lower and upper bound of x_i, respectively.

Let $\vec{u} = (u_1, \cdots, u_m), \vec{v} = (v_1, \cdots, v_m)$ be two vectors. Here, three Pareto dominance concepts are defined as follows: 1) $\vec{u}$ is said to *weakly dominate* $\vec{v}$, notated as $\vec{u} \preceq \vec{v}$ iff $\forall i : u_i \leq v_i$; 2) $\vec{u}$ *dominates* $\vec{v}$, notated as $\vec{u} \prec \vec{v}$ iff $\forall i : u_i \leq v_i$ and $\exists i : u_i < v_i$; and 3) $\vec{u}$ *strongly dominates* $\vec{v}$, notated as $\vec{u} \prec\prec \vec{v}$ iff $\forall i : u_i < v_i$.

Other important definitions of MOPs are outlined:

Pareto Set. All solution vectors in the variable space that non-dominates each other form the *Pareto Set* (PS), and notated as $PS = \{\vec{x} \in \Omega | \nexists \vec{y} \in \Omega : F(\vec{y}) \preceq F(\vec{x})\}$.

Pareto front. The *Pareto front* (PF) lies in the objective space, and has relation to PS such that $PF = \{F(\vec{x}) | \vec{x} \in PS\}$. In general, it is impossible to find all solutions on the continuous PFs. Hence, a finite number of non-dominated solutions that approximates the true PFs is termed as P.

Optimal Solution Set. The *optimal solution set* obtained by the optimizers (e.g., MOEAs) is termed as S.

Reference Set. The *reference set* is designed with predefined solutions, and is defined as R.

Good (Utopian) Point and Bad Point. A *Good Point* is termed as $P_G = (f_1^G, \cdots, f_m^G)$, where $\forall \vec{x} \in PS : P_G \prec F(\vec{x})$. A *Bad Point* is defined as $P_B = (f_1^B, \cdots, f_m^B)$, $\forall \vec{x} \in PS : P_B \succ F(\vec{x})$.

To evaluate the quality of set S, three major performance criteria have been considered in MOO: 1) the number of non-dominated solutions in S, 2) the convergence of S to the true PFs, and 3) the diversity of S in the objective space [1–5, 52]. In particular, an optimal solution set with large number of non-dominated solutions, approaching the true PFs and scattering evenly are generally desirable.

III. Performance Metrics in MOO

In the design of MOO metrics, three major performance criteria, namely capacity, convergence and diversity, have typically been taken into considerations. Based on these criteria, we categorize the MOO metrics into four core groups:

- *Capacity* metrics: This group metrics tally the number or ratio of non-dominated solutions in S that satisfies given predefined requirements.
- *Convergence* metrics: These are metrics for measuring the proximity of optimal solution set S to PF (P).
- *Diversity* metrics: These metrics include two forms of information: 1) Distribution measures how evenly scattered are the solutions of S in the objective space, and 2) Spread indicates how well do the solutions of S arrive at the extrema of true PFs.
- *Convergence–Diversity* metrics: They indicate both the convergence and diversity of S on a single scale.

A. Capacity Metrics

Capacity metrics quantify the number or ratio of non-dominated solutions in S that conforms to the predefined requirements. In general, a large number of non-dominated solutions in S is preferred.

In [41], the *Overall Non-dominated Vector Generation* (ONVG) metric is introduced as:

$$\text{ONVG}(S) = |S|. \tag{2}$$

This gives the number of the non-dominated solutions in the optimal solution set (S). Hereafter $|\cdot|$ defines the cardinality or number of elements in the set, unless explicitly indicated.

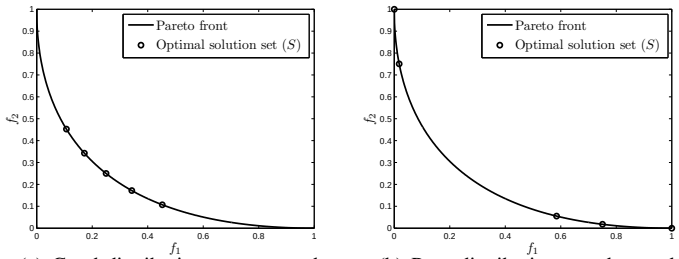
The *Overall Non-dominated Vector Generation Ratio* (ONVGR) [41], which gives the capacity ratio of the optimal solution set (S) with respect to the PF (P), is given as:

$$\text{ONVGR}(S, P) = \frac{|S|}{|P|}. \tag{3}$$

Several metrics have also been designed by considering the search stages. For instance, [41] defined three metrics: 1) *Generational Non-dominated Vector Generation* $\text{GNVG}(S) = |S(t)|$, 2) *Generational Non-dominated Vector Generation Ratio* $\text{GNVGR}(S, P) = \frac{|S(t)|}{|P|}$, and 3) *Non-dominated Vector Additional* $\text{NVA}(S, t) = \text{GNVG}(S, t) - \text{GNVG}(S, t-1)$, where t is the generation index. Thus, they measure the capacity of S or the changes in capacity along the MOO search.

While the above five metrics are associated with the size of set S, the *Error Ratio* (ER) in [53], on the other hand, considered the solution intersections between S and PF (P). It takes the form of:

$$\text{ER}(S, P) = 1 - \frac{|S \bigcap P|}{|P|}, \tag{4}$$



(a) Good distribution, poor spread (b) Poor distribution, good spread

Fig. 1. Two components (distribution and spread) in *diversity* metrics.

where $S \bigcap P$ denotes the solutions existing in both S and P. By replacing P with reference set R, the *Ratio of the Reference Points Found* in [54, 55] is proposed as $C1_R = \frac{|S \bigcap R|}{|R|}$.

As mentioned in Section II, set P consists of a finite number of non-dominated solutions that approximate the true PFs. Hence, it is often infeasible to arrive at the exact same solutions (usually comprising of real values in the objective space) in both S and P. For this reason, the metric *Non-dominated Points by Reference Set* ($C2_R$) [55], which adopts the Pareto dominance concept, is introduced as follows:

$$C2_R(S,R) = \frac{|\{\vec{s} \in S | \nexists \vec{r} \in R : \vec{r} \prec \vec{s}\}|}{|S|}. \quad (5)$$

In other words, Eq. 5 estimates how many solutions of S are non-dominated by reference set R.

In contrast to the eight *capacity* metrics presented, which only elicit information from one optimal solution set, the *Coverage of Two Sets* (or Metric C) in [44, 45] concentrates on the overlaps between two optimal solution sets.

$$C(S_1,S_2) = \frac{|\{\vec{s}_2 \in S_2 | \exists \vec{s}_1 \in S_1 : \vec{s}_1 \preceq \vec{s}_2\}|}{|S_2|}. \quad (6)$$

Note that this metric is independent of PF (P) or reference set (R). Based on pairwise comparisons between solutions, the computational complexity of $C1_R$ and $C2_R$ is $O(m|S| \cdot |R|)$, and that of $C(S_1, S_2)$ is $O(m|S_1| \cdot |S_2|)$.

B. Convergence Metrics

Convergence metrics measure the degree of proximity based on the distance between the solutions in S to those in PF (P).

The *Generational Distance* (GD) metric [6, 42] is among those commonly used in MOEAs, and has the formulation as:

$$\text{GD}(S,P) = \frac{(\sum_{i=1}^{|S|} d_i^q)^{1/q}}{|S|}, \quad (7)$$

where $d_i = \min_{\vec{p} \in P} ||F(\vec{s}_i) - F(\vec{p})||$, $\vec{s}_i \in S$ and $q = 2$. Thus, d_i is the smallest distance from $\vec{s} \in S$ to the closest solution in P. Hereafter $||\cdot||$ denotes the Euclidean distance unless explicitly indicated. Two similar metrics have been proposed as γ index [52] and M_1^* [56] with $q = 1$.

Among the *convergence* metrics, Zitzler et al. introduced the commonly used metric *ϵ-indicator* [47] as follows:

$$I_{\epsilon+}^1(S,P) = \inf_{\epsilon \in R} \{\forall \vec{p} \in P | \exists \vec{s} \in S : \vec{s} \preceq \vec{p} + \epsilon\}, \quad (8)$$

and $\vec{s} \preceq \vec{p} \cdot \epsilon$ is an alternative formulation of $\vec{s} \preceq \vec{p} + \epsilon$. In these two metrics, ϵ thus defines the value required to translate/scale the optimal solution set S such that S dominates P.

The metric *Seven Points Average Distance* (SPAD) in [57], on the other hand, considers the solution distance between S and reference set R. The formulation of SPAD is defined as:

$$\text{SPAD}(S,R) = \frac{\sum_{i=1}^{|R|} d_i}{|R|}, \quad (9)$$

where $d_i = \min_{\vec{s} \in S} ||F(\vec{r}_i) - F(\vec{s})||$, $\vec{r}_i \in R$. The seven points in R are generated as: $(0, \frac{1}{3} f_2^{max})$, $(0, \frac{2}{3} f_2^{max})$, $(0, f_2^{max})$, $(0, 0)$, $(\frac{1}{3} f_1^{max}, 0)$, $(\frac{2}{3} f_1^{max}, 0)$, $(f_1^{max}, 0)$, where f_1^{max} and f_2^{max} denote the maximum values of objective 1 and objective 2, respectively. Thus, SPAD only applies to 2-dimensional PFs. Another constraint of SPAD is that the solutions of R are strictly linearly distributed, irregardless of the shapes of PFs.

Based on the analysis of the above metrics, the time complexity of GD, γ index, M_1^* and ϵ-indicator is $O(m|S| \cdot |P|)$, and that of SPAD is $O(m|S|)$.

C. Diversity Metrics

Diversity metrics indicate the distribution and spread of solutions in the optimal solution set S. To illustrate the difference between distribution and spread, Fig. 1 showcases two representative examples. Particularly, the 5 non-dominated solutions of S in Fig. 1(a) possess good distribution but poor spread, since S does not contain the extreme points $(0, 1), (1, 0)$ of the 2-dimensional PF. On the other hand, Fig. 1(b) illustrates the example of an optimal solution set with good spread but unfavorable distribution.

1) Distribution in diversity metrics: Distribution is derived from the discrepancy of pairwise solutions in set S.

In [52], Deb et al. proposed a metric Δ' that compares all the solutions' consecutive distances with the average distance:

$$\Delta'(S) = \sum_{i=1}^{|S|-1} \frac{(d_i - \bar{d})}{|S| - 1}, \quad (10)$$

where d_i is the Euclidean distance between consecutive solutions in S, and $\bar{d}$ is the average of d_i. If all the pair of consecutive solutions share equal distance, then $d_i = \bar{d}$, $\Delta'(S) = 0$, and S has a perfect distribution. To find consecutive solutions, the prerequisite of this metric is to sort the solutions of S by lexicography order.

Two similar metrics have also been introduced in [56, 57]. The M_3^* metric [56] considers the maximum distance instead of the average distance in Δ'. In [57], the *Spacing* (SP) metric is designed as $\text{SP}(S) = \sqrt{\sum_{i=1}^{|S|} (d_i - \bar{d})^2 / (|S| - 1)}$, where $d_i = \min_{\vec{s}_j \in S, \vec{s}_j \neq \vec{s}_i} ||F(\vec{s}_i) - F(\vec{s}_j)||$ and $s_i \in S$. In contrast to the consecutive distance in Δ', metric SP calculates the closest distance of pairwise solutions in S.

In addition to the parameter-free metrics Δ', M_3^* and SP, the following two metrics are designed with user-specified parameters. The $M2^*$ metric in [56] is equipped with a niche radius σ and takes the form of:

$$M_2^*(S) = \frac{\sum_{\vec{s}_1 \in S} |\{\vec{s}_2 \in S | \ ||\vec{s}_1 - \vec{s}_2|| < \sigma\}|}{|S| - 1}. \quad (11)$$

For a solution $\vec{s}_1 \in S$, $M_2^*(S)$ measures how many solutions $\vec{s}_2 \in S$ are located in its local vicinity $||\vec{s}_1 - \vec{s}_2|| < \sigma$.

