A grid-based inverted generational distance for multi/many-objective optimization

Cai, Xinye; Xiao, Yushun; Li, Miqing; Hu, Han ; Ishibuchi, Hisao; Li, Xiaoping

DOI:
10.1109/TEVC.2020.2991040

A Grid-Based Inverted Generational Distance for Multi/Many-Objective Optimization

Xinye Cai, *Member, IEEE,* Yushun Xiao, Miqing Li, *Member, IEEE,* Han Hu, Hisao Ishibuchi *Fellow, IEEE,*, Xiaoping Li, *Senior Member, IEEE,*

***Abstract*—Assessing the performance of Pareto front (PF) approximations is a key issue in the field of evolutionary multi/many-objective optimization. Inverted Generational Distance (IGD) has been widely accepted as a performance indicator for evaluating the comprehensive quality for a PF approximation. However, IGD usually becomes infeasible when facing a real-world optimization problem as it needs to know the true PF *a priori*. In addition, the time complexity of IGD grows quadratically with the size of the solution/reference set. To address the aforementioned issues, a grid-based IGD (Grid-IGD) is proposed to estimate both convergence and diversity of PF approximations for multi/many-objective optimization. In Grid-IGD, a set of reference points is generated by estimating PFs of the problem in question, based on the representative nondominated solutions of all the approximations in a grid environment. To reduce the time complexity, Grid-IGD only considers the closest solution within the grid neighborhood in the approximation for every reference point. Grid-IGD also possesses other desirable properties such as Pareto compliance, immunity to dominated/duplicate solutions and no need of normalization. In the experimental studies, Grid-IGD is verified on both the artificial and real PF approximations obtained by five many-objective optimizers. Effects of the grid specification on the behavior of Grid-IGD are also discussed in detail theoretically and experimentally.**

***Index Terms*—Many-objective optimization, Inverted generational distance, Performance indicator, Grid system.**

I. Introduction

Many real-world optimization problems involve the simultaneous optimization of multiple conflicting objectives. Unlike a single-objective optimization problem, a multiobjective optimization problem (MOP) has a set of Pareto optimal solutions, concerning the trade-offs between different objectives. The set of all the Pareto optimal solutions is usually called the *Pareto set* (*PS*) and its mapping in the objective space is called the *Pareto front* (*PF*) [1]. A PF approximation apparently can help decision makers for understanding the tradeoff relationship among different objectives and selecting their preferred solutions. Over the past decades, multiobjective evolutionary algorithms (MOEAs) [2], [3], [4], [5] have been accepted as a major methodology for approximating the PFs in nonlinear MOPs [6], [7].

In the field of MOEAs, the performance evaluation is a critical issue. The quality of an obtained solution set can be measured by performance indicators [8], [9] in one or several of the following aspects: 1) convergence, 2) spread (i.e., coverage [10] or extensity [11]) and 3) uniformity. The combination of the latter two is usually called diversity of a solution set [12], [11].

The performance indicators used to assess solution sets obtained by multiobjective optimizers are critical. An ideal indicator should be able to correctly reflect the quality of a solution set. To achieve such a goal, a performance indicator is desirable to have as more following properties as possible [9].

1) Less prior problem information needed: Many existing indicators require additional problem information, e.g., the true Pareto front or a reference point. As the accuracy of such indicators largely depend on those references, it is desirable for indicators to have as little reference information as possible.
2) No need of scaling and normalization: For some indicators, scaling (or normalization) may need for indicators whose calculation involves objective blending. Nevertheless, an indicator may not need such an operation when different objectives may contribute equally to the indicator values. Apparently, indicators without requiring scaling and normalization are more desirable.
3) No effect of adding dominated or duplicate solutions: as dominated or duplicate solutions do not provide any useful information for the decision maker in the context of Pareto optimality, indicators that have no effect of adding dominated or duplicate solutions are more desirable.
4) Low computational cost: The increasing number of objectives results in an exponential increase in the time and space complexity for some commonly used indicators, e.g., hypervolume [13] and hyperarea difference [14]. Such indicators are incapable of computing high-dimensional PFs. It is more desirable to design an

Xinye Cai, Yushun Xiao and Han Hu are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, 210016, China, and also with the Collaborative Innovation Center of Novel Software Technology and Industrialization, Nanjing 210023, China (e-mail: xinye@nuaa.edu.cn, huhan@nuaa.edu.cn).

Miqing Li is with Centre of Excellence for Research in Computational Intelligence and Applications (CERCIA) in the School of Computer Science at the University of Birmingham.