In [48], Tan et al. proposed the parameter dependent metric *Uniform Distribution* (UD) as follows:

$$\text{UD}(S) = \frac{1}{1 + D_{nc}}, \tag{12}$$

where $D_{nc} = \sqrt{\sum_{\vec{s}_i \in S}(nc(\vec{s}_i) - \overline{nc}(\vec{s}))^2/(|S|-1)}$, $nc(\vec{s}_i) = |\{\vec{s}_j \in S|\ ||\vec{s}_i - \vec{s}_j|| < \sigma\}| - 1$, and $\overline{ns}(\vec{s})$ is the average of $nc(\vec{s}_i)$. The computational complexity of Δ', M_3^*, SP, M_2^* and UD is derived as $O(m|S|^2)$.

It is worth noting that the diversity metrics presented thus far are important indicators on the distribution property of S. They however do not capture the spread characteristics of S. As shown in Fig. 1(a), a perfect distribution of solutions in S is indicated on all 5 metrics Δ', M_3^*, SP, M_2^* and UD. However, the spread property of S is low as it does not cover the PF completed (e.g., extreme points are left unexplored).

2) Spread in diversity metrics: Spread quantifies how much of the extreme regions are covered by set S.

The *Overall Pareto Spread* (OS) in [51] is proposed as:

$$\text{OS}(S, P_G, P_B) = \prod_{k=1}^{m} \frac{|\underset{\vec{s}\in S}{max}\, f_k(\vec{s}) - \underset{\vec{s}\in S}{min}\, f_k(\vec{s})|}{|f_k(P_B) - f_k(P_G)|}, \tag{13}$$

where $\underset{\vec{s}\in S}{max}\, f_k(\vec{s}), \underset{\vec{s}\in S}{min}\, f_k(\vec{s})$ are the maximum and minimum values of the k^{th} objective in S, respectively. The metric OS has a computational complexity of $O(m|S|)$. From metric OS, there exists a contradiction between distribution and spread. For instance, suppose $P_G = (0,0), P_B = (1,1)$, the spread of solutions in Fig. 1(b) is superior to that of Fig. 1(a) under the OS metric, whereas solutions of Fig. 1(b) is inferior to those of Fig. 1(a) in terms of distribution.

3) Distribution and Spread in diversity metrics: The following MOO metrics in this section consider the distribution and spread of optimal solution set S simultaneously.

The metric Δ, as introduced by Deb et al. [6], is commonly used in MOEAs. The formulation of Δ is derived as follows:

$$\Delta(S, P) = \frac{d_f + d_l + \sum_{i=1}^{|S|-1} |d_i - \bar{d}|}{d_f + d_l + (|S|-1)\bar{d}}, \tag{14}$$

where d_i is the Euclidean distance between consecutive solutions and $\bar{d}$ is the average of d_i. The terms d_f and d_l are the minimum Euclidean distances from solutions in S to the extreme (bounding) solutions of the PF (P).

The consecutive sorting in metric Δ makes it applicable to 2-dimensional PFs only. To cope with high dimensional MOPs, Zhou et al. [43] introduced the *Generalized Spread* (Δ^*) as an extension of Δ, which takes the form:

$$\Delta^*(S, P) = \frac{\sum_{k=1}^{m} d(\vec{e}_k, S) + \sum_{i=1}^{|S|} |d_i - \bar{d}|}{\sum_{k=1}^{m} d(\vec{e}_k, S) + (|S|)\bar{d}}, \tag{15}$$

where $d(\vec{e}_k, S) = \min_{\vec{s}\in S} ||F(\vec{e}_k) - F(\vec{s})||$ and $\vec{e}_k \in P$ is the extreme solutions on the k^{th} objective. Another distance in Δ^* is $d_i = \min_{\vec{s}_j \in S, \vec{s}_j \neq \vec{s}_i} ||F(\vec{s}_i) - F(\vec{s}_j)||$ to identify the closest pairwise solutions in S, and $\bar{d}$ is the average of d_i.

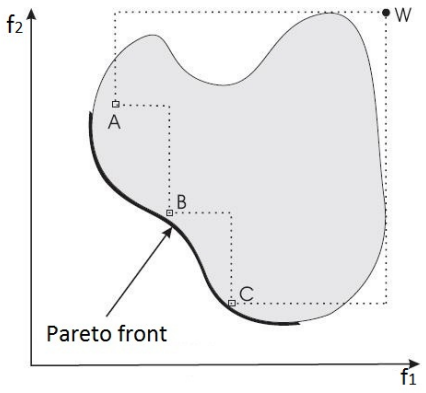


Fig. 2. The performance metric *Hypervolume* (HV) in MOO.

From Eqs. 14-15, both Δ and Δ^* share a computational complexity of $O(m|S|^2 + m|S| \cdot |P|)$. In some special cases, metric Δ^* may bring misleading information due to the closest distance calculations. In Fig. 1(b), the optimal solution set S comprises five points: $A(0,1)$, $B(0.02, 0.75)$, $C(0.58, 0.06)$, $D(0.75, 0.02)$, $E(1,0)$. Under the closest distance concept of metric Δ^*, the distance values of AB, CD would be considered twice, whereas, BC is disregarded.

In addition to those presented, other notable *diversity* metrics have also been introduced in the MOO context [51, 58, 59]. Instead of describing all of them individually, we refer the readers to the existing literature [51, 58, 59] for the details. Nevertheless, the major ideas are summarized in what follows.

- The *Entropy-based* metric [58] measures the uniformity and coverage. It employs influence functions to estimate the solution densities. The metric value is then calculated based on Shannon entropy in the discrete objective space.
- Metrics *NDC*$_\mu$ and *CL*$_\mu$ [51] divide the objective space into $(\frac{1}{\mu})^m$ grids ($\mu \in [0,1]$). The metric value is evaulated based on the number of grids containing solutions.
- Metrics σ and $\bar{\sigma}$ [59] divide the objective space into equal angles based on a set of reference lines that emanate from the origin. The metric value is the number of reference lines that are sufficiently close to the solutions of S at a predefined Euclidean distance.

A potential constraint of the above five metrics (i.e., *entropy-based* metric, *NDC*$_\mu$, *CL*$_\mu$ and σ, $\bar{\sigma}$), nonetheless, is the sensitivity to the number of grids, subregions or angles in the objective space, which are defined by the user-specified parameters $|grids|$, μ, σ, respectively.

D. Convergence–Diversity Metrics

Convergence–Diversity metrics measure the quality of the optimal solution set S in terms of convergence and diversity on a single scale.

In [44, 45], Zitzler et al. proposed the popular performance metric *Hypervolume* (HV) as:

$$\text{HV}(S, R) = volume(\bigcup_{i=1}^{|S|} v_i). \tag{16}$$

TABLE I
A CLASSIFICATION OF PERFORMANCE METRICS IN MULTIOBJECTIVE OPTIMIZATION (MOO).

Performance criteria		Performance metrics	Parameter requirements	Comparison set	Computational compleixity
Capacity		ONVG, GNVG, NVA	None	None	Low
		ONVGR, GNVGR, ER	None	Pareto front (P)	Low
		$C1_R$, $C2_R$	None	Reference set (R)	$O(m\|S\| \cdot \|R\|)$
		Metric C	None	Optimal solution set (S)	$O(m\|S_1\| \cdot \|S_2\|)$
Convergence		GD, ϵ-indicator, γ index, M_1^*	None	Pareto front (P)	$O(m\|S\| \cdot \|P\|)$
		SPAD	None	Pareto front (R)	$O(m\|S\|)$
Diversity	Distribution	Δ', M_3^*, SP	None	None	$O(m\|S\|^2)$
		M_2^*, UD	Niche radius (σ)	None	$O(m\|S\|^2)$
	Spread	OS	None	Good, Bad points (P_G, P_B)	$O(m\|S\|)$
	Distribution–Spread	Δ, Δ^*	None	Pareto front (P)	$O(m\|S\|^2 + m\|S\| \cdot \|P\|)$
		Entropy, NDC$_\mu$, CL$_\mu$, σ, $\bar{\sigma}$	$\|grids\|$, μ, $\|lines\|$	None	High
Convergence–Diversity		HV, HD, HR, Metric S, Metric D	None	Reference set (R)	$O(\|S\|^{m-1})$
		IGD, Δ_q, MPFE	None	Pareto front (P)	$O(m\|S\| \cdot \|P\|)$

This gives the volume (in the objective space) that is dominated by the optimal solution set S. Using the example in Fig. 2, where $S = \{A, B, C\}$ is attained when minimizing a bi-objective MOP. The HV(S, R) is thus the area $ABCWA$ enclosed by the discontinuous boundary, where reference set $R = \{W\}$[2]. Mathematically, for each solution $\vec{s}_i \in S$, a hypercube v_i is constructed with the reference set and the solution $\vec{s}_i$ as the diagonal corners of the hypercube.

Hypervolume (HV) is a set quality measure that is strictly monotonic with regards to the Pareto dominance concept [60, 61]. It quantifies and encapsulates both the convergence and diversity information of S. In particular, the closer are the solutions of S to the true PFs, the larger is the value of HV. At the same time, a higher HV could also indicate solutions of S are scattered more evenly in the objective space.

Other metrics adopting the similar concept have also been introduced in the MOO context, such as *Hypervolume Difference* (HD) [51], *Hypervolume Ratio* (HR), Metric S [44], *Coverage Difference of Two Sets* or Metric D [62].

The downside of HV lies in the high computational complexity of $O(|S|^{m-1})$, which can become intractable when the number of objectives is large (e.g., $m \geq 4$). Although some research efforts have been devoted to reduce the computational burden, such as Monte Carlo sampling [60, 63], the accuracy of HV is however compromised. For the fast approaches on calculating the exact HV, the reader is referred to [61, 64–69].

The metric *Inverted Generational Distance* (IGD) in [9, 42, 47], which has a similar formulation to GD, is introduced as:

$$\text{IGD}(P, S) = \frac{(\sum_{i=1}^{|P|} d_i^q)^{1/q}}{|P|}, \tag{17}$$

where $d_i = \min_{\vec{s} \in S} ||F(\vec{p}_i) - F(\vec{s})||$, $\vec{p}_i \in P$, $q = 2$, and d_i is the smallest distance of $\vec{p} \in P$ to the closest solutions in S.

The metric *Averaged Hausdorff Distance* (Δ_q) [70], which combines the properties of GD and IGD, is defined as:

$$\Delta_q(S, P) = \max(\text{GD}(S, P), \text{IGD}(P, S)). \tag{18}$$

The parameter q in Δ_q is a positive integer. A larger q value penalizes the outlier solutions in S more. Both IGD and Δ_q share a computational complexity of $O(m|S| \cdot |P|)$.

[2]Point W can be simply attained by constructing a vector of worst objective function values. Meanwhile, the number of solutions in R is not limit to one.

Instead of measuring the average distance in IGD, the *Maximum Pareto Front Error* (MPFE) [57] is defined as:

$$\text{MPFE}(P, S) = \max_{\vec{p} \in P} \sqrt{\min_{\vec{s} \in S} \sum_{k=1}^{m} |f_k(\vec{s}) - f_k(\vec{p})|^2}. \tag{19}$$

This metric finds the maximum distance from solutions in P to the closest solution in S. In contrast to HV and IGD, MPFE focuses on the extreme solutions rather than the entire solution set. In particular, when some outliers exist in S, MPFE is inclined to show the worst value of S and ignores the information of other high quality solutions.

To summarize, Table I classifies the major MOO metrics proposed to date in terms of performance criteria, parameter requirements, comparison set and computational complexity. In what follows, we highlight and analyze the potential issues of some of these metrics in Section IV. The empirical investigation on the relationships among representative group metrics, with special focus on the consistencies or contradictions on PFs of diverse geometrical shapes, is then presented in Section V.

IV. INSIGHTS ON EXTREME CASES OF MOO METRICS

In this section, we highlight and analyse the inadequacies of MOO metrics. In particular, on extreme cases, some MOO metrics are noted to omit partial information that describes the true quality of the optimal solution set S.

The *Capacity* metrics focus on tallying the number of solutions in the optimal solution set S. Hence, they are not designed to provide the convergence and diversity information of S. For instance, two optimal solution sets S_1 and S_2 of Fig. 3 are obtained by minimizing a bi-objective MOP. Notably, S_2 is better than S_1 in terms of convergence, since the solutions of S_2 approach the origin $(0, 0)$ closer than those of S_1. However, under Metric C, both of them share a common value of $C(S_1, S_2) = C(S_2, S_1) = 0.5$. In MOO, *capacity* metric is generally used as the basic criterion for assessing the optimal solution sets [4, 5]. If two optimal solution sets obtain the same number of solutions, we can then compare the convergence and diversity information in a fair and detail manner.

With regards to the *convergence* metrics, two critical issues can be asserted. First of all, most of them (e.g., GD and $I_{\epsilon+}^1$)

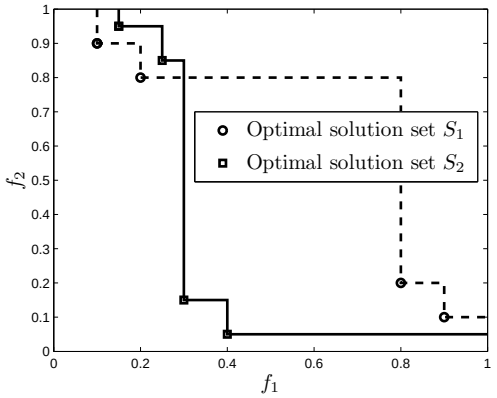


Fig. 3. The inadequacy of Metric C, $C(S_1, S_2) = C(S_2, S_1) = 0.5$ [46].

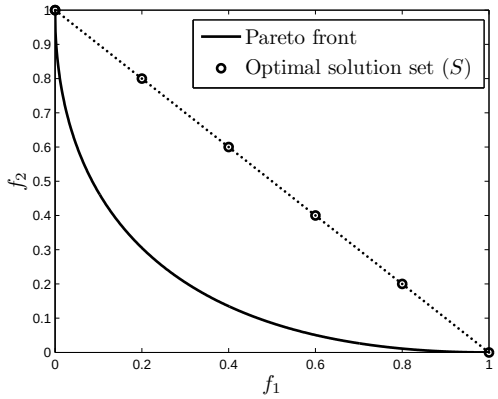


Fig. 4. The inadequacy of Δ and Δ^*, $\Delta(S, P) = \Delta^*(S, P) = 0$.

require the PF (P) to be predefined. In practice, however, it is almost impossible to know the true PFs, especially when no prior knowledge about the problem is available. On the other hand, even if the geometrical characteristics of PFs is known, constructing P for many objectives, discrete and asymmetric PFs, is non-trivial. Secondly, *convergence* metrics often omit the diversity property of S. Taking Fig. 1 as the illustration example, which comprises two optimal solution sets with perfect convergence (i.e., GD$(S, P) = 0$), but both display poor diversity.

Among the *diversity* metrics, only a few of them (i.e., Δ and Δ^*) are designed with comparison sets, e.g., PF (P) or reference set R. In addition, Δ and Δ^* only consider the extreme solutions of PF (P), e.g., d_f and d_l in Eq. 14. Hence, a large number of solutions in P are not used in the comparison with solutions in S. In Fig. 4, the solutions in S are scattered evenly along the linear dash line, and $\Delta(S, P) = \Delta^*(S, P) = 0$. This means that S has a perfect diversity. However, they do not assert whether the solutions had converged to the true PFs.

On the other hand, the *convergence–diversity* metrics measure two types of information on a single scale. As mentioned in Section III-D, the high computational complexity makes it cumbersome to apply HV in MOO. In addition, the potential limitation of the other three metrics (i.e., IGD, Δ_q and MPFE) lies in the need to construct the comparison sets. In HV, the comparison set is the reference set R. As shown in Fig. 2, R can be constructed with ease using single point W. However, the comparison set for IGD, Δ_q and MPFE is P, which is the set of representative solutions on PF. For instance, a popular method [4, 5] is to divide the objectives evenly and sample a large number of solutions on the true PF to form P. In Fig. 4, set P of $10,000$ points can be obtained by dividing $f_1 \in [0, 1]$ into $10,000$ segments equally. However, this results in unevenly distributed solutions, where nearly 5,000 points are crowded in ($f_1 \in [0.5, 1]$, $f_2 \in$ [**0, 0.1**]) when the PF is convex. Thus, this method is insufficient to approximate the true PFs. For these metrics, i.e., IGD, Δ_q and MPFE, where the reliance on set P is high, the ability to construct well scattered solutions of P is crucial for assessing the true quality of set S.

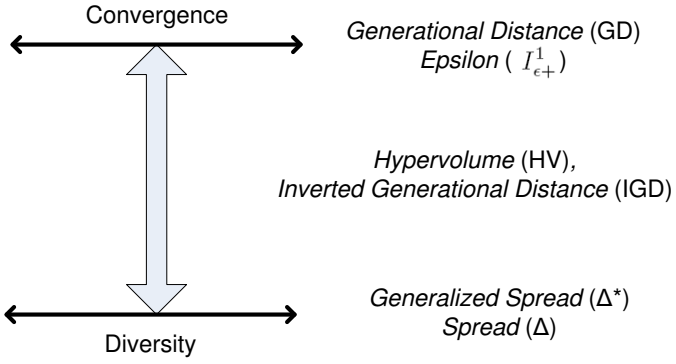


Fig. 5. Relationships of GD, $I^1_{\epsilon+}$, Δ, Δ^*, IGD and HV as introduced in [4].

V. Empirical Studies and Analyses

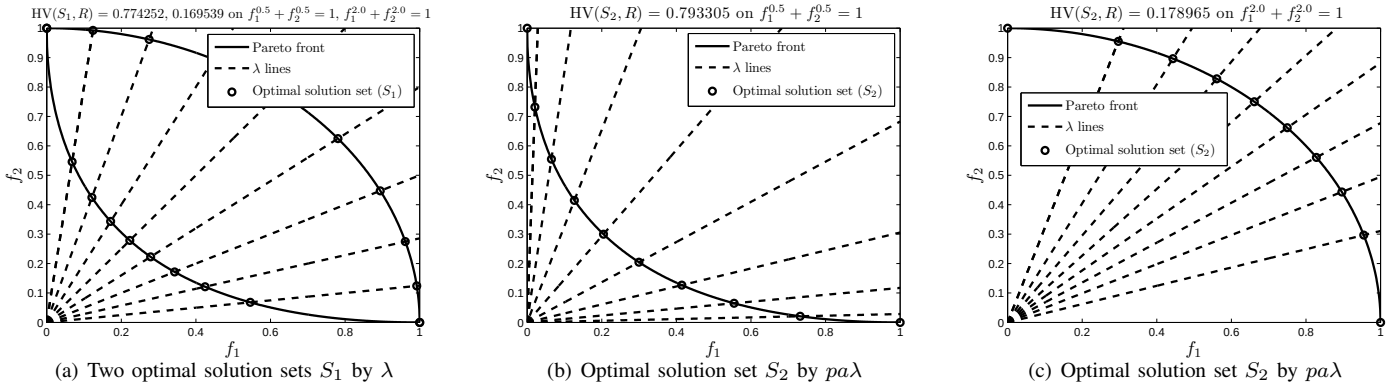
With the categorization of MOO metrics based on the performance criteria, namely *capacity*, *convergence*, *diversity*, *convergence–diversity*, in this section, a systematic investigation on the relationships among metrics is presented. After constructing various Pareto fronts (PFs) and reference sets, six representative metrics are investigated on two optimal solution sets. In the experiments, we used hypervolume (HV) [4, 5, 16, 60], which is the most widely accepted metric in the MOO community, as the baseline to study the relationships among metrics on different geometrical shapes of PFs.

Two representative metrics from each of the following categories are considered in the study[3]:

- *Convergence* metrics: GD, $I^1_{\epsilon+}$;
- *Diversity* metrics: Δ, Δ^*;
- *Convergence–Diversity* metrics: IGD, HV.

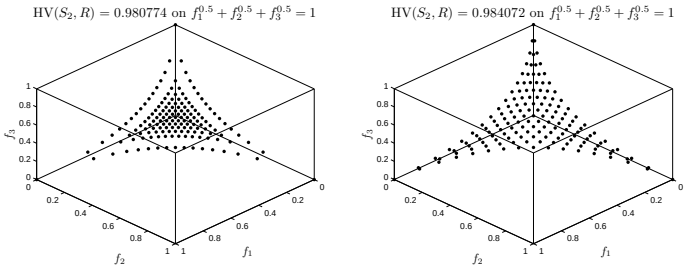
In general, an optimal solution set with small GD, $I^1_{\epsilon+}$, Δ, Δ^*, IGD and large HV is desirable. In the literature [4, 5], the relationships of six metrics were intuitively and briefly discussed (As shown in Fig. 5). In this paper, we not only validate the relationships, but also provide the details on the potential consistencies and contradictions among the six metrics when the PFs are convex and concave.

[3]Since *capacity* metrics only provide cardinality information, it does not serve as meaningful when used alone to assess the optimal solution set.



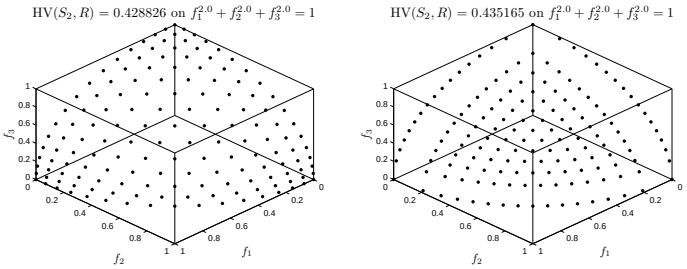
(a) Two optimal solution sets S_1 by λ (b) Optimal solution set S_2 by $pa\lambda$ (c) Optimal solution set S_2 by $pa\lambda$

Fig. 6. 10 intersection points along 2-dimensional PFs. (a) HV(S_1, R) = 0.774252, 0.169539 by λ method [8, 9] on convex and concave PFs, respectively, (b) HV(S_2, R) = 0.793305 by $pa\lambda$ method [16] on convex PF $f_1^{0.5} + f_2^{0.5} = 1$, (c) HV($S_2, R$) = 0.178965 by $pa\lambda$ method on concave PF $f_1^{2.0} + f_2^{2.0} = 1$.



(a) Optimal solution set S_1 by λ (b) Optimal solution set S_2 by $pa\lambda$

Fig. 7. 153 points along 3-dimensional convex PF $f_1^{0.5} + f_2^{0.5} + f_3^{0.5} = 1$. (a) HV($S_1, R$) = 0.980744 by λ method [8, 9], (b) HV(S_2, R) = 0.984072 by $pa\lambda$ method [16].



(a) Optimal solution set S_1 by λ (b) Optimal solution set S_2 by $pa\lambda$

Fig. 8. 153 points along 3-dimensional concave PF $f_1^{2.0} + f_2^{2.0} + f_3^{2.0} = 1$. (a) HV($S_1, R$) = 0.428826 by λ method [8, 9], (b) HV(S_2, R) = 0.435165 by $pa\lambda$ method [16].