Hisao Ishibuchi is with Shenzhen Key Laboratory of Computational Intelligence, University Key Laboratory of Evolving Intelligent Systems of Guangdong Province, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China.

Xiaoping Li is with the School of Computer Science and Engineering, Southeast University, Nanjing, 211189, China; and also with the Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing, 211189, China (Tel: 86-25-52091916; Fax: 86-25-52091916). (e-mail: xpli@seu.edu.cn).

This work was supported in part by the National Natural Science Foundation of China (NSFC) under grant 61300159, 61732006, 61876075, 61572127, 61872077 and 61832004, by the Natural Science Foundation of Jiangsu Province of China under grant BK20181288, by China Postdoctoral Science Foundation under grant 2015M571751 and by the Aeronautical Science Foundation of China under grant 20175552042.

indicator with low computational cost.

5) Pareto compliance property: Pareto compliance property is very important for the indicator design. The formal definition of strict/weak Pareto compliance can be referred to Section II. A. Unfortunately, many widely used indicators, e.g., Inverted Generational Distance (IGD) [15], [16], are Pareto non-compliant.

Some comprehensive performance indicators, such as Hypervolume (HV) [13] and Inverted Generational Distance (IGD) [15], [16], are very popular in multiobjective evolutionary optimization community [8], [9]. However, they do not own several above-mentioned desirable properties. For instance, although Monte Carlo sampling-based approximation can significantly reduce the computational cost of calculating HV and makes it possible to use for high-dimensional PFs [17], the proper choice of the reference point is a tedious task, which will largely affect the ability of HV to distinguish the quality of PF approximations [18]. For IGD, a reference set that can well-represent the real PF is required, which apparently is very difficult to meet for real-world optimization problems. Recent studies also show that IGD-based comparison results largely depend on the specification of such a reference set [19].

To meet the above-mentioned properties, we propose a grid-based IGD (Grid-IGD) as a comprehensive performance indicator for multiobjective optimization. The major motivations of this work can be summarized as follows:

1) For a real-world MOP, IGD is generally infeasible as the true PF in question is usually unknown in advance. A grid system can provide a set of reference points by adopting the utopia point (i.e., the left bottom corner for a minimization problem) of each grid that the non-dominated solutions are located in. Such a reference set automatically generated from the solution sets by a grid system makes Grid-IGD require very little prior problem information and have no need of normalization before calculation.
2) As referenced in [20], when the true PF of an MOP is unknown, the current mainstream methods use all the non-dominated solutions as the reference points [21], [19]. The main disadvantage of this approach is that the generated reference points are always not uniform over the entire PF. Such a biased distribution of the reference points may lead to the biased comparison results [20]. How to design a method that is easy to implement and can sample well-distributed reference points in the high-dimensional space, is of great importance for calculating IGD. A grid system, as an inherent diversity maintenance mechanism, can be suitable for such a task.
3) The time complexity of IGD is $O(mMN)$ where m is the number of objectives, M is the number of reference points and N is the size of solution set. The time complexity is usually significantly larger than $O(mN^2)$ as the number of reference points is usually much larger than the size of PF approximation. In IGD, calculating the distance of a reference point to a far away solution appears to be redundant as it only takes the closest distance into account. When the grid system is introduced, the grid distance between solutions can be seen as a form of the definition for solutions' neighborhood. If each reference point only considers its neighborhood, it will lead to a much less time complexity, which is especially desirable for evaluating the performance of high-dimensional PF approximations, as well as using IGD potentially as an online indicator.
4) IGD is Pareto non-compliant, not immune to dominated solutions, and requires normalization before calculation. The lack of these desirable properties seriously limit its use and may also cause misleading comparison results. A comprehensive indicator that possesses most or even all of the aforementioned desirable properties needs to be further designed.

The rest of this paper is organized as follows. Related work on the proposed indicator, i.e., Grid-IGD, is introduced in Section II. Section III elaborates Grid-IGD. In Section IV, the systematic experiments are conducted to verify the effectiveness of Grid-IGD. Finally, the paper is concluded in Section V where future research directions are also suggested.