A. PFs and Reference Sets Design

In general, the true PFs have many different shapes and forms, which can be convex, concave, discrete, no-differential, multi-modal, asymmetric, etc [1–5]. Instead of considering all forms of PFs, our interest in the current paper is on the symmetric and continuous PFs. Without loss of generality, we consider the true PFs as

$$f_1^p + f_2^p + \cdots + f_m^p = 1, \tag{20}$$

where the objectives are normalized in the range $[0, 1]$ and $p \in (0, \infty)$ is the parameter to control the geometrical shapes of PFs. As shown in Fig. 6(b), when $p = 0.5$, the PF is convex in the 2-dimensional objective space. On the other hand, the PF is 2-dimensional concave when $p = 2$ in Fig. 6(c).

To obtain P, a number of non-dominated solutions need to be selected to approximate the true PFs. A simple approach for constructing P is to adopt the popular λ method [8, 9]. Let $\lambda = (\lambda_1, \cdots, \lambda_m)^T$ be a weight vector, where $\lambda_i \geq 0$ and $\sum_{i=1}^m \lambda_i = 1$. As shown in Fig. 6, each weight vector defines a λ line of predefined gradient. All the weight vectors then take values from the following set:

$$\{\frac{0}{H}, \frac{1}{H}, \cdots, \frac{H}{H}\},$$

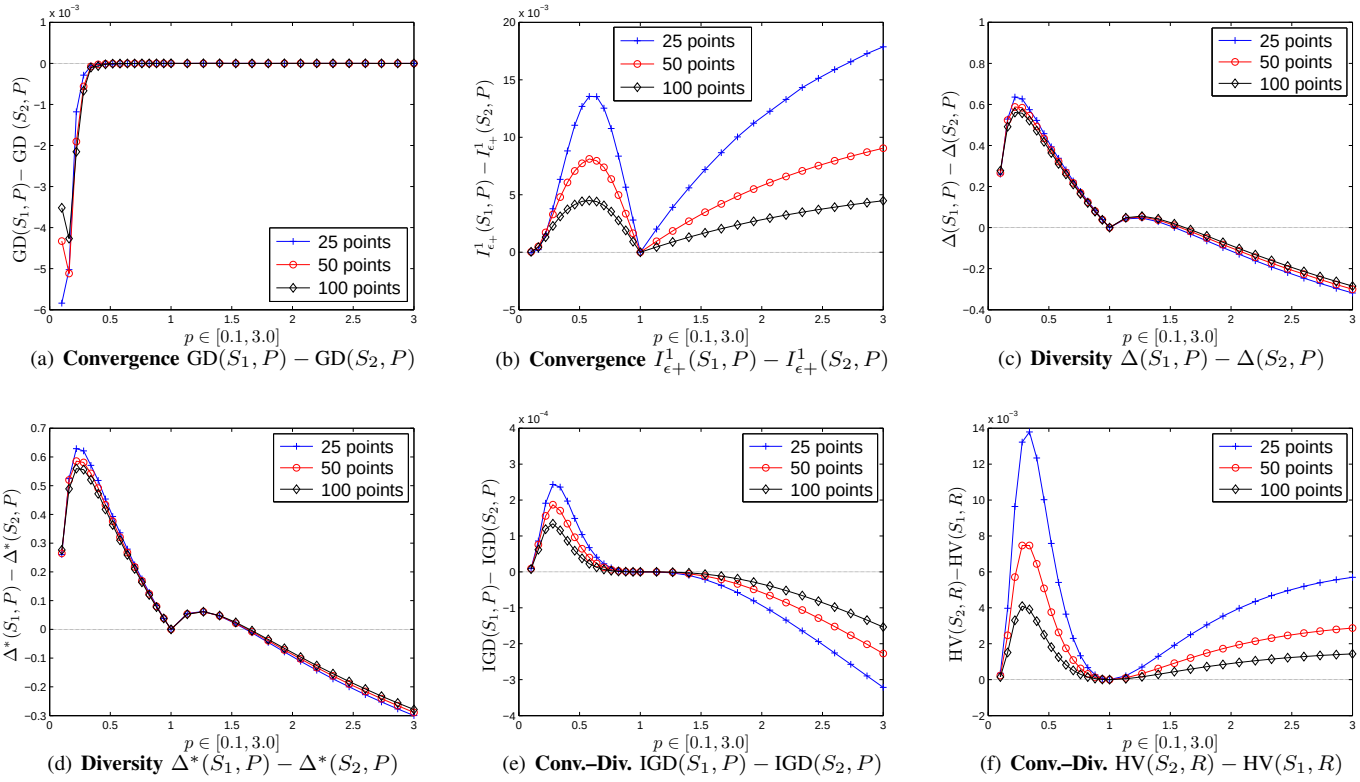
where H is a control parameter and the number of weight vectors is C_{H+m-1}^{m-1}. Every λ line meets the PF and each intersect point denotes a solution. In this case, the number of points in P is thus $|P| = C_{H+m-1}^{m-1}$. Taking Fig. 6 as an illustration example, we set $H = 9, m = 2$, and the $C_{10}^1 = 10$ weight vectors can be derived as $\{(\frac{0}{9}, \frac{9}{9}), (\frac{1}{9}, \frac{8}{9}) \cdots, (\frac{9}{9}, \frac{0}{9})\}$. The 10 solutions are then obtained by solving Eq. 20 based on the 10 weight vectors.

In MOO, set P typically comprises many non-dominated solutions to approximate the true PFs at reasonable accuracies [4, 5]. In the experimental studies, the parameters are set as $H = 10,000$ and $|P| = 10,001$ on the bi-objective PFs. Their values are $H = 140$ and $|P| = 10,011$ on the tri-objective PFs. They are $H = 38, |P| = 10,660$ on the quad-objective PFs. In comparison, the reference set R is easy to construct, which is generated as $R = \{(1, 1)\}, \{(1, 1, 1)\}, \{(1, 1, 1, 1)\}$ on 2, 3, 4-dimensional PFs, respectively.

B. Optimal Solution Sets

The goal in MOO is to find the optimal solution set S with both good convergence and diversity. The special case of solutions in S that are not on PFs (Called *poor convergence*) will not be considered in the present study. As shown in Fig. 4, it may not make good sense to evaluate the diversity of S alone if it does not satisfy the basic requirement of convergence.

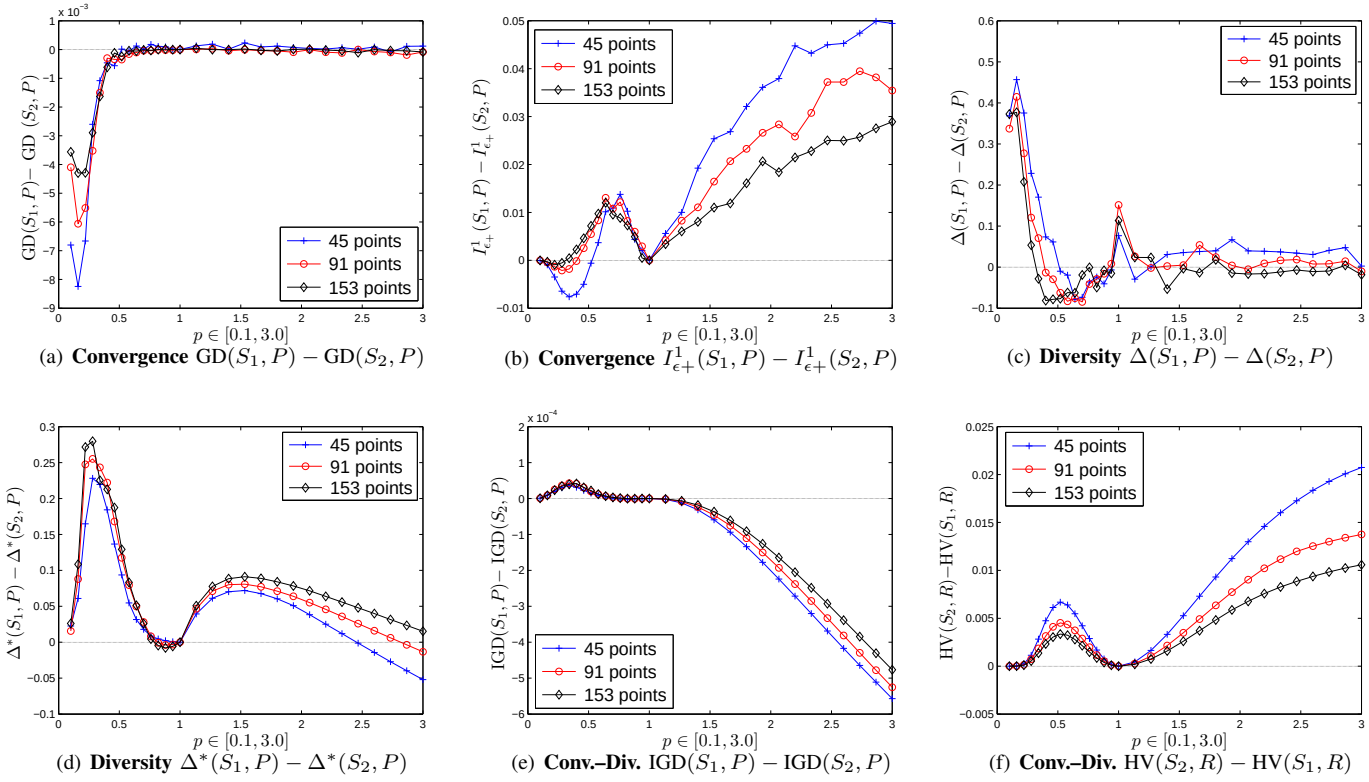
In this paper, two optimal solution sets (S_1 and S_2) are constructed on the true PFs (Called *good convergence*) with different diversities. The set S_1 is attained based on the λ method [8, 9]. The Pareto-adaptive weight vectors ($pa\lambda$ method [16]) is adopted to maximize HV and generate S_2.



(a) **Convergence** $\mathrm{GD}(S_1,P) - \mathrm{GD}(S_2,P)$ (b) **Convergence** $I^1_{\epsilon+}(S_1,P) - I^1_{\epsilon+}(S_2,P)$ (c) **Diversity** $\Delta(S_1,P) - \Delta(S_2,P)$

(d) **Diversity** $\Delta^*(S_1,P) - \Delta^*(S_2,P)$ (e) **Conv.–Div.** $\mathrm{IGD}(S_1,P) - \mathrm{IGD}(S_2,P)$ (f) **Conv.–Div.** $\mathrm{HV}(S_2,R) - \mathrm{HV}(S_1,R)$

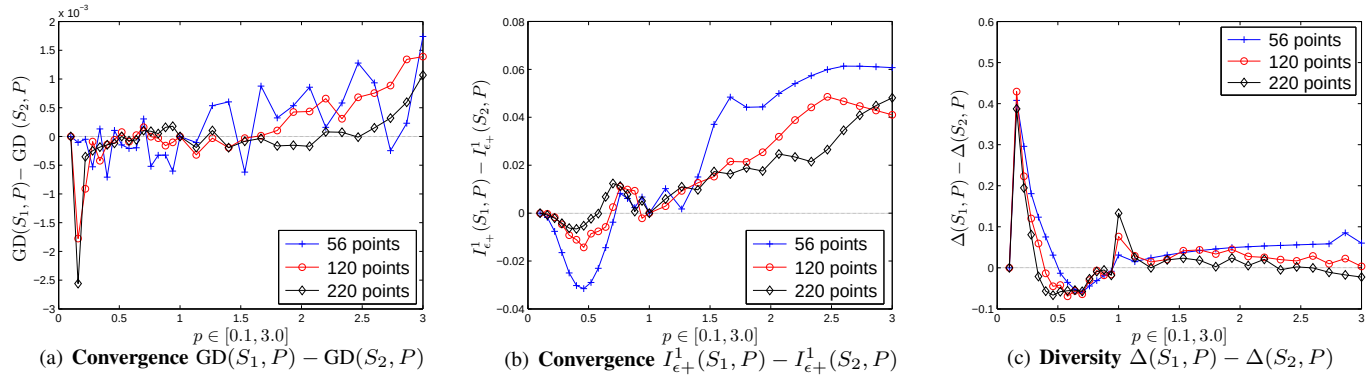
Fig. 9. 2-dimensional PFs: Differences in the metrics of S_1 and S_2 for cases of 25, 50 or 100 points along $f_1^p + f_2^p = 1, p \in [0.1, 3.0]$. The optimal solution sets S_1 and S_2 are obtained based on λ [8, 9] and $pa\lambda$ [16], respectively. A small GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD while large HV is desirable.



(a) **Convergence** $\mathrm{GD}(S_1,P) - \mathrm{GD}(S_2,P)$ (b) **Convergence** $I^1_{\epsilon+}(S_1,P) - I^1_{\epsilon+}(S_2,P)$ (c) **Diversity** $\Delta(S_1,P) - \Delta(S_2,P)$

(d) **Diversity** $\Delta^*(S_1,P) - \Delta^*(S_2,P)$ (e) **Conv.–Div.** $\mathrm{IGD}(S_1,P) - \mathrm{IGD}(S_2,P)$ (f) **Conv.–Div.** $\mathrm{HV}(S_2,R) - \mathrm{HV}(S_1,R)$

Fig. 10. 3-dimensional PFs: Differences in the metrics of S_1 and S_2 for cases of 45, 91 or 153 points along $f_1^p + f_2^p + f_3^p = 1, p \in [0.1, 3.0]$. The optimal solution sets S_1 and S_2 are obtained based on λ [8, 9] and $pa\lambda$ [16], respectively. A small GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD while large HV is desirable.



(a) **Convergence** $\mathrm{GD}(S_1,P) - \mathrm{GD}(S_2,P)$ (b) **Convergence** $I^1_{\epsilon+}(S_1,P) - I^1_{\epsilon+}(S_2,P)$ (c) **Diversity** $\Delta(S_1,P) - \Delta(S_2,P)$

(d) **Diversity** $\Delta^*(S_1,P) - \Delta^*(S_2,P)$ (e) **Conv.–Div.** $\mathrm{IGD}(S_1,P) - \mathrm{IGD}(S_2,P)$ (f) **Conv.–Div.** $\mathrm{HV}(S_2,R) - \mathrm{HV}(S_1,R)$

Fig. 11. 4-dimensional PFs: Differences in the metrics of S_1 and S_2 for cases of 56, 120 or 220 points along $f_1^p + f_2^p + f_3^p + f_4^p = 1, p \in [0.1, 3.0]$. The optimal solution sets S_1 and S_2 are obtained based on λ [8, 9] and $pa\lambda$ [16], respectively. A small GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD while large HV is desirable.

In Fig. 6(a), the optimal solution set S_1 has $\mathrm{HV}(S_1,R)$ = 0.774252 when the PF is $f_1^{0.5} + f_2^{0.5} = 1$, and $\mathrm{HV}(S_1,R)$ = 0.169539 for $f_1^2 + f_2^2 = 1$. In Fig. 6(b), the hypervolume of S_2 is $\mathrm{HV}(S_2,R)$ = 0.793305 on the convex PF. The solutions of S_2 are distributed at near the two extreme points of PF and scattered more evenly than those of S_1. On the other hand, Fig. 6(c) depicts $\mathrm{HV}(S_2,R)$ = 0.178965 on the concave PF. The solutions of S_2 are close to the median of PF and exhibit a better distribution than those of S_1 in the central PF.

Figs. 7-8 depict two optimal solution sets (S_1 and S_2) on 3-dimensional PFs, which are also constructed based on the λ method [8, 9] and $pa\lambda$ method [16], respectively. The convex PF ($f_1^{0.5} + f_2^{0.5} + f_3^{0.5} = 1$) of Fig. 7 has $\mathrm{HV}(S_1,R)$ = 0.980774 and $\mathrm{HV}(S_2,R)$ = 0.984072. The solutions of S_2 in Fig. 7(b) are noted to be more evenly spread than those of S_1 in Fig. 7(a) on all three objectives. On the other hand, for the concave PF ($f_1^{2.0} + f_2^{2.0} + f_3^{2.0} = 1$) of Fig. 8, the hypervolumes are $\mathrm{HV}(S_1,R)$ = 0.428826 and $\mathrm{HV}(S_2,R)$ = 0.435165. Solutions of S_1 in Fig. 8(a) are crowded near the three extreme points $\{(1,0,0),(0,1,0),(1,0,0)\}$, while they are sparse in the central region of PF. Furthermore, in contrast to S_1, the solutions of S_2 in Fig. 8(b) are crowded near the center of the PF $(\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}})$.

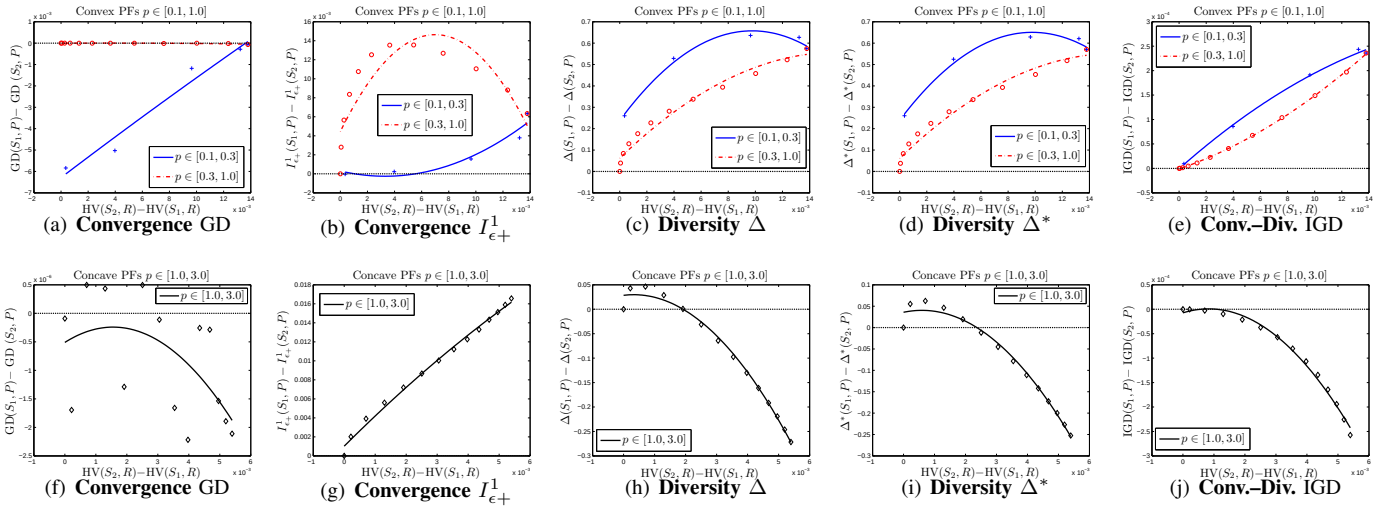
C. Performance Metric Studies on 2-dimensional PFs

Upon constructing the set P, which represents the true PFs (Section V-A), we can then proceed to assess and compare the two optimal solution sets S_1 and S_2 (Section V-B) using the six representative metrics, first on 2-dimensional PFs. The member size of the optimal solution set is configured as $|S_1| = |S_2| \in \{25, 50, 100\}$ and the formulation of PFs is $f_1^p + f_2^p = 1$, where $p \in [0.1, 3.0]$.

In Fig. 9, the differences among the six MOO metrics for S_1 and S_2 on 2-dimensional PFs are presented. The values above the origin in Fig. 9 indicate that S_2 is superior to S_1. Fig. 9(a) depicts the differences between S_1 and S_2 on metric GD. Note that $\mathrm{GD}(S_2,P) - \mathrm{GD}(S_1,P)$ is near to zero when $p \in [0.3, 3.0]$. This implies that GD could not uncover the differences between the two optimal solution sets, which are revealed by the $I^1_{\epsilon+}$ metric of Fig. 9(b). The result of $I^1_{\epsilon+}$ is quite similar to HV in Fig. 9(f), where their curves exhibit similar trends, despite at the different scales. As the solution number of S increases from 25 to 50 and 100, $|I^1_{\epsilon+}(S_2,P) - I^1_{\epsilon+}(S_1,P)|$ becomes smaller. The reason is that a small ϵ is need to make S_1 and S_2 dominate the PFs (P) by Eq. 8 when the solution number of S is large.

The two plots, Figs. 9(c)-(d), summarize the differences between S_1 and S_2 on the *diversity* metrics (Δ and Δ^*). Notably, both metrics arrive at similar results. The optimal solution set S_2 exhibits superior diversity over S_1 when $p \in [0.1, 1.5)$ (i.e., $\Delta(S_2,P) < \Delta(S_1,P)$ and $\Delta^*(S_2,P) < \Delta^*(S_1,P)$), whereas, S_2 is inferior to S_1 when $p \in [1.5, 3.0]$. As shown in Fig. 6, solutions in S_2 are scattered on convex PFs, however, they tend to congregate together when the PF is concave. Overall, both Δ and Δ^* share similar results with HV on convex PFs, but are in conflict with HV on concave PFs.

Taking focus on the *convergence–diversity* metrics (IGD and



(a) **Convergence** GD (b) **Convergence** $I^1_{\epsilon+}$ (c) **Diversity** Δ (d) **Diversity** Δ^* (e) **Conv.–Div.** IGD

(f) **Convergence** GD (g) **Convergence** $I^1_{\epsilon+}$ (h) **Diversity** Δ (i) **Diversity** Δ^* (j) **Conv.–Div.** IGD

Fig. 12. The relationships of five metrics (GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD) to the baseline HV on 2-dimensional convex PFs (a-e) and concave PFs (f-j). The $I^1_{\epsilon+}$, Δ, Δ^* and IGD are consistent with HV on 2-dimensional convex PFs, whereas, Δ , Δ^* and IGD contradict with HV on 2-dimensional concave PFs.

HV) in Figs. 9(e)-(f), both S_1 and S_2 share the same scattered solutions and $\mathrm{HV}(S_2, R) = \mathrm{HV}(S_1, R)$ when $p = 1$ (i.e., the PF is a straight line). When $p \neq 1$, S_2 displays better HV than S_1, i.e., $\mathrm{HV}(S_2, R) > \mathrm{HV}(S_1, R)$. In Fig. 9(e), the IGD value of S_2 is superior to S_1 when $p \in [0.1, 1.5)$, but S_2 becomes inferior to S_1 under the same IGD metric when $p \in [1.5, 3.0]$. This indicates that IGD shares similar estimations with HV on convex PFs, but displays conflicting information to HV on concave PFs.

To summarize, from the studies of the MOO metrics on 2-dimensional PFs, metric GD is found to be unable to provide diversity information on S_1 and S_2. The results of metric $I^1_{\epsilon+}$, on the other hand, is similar to those of HV. Other three metrics Δ, Δ^* and IGD share two characteristics: they display similar trends with HV on 2-dimensional convex PFs, but in some special cases (i.e., $\{f_1^p + f_2^p = 1 | 1.5 \leq p \leq 3\}$), they are in conflict with HV on 2-dimensional concave PFs.

D. Performance Metric Studies on 3-dimensional PFs

In this section, we extend the MOO metric studies to 3-dimensional PFs. The parameter H is set as $H \in \{8, 12, 16\}$ and the number of solutions in S_1, S_2 are thus $C^2_{H+2} = \{45, 91, 153\}$, respectively. The 3-dimensional PFs are formulated as $f_1^p + f_2^p + f_3^p = 1$, where $p \in [0.1, 3.0]$. Figs. 10(a)-(f) present the assessments of the optimal solution sets S_1 and S_2 on the six representative metrics, i.e., GD, $I^1_{\epsilon+}$, Δ, Δ^*, IGD and HV, respectively.

Fig. 10(a) shows similar GD values on S_1 and S_2 when $p \in [0.3, 3.0]$, despite the two sets of solutions being unique to one other. This indicates that metric GD cannot distinguish between S_1 and S_2 on their diversities. With respect to metric $I^1_{\epsilon+}$ in Fig. 10(b), it displays similar shapes to HV in Fig. 10(f), except for $p \in [0.1, 0.5]$. This means that an optimal solution set with a better HV will likely also imply a better $I^1_{\epsilon+}$. From the results in Fig. 10(c), metric Δ does not display a stable trend, since its consecutive sorting procedure only works well on 2-dimensional PFs (See Section III-C3). Figs. 10(d)-(e) show that the two metrics, Δ^* and IGD, exhibit similar trends, although at different scales. In particular, both Δ^* and IGD share similar estimations with HV on the 3-dimensional convex PFs. However, when the PFs are 3-dimensional concave, i.e., $\{f_1^p + f_2^p + f_3^p = 1 | 1.5 \leq p \leq 3\}$, both Δ^* and IGD are observed to be in conflict with HV.

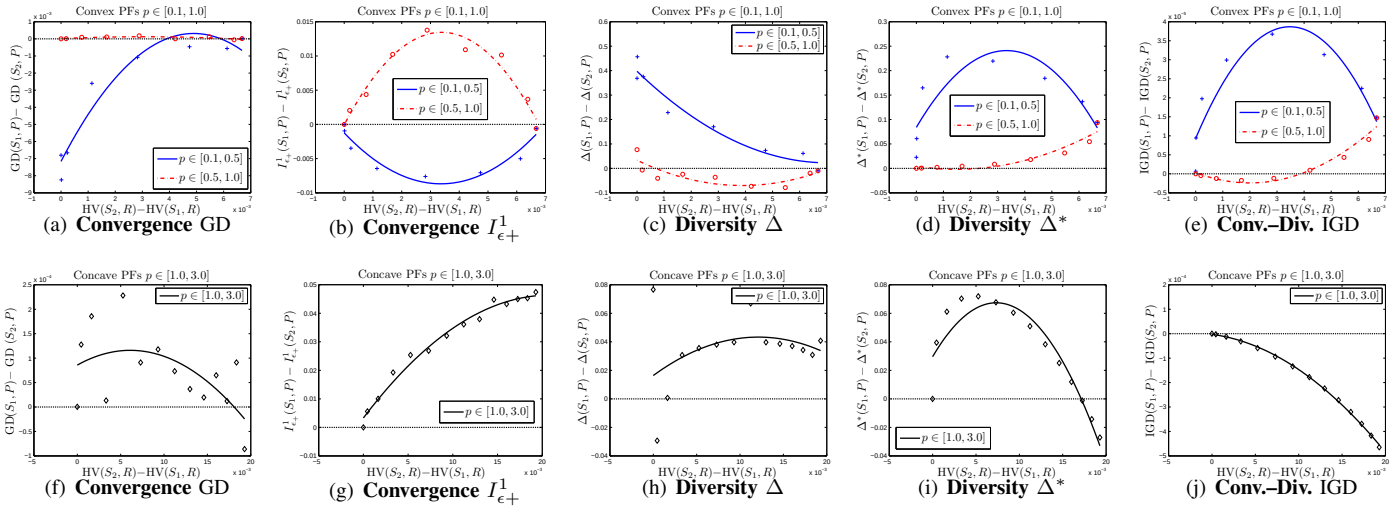
E. Performance Metric Studies on 4-dimensional PFs

In this section, we proceed further to the study of the six MOO metrics on 4-dimensional PFs. The parameter H is configured as $H \in \{5, 7, 9\}$ and the number of solutions in S_1, S_2 are $C^3_{H+3} = \{56, 120, 220\}$, respectively. The 4-dimensional PFs are formulated as $f_1^p + f_2^p + f_3^p + f_4^p = 1$, where $p \in [0.1, 3.0]$. Figs. 11(a)-(f) show the measurements of the optimal solution sets S_1 and S_2 on the six representative metrics, i.e., GD, $I^1_{\epsilon+}$, Δ, Δ^*, IGD and HV, respectively.

The results in Figs. 11(a, c) do not display clear trends on the two metrics GD and Δ, which is similar to the results on 2, 3-dimensional PFs. Hence we are unable to draw any concrete conclusions on metrics GD and Δ. On the other hand, the results of metric $I^1_{\epsilon+}$ in Fig. 11(b) shows similar trends to HV in Fig. 11(f), except for $p \in [0.1, 0.7]$. This indicates that an optimal solution set with a large HV will likely also report a small $I^1_{\epsilon+}$. From the results of Fig. 11(d), metric Δ^* is in conflict with HV on the 4-dimensional convex PFs ($p \in [0.1, 1.0)$), whereas it shares a similar estimation with HV on the 4-dimensional concave PFs ($p \in [1.0, 3.0]$). Last but not least, metric IGD in Fig. 11(e) exhibits similar and conflicting trends to HV on the 4-dimensional convex and concave PFs, respectively.

F. Relationship of MOO Metrics

In this section, we analyse the relationships of the six representative metrics statistically and systematically. In particular, we discuss the case where $|S_1| = |S_2| = 25, 45, 56$ for 2, 3, 4-dimensional PFs, respectively.



(a) **Convergence** GD (b) **Convergence** $I^1_{\epsilon+}$ (c) **Diversity** Δ (d) **Diversity** Δ^* (e) **Conv.–Div.** IGD

(f) **Convergence** GD (g) **Convergence** $I^1_{\epsilon+}$ (h) **Diversity** Δ (i) **Diversity** Δ^* (j) **Conv.–Div.** IGD

Fig. 13. The relationships of five metrics (GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD) to the baseline HV on 3-dimensional convex PFs (a-e) and concave PFs (f-j). The $I^1_{\epsilon+}$, Δ^* and IGD are consistent with HV on 3-dimensional convex PFs, whereas, Δ^* and IGD contradict with HV on 3-dimensional concave PFs.

Suppose two metrics M_1 and M_2 are considered for assessing the two optimal solution sets S_1 and S_2 (See Section V-B), the following concepts define the *Consistency* and *Contradiction* relationships between M_1 and M_2.

- **Consistency**:
  1) $S_1 \prec S_2$ on the two metrics (M_1, M_2);
  2) $\mathrm{Corr}(M_1(S_1) - M_1(S_2), M_2(S_1) - M_2(S_2)) > 0$.

  In the first condition[4], S_1 dominates S_2, which means that S_1 reports better results than S_2 on both metrics (M_1, M_2). The second condition[5] reveals a positive correlation on the metric value differences.
- **Contradiction**:
  1) $S_1 \not\prec S_2$ on the two metrics (M_1, M_2);
  2) $\mathrm{Corr}(M_1(S_1) - M_1(S_2), M_2(S_1) - M_2(S_2)) < 0$.

  In the first condition[6], S_1 and S_2 are non-dominated. This implies that S_1 is better than S_2 on one metric, but worse on the other metric. The second condition indicates a negative correlation on the metric value differences.

Figs. 12-14 show the relationships of the five metrics (GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD) with the HV as baseline, on 2, 3, 4-dimensional PFs, respectively. Since a large HV and small GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD are desirable, we set the x-axis as $\mathrm{HV}(S_2) - \mathrm{HV}(S_1)$ and the y-axis is set in a reverse order for the other five metrics (e.g., $\mathrm{GD}(S_1) - \mathrm{GD}(S_2)$).

In Figs. 12(a-e), for the 2-dimensional convex PFs $\{f_1^p + f_2^p = 1 | 0.1 \leq p \leq 1\}$, the five metrics show positive correlations to the baseline HV, except for GD with $p \in [0.3, 1.0]$ and $I^1_{\epsilon+}$ with $p \in [0.3, 0.6]$. For instance, when $p \in [0.1, 0.3]$, the correlation coefficients of GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD to HV are 0.986, 0.889, 0.831, 0.985, 0.994, respectively. The curve of GD in Fig. 12(a), which is below the origin, indicates a conflict. This indicates that an optimal solution set with better HV (e.g., $\mathrm{HV}(S_2) > \mathrm{HV}(S_1)$) will perform poorer on GD (e.g., $\mathrm{GD}(S_2) > \mathrm{GD}(S_1)$). In addition, $I^1_{\epsilon+}$, Δ, Δ^* and IGD are consistent with HV on the 2-dimensional convex PFs, since their curves are above the origin and they are positively correlated to HV. A special case is that $I^1_{\epsilon+}$ has a negative correlation to HV when $p \in [0.3, 0.6]$, i.e., $I^1_{\epsilon+}(S_1) - I^1_{\epsilon+}(S_2)$ decreases when $\mathrm{HV}(S_2) - \mathrm{HV}(S_1)$ increases in Fig. 12(b).

Figs. 12(f-j) showcase the relationships of the five metrics to the baseline HV on 2-dimensional concave PFs $\{f_1^p + f_2^p = 1 | 1 \leq p \leq 3\}$. Fig. 12(g) shows that $I^1_{\epsilon+}$ is consistent with HV, whereas, IGD contradicts with HV in Fig. 12(j). In particular, the correlation coefficient of $I^1_{\epsilon+}$ to HV is 0.996, and that of IGD to HV is -0.934. In addition, GD does not show clear trends in Fig. 12(f) due to the high p-value ($p = 0.0846 > 0.05$) under Pearson statistical test[7]. The other two metrics Δ and Δ^* show negative correlations to HV at $-0.955, -0.978$, respectively. At the same time, the curves of Δ and Δ^* are above the origin when $p \in [1.0, 1.5]$ in Figs. 12(h-i). This implies that an optimal solution set with good HV (e.g., $\mathrm{HV}(S_2) > \mathrm{HV}(S_1)$) is likely to also show good Δ and Δ^* (e.g., $\Delta(S_2) < \Delta(S_1)$ and $\Delta^*(S_2) < \Delta^*(S_1)$). On the other hand, Δ and Δ^* contradict with HV on $\{f_1^p + f_2^p = 1 | 1.5 \leq p \leq 3\}$, since their curves are below the origin and they also exhibit negative correlations to HV.

Fig. 13 shows the relationships of the five metrics to the baseline HV on 3-dimensional PFs. From Figs. 13(a, f), GD does not exhibit stable trends. In general, $I^1_{\epsilon+}$ is consistent with HV, because the curve of $I^1_{\epsilon+}$ is above the origin and it is positively correlated to HV in most cases, when $p \in [0.8, 3.0]$. In addition, the metric Δ in Fig. 13(c) shows a negative correlation to HV on the 3-dimensional convex PFs, whereas, it is positively correlated to HV on the 3-dimensional concave PFs (Fig. 13(h)). On the other hand, metrics Δ^* and IGD share similar trends. Both Δ^* and IGD are consistent with HV on the 3-dimensional convex PFs, except for $p \in [0.3, 0.5]$, whereas, they contradict with HV on $\{f_1^p + f_2^p + f_3^p = 1 | 1.5 \leq p \leq 3\}$.

[4] $S_1 \prec S_2$ means S_1 is better than S_2 on both metrics (M_1, M_2).
[5] $\mathrm{Corr}(X, Y) = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$.
[6] $S_1 \not\prec S_2$ means S_1 non-dominates S_2 on the two metrics (M_1, M_2).
[7] If $p < 0.05$, the correlation of two metrics is significant with 95% confidence level. Otherwise, the correlation of two metrics is insignificant.

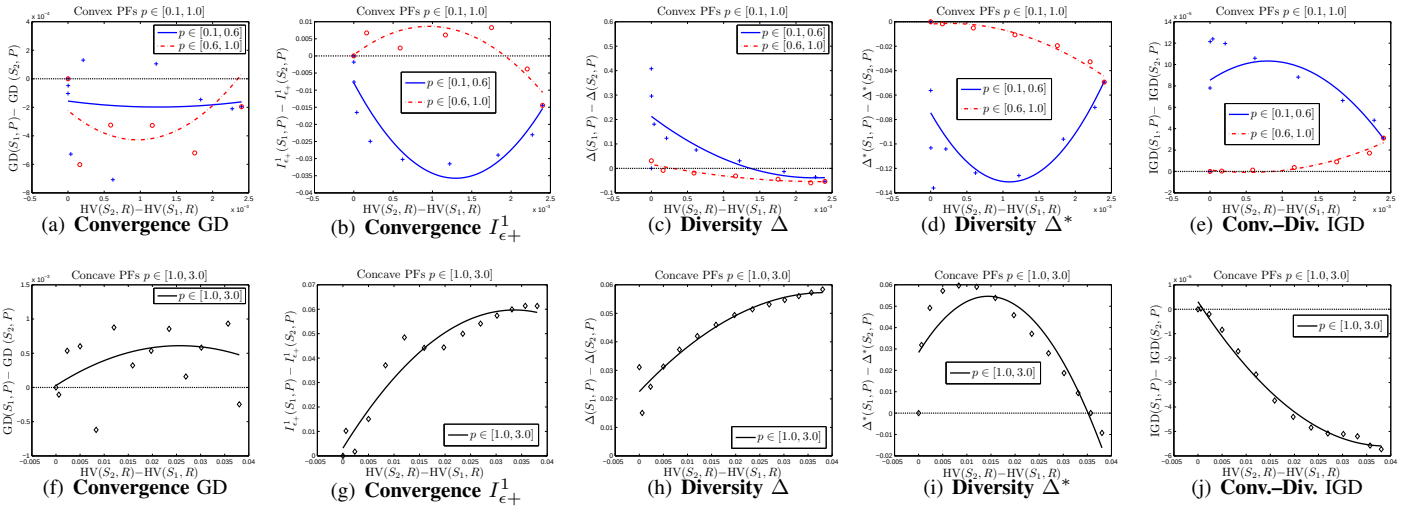


Fig. 14. The relationships of five metrics (GD, $I^1_{\epsilon+}$, Δ, Δ^* and IGD) to the baseline HV on 4-dimensional convex PFs (a-e) and concave PFs (f-j). The $I^1_{\epsilon+}$ and IGD are consistent with HV on 4-dimensional convex PFs, whereas, Δ^* and IGD contradict with HV on 4-dimensional concave PFs.

Fig. 14 summarizes the relationships of the five metrics to HV on 4-dimensional PFs. From Figs. 14(a, f), GD once again does not exhibit clear relationships to HV. On the other hand, in Figs. 14(b, g), $I^1_{\epsilon+}$ is consistent with HV in most cases, when $p \in [0.7, 3.0]$. Further, metric Δ displays contradiction with HV on the 4-dimensional convex PFs (Fig. 14(c)), whereas, it is consistent with HV on the 4-dimensional concave PFs (Fig. 14(h)). In Figs. 14(d, i), metric Δ^* displays negative correlations to HV. Last but not least, Figs. 14(e, j) indicate that IGD is consistent with HV on the 4-dimensional convex PFs, except for $p \in [0.3, 0.6]$, whereas, it contradicts with HV on $\{f_1^p + f_2^p + f_3^p + f_4^p = 1 | 1.0 \leq p \leq 3\}$.

From our studies on the optimal solution sets (S_1 and S_2) on 2, 3, 4-dimensional PFs, we summarize the characteristics and relationships of the six representative metrics (GD, $I^1_{\epsilon+}$, Δ, Δ^*, IGD and HV) considered in what follows.

- HV, which gives the convergence and diversity performance on a single scale, is identified as one of the most important metric to consider in MOO. Meanwhile, HV also requires little prior knowledge relative to the other five metrics. In particular, the comparison set of HV and other five metrics are reference set R and PF (P), respectively. It is easy to construct R but hard to generate P. The reasons are that constructing P requires the geometrical characteristics of the true PFs and $|R| \ll |P|$ (e.g., $|R| = 1, |P| = 10,001$ for 2-dimensional PFs in Section V-A).
- The metric GD is solely designed for measuring solution set convergence. Hence, it does not provide diversity information.
- The metric $I^1_{\epsilon+}$ is consistent with HV. Although $I^1_{\epsilon+}$ belongs to the family of *convergence* metrics, it can measure diversity when P approximates the true PFs accurately.
- The metric Δ works well for measuring diversity on bi-objective PFs, but it is not suitable for high dimensional PFs ($m \geq 3$). The metric Δ^* is an extension of Δ to deal with different forms of PFs.
- The metric IGD is consistent with HV on convex PFs. On concave PFs, in some special case $\{f_1^p + f_2^p = 1, f_1^p + f_2^p + f_3^p = 1, f_1^p + f_2^p + f_3^p + f_4^p = 1 | 1.5 \leq p \leq 3\}$, the metric IGD contradicts with HV.

G. Discussion of MOO Metrics

In this section, we present a discussion which we hope would serve as a guide on the appropriate use of MOO metrics. In particular, we summarize the three key points as follows:

1) As an important performance criterion in MOO, *capacity* is commonly used as the prerequisite ahead of the other criteria, when one attempts to measure the quality of optimal solution sets [4, 5]. The reasons being that all the other criteria would become statistical insignificant if the capacity measure, for example, the number of non-dominated solutions in any two optimal sets for comparison, differs or is too small in size. Hence, capacity metric serves as the paramount criterion for assessing multiobjective search algorithms.

2) On the convex PFs, two metrics, $I^1_{\epsilon+}$ and IGD, have shown high consistencies to metric HV (Section V-F). This indicates that these three metrics, i.e., $I^1_{\epsilon+}$, IGD and HV, can be jointly used to assess solution sets optimality on the convex PFs [4, 5]. In particular, IGD or HV gives the convergence and diversity information of solution sets on a single scale. In general, IGD is less costly to compute than HV, especially for the high dimensional PFs. However, as shown in Section V-A, the comparison set P of IGD can be more difficult to construct than the reference set R in HV. On the other hand, when interests are solely on the convergence quality of the optimal solution set, then, $I^1_{\epsilon+}$ can be adopted independently.

3) On the concave PFs, it is worth noting that only one metric $I^1_{\epsilon+}$ is found to be consistent with HV. As shown in Section V-F, IGD exhibit contradictions to HV. This indicates that any attempts to report the IGD measures jointly with HV on concave PFs may not make good sense. From here, the present study thus highlights the important need for the design of new *diversity* metrics that are consistent and appropriate for use jointly with HV, when dealing with concave PFs.

VI. Conclusion and Future Work

In this paper, we have classified the performance metrics of Multiobjective Optimization (MOO) into four groups, namely *Capacity*, *Convergence*, *Diversity*, *Convergence–Diversity*. With the presence of extreme cases, the inadequacies of some MOO metrics are analysed. Then, the relationship among representative metrics are investigated via empirical studies. In particular, metrics $I_{\epsilon+}^1$ and IGD are found to be consistent with HV on convex Pareto fronts (PFs). When the PFs are concave, however, IGD displayed contradictions to HV. As such, there is thus room for the introduction of new MOO metrics that are appropriate for use with HV on concave PFs. Another future work is to investigate the relationship of MOO metrics on PFs of different geometrical characteristics, such as discrete, many-objective, asymmetric PFs, etc.

The Matlab source codes of Pareto-adaptive weight vectors ($pa\lambda$) [16], and the relationships among the six metrics (GD, $I_{\epsilon+}^1$, Δ, Δ^*, IGD and HV) are at http://www.ntu.edu.sg/home/asysong/MOPmetrics-matlab.rar.

Acknowledgment

This work is partially supported under the A*Star-TSRP funding, the Singapore Institute of Manufacturing Technology-Nanyang Technological University Joint Laboratory and Collaborative research Programme on Complex Systems, and the Center for Computational Intelligence (C2I) at NTU.

References

[1] K. Deb. *Multi-objective optimization using evolutionary algorithms*, volume 16. Wiley, 2001.

[2] C.A.C. Coello. Evolutionary multi-objective optimization: a historical view of the field. *IEEE Comput. Intell. Mag.*, 1(1):28–36, 2006.

[3] H.A. Abbass, A. Bender, S. Gaidow, and P. Whitbread. Computational red teaming: Past, present and future. *IEEE Comput. Intell. Mag.*, 6(1):30–42, 2011.

[4] J.J. Durillo, A.J. Nebro, and E. Alba. The jmetal framework for multi-objective optimization: Design and architecture. In *Proc. IEEE Cong. Evol. Comput.*, pages 1–8, 2010.

[5] J.J. Durillo and A.J. Nebro. jMetal: A java framework for multi-objective optimization. *Advances in Engineering Software*, 42(10):760–771, 2011.

[6] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.*, 6(2):182–197, 2002.

[7] E. Zitzler, M. Laumanns, and L. Thiele. SPEA2: Improving the strength pareto evolutionary algorithm. *Computer Engineering and Networks Laboratory (TIK), Zurich, Switzerland*, (103), 2001.

[8] Q. Zhang and H. Li. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.*, 11(6):712–731, 2007.

[9] H. Li and Q. Zhang. Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II. *IEEE Trans. Evol. Comput.*, 13(2):284–302, 2009.

[10] E. Zitzler and S. Kúnzli. Indicator-based selection in multiobjective search. In *Proc. Parall. Problem Solv. Nature*, pages 832–842. Springer, 2004.

[11] N. Beume, B. Naujoks, and M. Emmerich. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 181(3):1653–1669, 2007.

[12] S. Kukkonen and J. Lampinen. GDE3: The third evolution step of generalized differential evolution. In *Proc. IEEE Cong. Evol. Comput.*, volume 1, pages 443–450, 2005.

[13] A.J. Nebro, J.J. Durillo, J. Garcia-Nieto, Coello C.A.C., F. Luna, and E. Alba. SMPSO: A new PSO-based metaheuristic for multi-objective optimization. In *Proc. IEEE Sympo. Comput. Intell. in MiultiCriteria Decision Making*, pages 66–73, 2009.

[14] S. Jiang, J. Zhang, and Y.S. Ong. Asymmetric pareto-adaptive scheme for multiobjective optimization. In *Proc. Conf. Artif. Intell.*, pages 351–360. Springer, 2011.

[15] I.W. Tsang C.W. Seah, Y.S. Ong and S. Jiang. Pareto rank learning in multi-objective evolutionary algorithms. In *Proc. IEEE Cong. Evol. Comput.*, pages 1–8, 2012.

[16] S. Jiang, Z. Cai, J. Zhang, and Y.S. Ong. Multiobjective optimization by decomposition with pareto-adaptive weight vectors. In *Proc. Conf. Natural Computation*, volume 3, pages 1260–1264, 2011.

[17] S. Jiang, J. Zhang, and Y.S. Ong. A multiagent evolutionary framework based on trust for multiobjective optimization. In *Proc. Conf. Autonomous Agents and Multiagent Systems*, pages 299–306, 2012.

[18] M.N. Le, Y.S. Ong, S. Menzel, C.W. Seah, and B. Sendhoff. Multi co-objective evolutionary optimization: Cross surrogate augmentation for computationally expensive problems. In *Proc. IEEE Cong. Evol. Comput.*, pages 1–8, 2012.

[19] C.K. Goh, K.C. Tan, C.Y. Cheong, and Y.S. Ong. Noise-induced features in robust multi-objective optimization problems. In *Proc. IEEE Cong. Evol. Comput.*, pages 568–575, 2007.

[20] Y. Wang, Z. Cai, G. Guo, and Y. Zhou. Multiobjective optimization and hybrid evolutionary algorithm to solve constrained optimization problems. *IEEE Trans. Sys., Man and Cybern., Part B*, 37(3):560–575, 2007.

[21] S. Mostaghim and J. Teich. Strategies for finding good local guides in multi-objective particle swarm optimization (MOPSO). In *Proc. IEEE Swarm Intell. Symposium*, pages 26–33, 2003.

[22] D. Liu, K.C. Tan, C.K. Goh, and W.K. Ho. A multiobjective memetic algorithm based on particle swarm optimization. *IEEE Trans. Sys., Man and Cybern., Part B*, 37(1):42–50, 2007.

[23] W.F. Leong and G.G. Yen. PSO-based multiobjective optimization with dynamic population size and adaptive local archives. *IEEE Trans. Sys., Man and Cybern., Part B*, 38(5):1270–1293, 2008.

[24] S. Bandyopadhyay, S.K. Pal, and B. Aruna. Multiobjective GAs, quantitative indices, and pattern classification. *IEEE Trans. Sys., Man and Cybern., Part B*, 34(5):2088–2099, 2004.

[25] X. Zou, Y. Chen, M. Liu, and L. Kang. A new evolutionary algorithm for solving many-objective optimization problems. *IEEE Trans. Sys., Man and Cybern., Part B*, 38(5):1402–1412, 2008.

[26] M. Daneshyari and G.G. Yen. Cultural-based multiobjective particle swarm optimization. *IEEE Trans. Sys., Man and Cybern., Part B*, 41(2):553–567, 2011.

[27] E. Masazade, R. Rajagopalan, P.K. Varshney, C.K. Mohan, G.K. Sendur, and M. Keskinoz. A multiobjective optimization approach to obtain decision thresholds for distributed detection in wireless sensor networks. *IEEE Trans. Sys., Man and Cybern., Part B*, 40(2):444–457, 2010.

[28] G.G. Yen and W.F. Leong. Dynamic multiple swarms in multiobjective particle swarm optimization. *IEEE Trans. Sys., Man and Cybern., Part A*, 39(4):890–911, 2009.

[29] R.V. Kulkarni and G.K. Venayagamoorthy. Particle swarm optimization in wireless-sensor networks: a brief survey. *IEEE Trans. Sys., Man and Cybern., Part C*, 41(2):262–267, 2011.

[30] S. Bandyopadhyay. Multiobjective simulated annealing for fuzzy clustering with stability and validity. *IEEE Trans. Sys., Man and Cybern., Part C*, 41(5):682–691, 2011.

[31] G. Avigad and A. Moshaiov. Interactive evolutionary multiobjective search and optimization of set-based concepts. *IEEE Trans. Sys., Man and Cybern., Part B*, 39(4):1013–1027, 2009.

[32] C.K. Goh and K.C. Tan. A competitive-cooperative coevolutionary paradigm for dynamic multiobjective optimization. *IEEE Trans. Evol. Comput.*, 13(1):103–127, 2009.

[33] C.K. Goh and K.C. Tan. An investigation on noisy environments in evolutionary multiobjective optimization. *IEEE Trans. Evol. Comput.*, 11(3):354–381, 2007.

[34] K. Tang, K.C. Tan, and H. Ishibuchi. Guest editorial: Memetic algorithms for evolutionary multi-objective optimization. *Memetic Computing*, 2(1):1–1, 2010.

[35] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima. Simultaneous use of different scalarizing functions in moea/d. In *Proc. Genet. Evol. Comput.*, pages 519–526. ACM, 2010.

[36] Z. Wang, K. Tang, and X. Yao. Multi-objective approaches to optimal testing resource allocation in modular software systems. *IEEE Trans. Reliability*, 59(3):563–575, 2010.

[37] Y. Mei, K. Tang, and X. Yao. Decomposition-based memetic algorithm for multiobjective capacitated arc routing problem. *IEEE Trans. Evol. Comput.*, 15(2):151–165, 2011.

[38] Z. Wang, T. Chen, K. Tang, and X. Yao. A multi-objective approach to redundancy allocation problem in parallel-series systems. In *Proc. IEEE Cong. Evol. Comput.*, pages 582–589. IEEE, 2009.
[39] K. Tang, Z. Wang, X. Cao, and J. Zhang. A multi-objective evolutionary approach to aircraft landing scheduling problems. In *Proc. IEEE Cong. Evol. Comput.*, pages 3650–3656. IEEE, 2008.
[40] A.J. Nebro, J.J. Durillo, and Carlos A. Coello Coello C.A. Coello Coello. Analysis of leader selection strategies in a multi-objective particle swarm optimizer. In *Proc. IEEE Cong. Evol. Comput.*, 2013.
[41] D.A. Van Veldhuizen and G.B. Lamont. On measuring multiobjective evolutionary algorithm performance. In *Proc. IEEE Cong. Evol. Comput.*, pages 204–211, 2000.
[42] D.A. Van Veldhuizen and G.B. Lamont. Multiobjective evolutionary algorithm test suites. In *Proc. ACM Symposium on Applied Computing*, pages 351–357, 1999.
[43] A. Zhou, Y. Jin, Q. Zhang, B. Sendhoff, and E. Tsang. Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion. In *Proc. IEEE Cong. Evol. Comput.*, pages 892–899, 2006.
[44] E. Zitzler and L. Thiele. Multiobjective optimization using evolutionary algorithms–A comparative case study. In *Proc. Parall. Problem Solv. Nature*, pages 292–301. Springer, 1998.
[45] E. Zitzler and L. Thiele. Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach. *IEEE Trans. Evol. Comput.*, 3(4):257–271, 1999.
[46] T. Okabe, Y. Jin, and B. Sendhoff. A critical survey of performance indices for multi-objective optimisation. In *Proc. IEEE Cong. Evol. Comput.*, volume 2, pages 878–885, 2003.
[47] E. Zitzler, L. Thiele, M. Laumanns, C.M. Fonseca, and V.G. da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Trans. Evol. Comput.*, 7(2):117–132, 2003.
[48] K.C. Tan, T.H. Lee, and E.F. Khor. Evolutionary algorithms for multi-objective optimization: performance assessments and comparisons. *Artif. Intell. Review*, 17(4):251–290, 2002.
[49] J. Knowles and D. Corne. On metrics for comparing nondominated sets. In *Proc. IEEE Cong. Evol. Comput.*, volume 1, pages 711–716, 2002.
[50] C. Grosan, M. Oltean, and D. Dumitrescu. Performance metrics for multiobjective optimization evolutionary algorithms. In *Proc. Applied and Industrial Mathematics*, 2003.
[51] J. Wu and S. Azarm. Metrics for quality assessment of a multiobjective design optimization solution set. *Journal of Mechanical Design*, 123(1):18–25, 2001.
[52] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In *Proc. Parall. Problem Solv. Nature*, pages 849–858, 2000.
[53] D.A. Van Veldhuizen. Multiobjective evolutionary algorithms: classifications, analyses, and new innovations. *Evol. Comput.*, 1999.
[54] P. Czyżżak and A. Jaszkiewicz. Pareto simulated annealinga metaheuristic technique for multiple-objective combinatorial optimization. *Journal of Multi-Criteria Decision Analysis*, 7(1):34–47, 1998.
[55] M.P. Hansen and A. Jaszkiewicz. Evaluating the quality of approximations to the non-dominated set. *Technical University of Denmark, Denmark, IMM Technical Report IMM-REP-1998-7*, 1998.
[56] E. Zitzler, K. Deb, and L. Thiele. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evol. Comput.*, 8(2):173–195, 2000.
[57] J.R. Schott. Fault tolerant design using single and multicriteria genetic algorithm optimization. Technical report, DTIC Document, 1995.
[58] A. Farhang-Mehr and S. Azarm. Diversity assessment of pareto optimal solution sets: An entropy approach. In *Proc. IEEE Cong. Evol. Comput.*, volume 1, pages 723–728, 2002.
[59] S. Mostaghim and J. Teich. A new approach on many objective diversity measurement. *Practical Appro. MultiObj. Optimi.*, page 254, 2005.
[60] J. Bader and E. Zitzler. HypE: An algorithm for fast hypervolume-based many-objective optimization. *Evol. Comput.*, 19(1):45–76, 2011.
[61] L. While and L. Bradstreet. Applying the WFG algorithm to calculate incremental hypervolumes. In *Proc. IEEE Cong. Evol. Comput.*, pages 1–8, 2012.
[62] E. Zitzler. Evolutionary algorithms for multiobjective optimization: Methods and applications. *PhD thesis*, 1999.
[63] K. Bringmann and T. Friedrich. Approximating the least hypervolume contributor: NP-hard in general, but fast in practice. *Theoretical Computer Science*, 425:104–116, 2012.
[64] B. Naujoks, N. Beume, and M. Emmerich. Multi-objective optimisation using S-metric selection: Application to three-dimensional solution spaces. In *Proc. IEEE Cong. Evol. Comput.*, volume 2, pages 1282–1289, 2005.
[65] C.M. Fonseca, L. Paquete, and M. López-Ibánez. An improved dimension-sweep algorithm for the hypervolume indicator. In *Proc. IEEE Cong. Evol. Comput.*, pages 1157–1163, 2006.
[66] N. Beume, C.M. Fonseca, M. López-Ibáñez, L. Paquete, and J. Vahrenhold. On the complexity of computing the hypervolume indicator. *IEEE Trans. Evol. Comput.*, 13(5):1075–1082, 2009.
[67] N. Beume. S-metric calculation by considering dominated hypervolume as Klee's measure problem. *Evol. Comput.*, 17(4):477–492, 2009.
[68] L. Bradstreet. *The Hypervolume Indicator for Multi-objective Optimisation: Calculation and Use*. PhD thesis, 2010.
[69] L. While, L. Bradstreet, and L. Barone. A fast way of calculating exact hypervolumes. *IEEE Trans. Evol. Comput.*, 16(1):86–95, 2012.
[70] O. Schutze, X. Esquivel, A. Lara, and C.A.C. Coello. Using the averaged hausdorff distance as a performance measure in evolutionary multiobjective optimization. *IEEE Trans. Evol. Comput.*, 16(4):504–522, 2012.

Siwei Jiang received the M.S. and Ph.D. degrees in computer science from the China University of Geosciences (CUG), Wuhan, China, in 2006 and 2011, respectively.

He is currently a Ph.D. student in School of Comouter Engineering, Nanyang Technological University (NTU), Singapore. His research interests include multiagent evolutionary algorithms, multiagent system, reputation systems and vehicle routing problems.

Yew-Soon Ong received the B.S. and M.S. degrees in Electrical and Electronics Engineering from Nanyang Technological University (NTU), Singapore, in 1998 and 1999, respectively, and the Ph.D. degree on artificial intelligence in complex design from the Computational Engineering and Design Center, University of Southampton, Southampton, U.K., in 2002.

He is currently an Associate Professor, Director of the Center for Computational Intelligence at the School of Computer Engineering, NTU, and co-Director of the SIMTECH-NTU Joint Lab on Complex Systems. His research interest in computational intelligence spans across memetic computation, evolutionary design, machine learning and agent-based systems. Dr. Ong is the Founding Technical Editor-in-Chief of the Memetic Computing Journal, the Chief Editor of the Springer book series on studies in adaptation, learning, and optimization, and an Associate Editor of the IEEE Computational Intelligence Magazine, IEEE Transactions on Evolutionary Computation, IEEE Transactions on Neural Network & Learning Systems, IEEE Transactions on Cybernetics and others.

Jie Zhang received the Ph.D. degree from the University of Waterloo, Waterloo, Canada, in 2009.

He is currently an Assistant Professor at the School of Computer Engineering, Nanyang Technological University (NTU), Singapore. His research interests include artificial intelligence and multiagent systems, trust modeling and incentive mechanisms, and mobile and vehicular ad hoc networks.

Liang Feng received his B.S. degree in School of Telecommunication and Information Engineering from Nanjing University of Posts and Telecommunications, China, in 2009. Currently, he is working towards the Ph.D degree in computer engineering at the Center for Computational Intelligence, School of Comouter Engineering, Nanyang Technological University (NTU). His primary research interests include evolutionary computation, memetic computing, and transfer learning, etc.