II. Background

A. Basic Definitions

A multiobjective optimization problem (MOP) can be defined as follows:

$$\begin{aligned} \text{minimize} \quad & F(x) = (f_1(x), \ldots, f_m(x))^T \\ \text{subject to} \quad & x \in \Omega \end{aligned} \tag{1}$$

where Ω is the *decision space*, $F : \Omega \to R^m$ consists of m real-valued objective functions. Note that an MOP is usually called a many-objective optimization problem (MaOPs) when $m > 3$. The *attainable objective set* is $\{F(x)|x \in \Omega\}$. Let $a, b \in R^m$, a is said to *dominate* b, denoted by $a \prec b$, if and only if $a_i \leq b_i$ for every $i \in \{1, \ldots, m\}$ and $a_j < b_j$ for at least one index $j \in \{1, \ldots, m\}$; a is said to *weakly dominate* b, denoted by $a \preceq b$, if and only if $a_i \leq b_i$ for every $i \in \{1, \ldots, m\}$[1]. Given a set S in R^m, a solution in S is called non-dominated in S if no other solution in S dominates it. A solution $x^* \in \Omega$ is *Pareto-optimal* if $F(x^*)$ is non-dominated in the attainable objective set. $F(x^*)$ is then called a *Pareto-optimal (objective) vector*. In other words, any improvement in one objective of a Pareto optimal solution must lead to deterioration in at least another objective. The set of all the Pareto-optimal points is called the *Pareto set* (*PS*) and the set of all the Pareto-optimal objective vectors is the *Pareto front* (*PF*) [1].

In [8], the above-mentioned Pareto dominance relation between solutions was extended to a relation between solution sets as follows. Let A and B be two solution sets. A is said to *dominate* B, denoted by $A \prec B$, if and only if $\forall b_j \in B, \exists a_i \in A : a_i \prec b_j$. A is said to *weakly dominate* B, denoted by $A \preceq B$, if and only if $\forall b_j \in B, \exists a_i \in A : a_i \preceq b_j$.

A quality indicator is said to be strictly Pareto compliant [8], [9], if and only if $\forall A, B : A \prec B \implies I(A) < I(B)$; where

[1] In the case of maximization, the inequality signs should be reversed.

$I(.)$ is a mapping from a set of objective vector to an indicator value. Similarly, a quality indicator is said to be weakly Pareto compliant [8], [9], if and only if $\forall A, B : A \preceq B \implies I(A) \leq I(B)$; where $I(.)$ is a mapping from a set of objective vector to an indicator value.

The *ideal* and *nadir objective vectors* can be used to define the ranges of PFs as follows. The *ideal objective vector* z^* is a vector $z^* = \{z_1^*, \ldots, z_m^*\}^T$, which can be computed by

$$z_j^* = \min_{x \in \Omega} f_j(x), j \in \{1, \ldots, m\}. \tag{2}$$

The *nadir objective vector* z^{nad} is a vector $z^{nad} = \{z_1^{nad}, \ldots, z_m^{nad}\}^T$, which can be computed by

$$z_j^{nad} = \max_{x \in PS} f_j(x), j \in \{1, \ldots, m\}. \tag{3}$$

B. Previous work

As mentioned in the last section, IGD and HV are very popular in the field of multiobjective evolutionary optimization. The definition of IGD [15], [16] and HV [13] are given as follows.

- Inverted Generational Distance (IGD) [15], [16]: Let P^* be a set of points uniformly sampled over the true PF, and S be the set of solutions obtained by an MOEA. The IGD value of S is computed as:

$$IGD(S, P^*) = \frac{\sum_{y \in P^*} dist(y, S)}{|P^*|} \tag{4}$$

  where $dist(y, S)$ is the Euclidean distance between a point $y \in P^*$ and its nearest neighbor in S, and $|P^*|$ is the cardinality of P^*. IGD calculates an average minimum distance from each point in P^* to those in S, which measures both convergence and diversity of a solution set S. The lower the IGD value is, the better the quality of S is.
- Hypervolume (HV) [13]: Let $r^* = (r_1^*, r_2^*, ..., r_m^*)^T$ be a reference point in the objective space that is dominated by all solutions in a PF approximation S. HV metric measures the size of the objective space dominated by the solutions in S and bounded by r^*.

$$HV(S) = VOL(\bigcup_{y \in S} [y_1, r_1^*] \times ...[y_m, r_m^*]) \tag{5}$$

  where $VOL(\bullet)$ indicates the Lebesgue measure.
  The main advantage of HV is its strictly Pareto compliance property [22]. One main drawback of the HV compared with the IGD is its large computational burden for many-objective optimization problem. Although some fast computational methods have been proposed for HV [23], [24], [25], it is still difficult to compute the exact value of HV for a large solution set with many objectives. Another weakness of HV is that it may be in favor of very non-uniform solution sets on a highly nonlinear Pareto front no matter what reference point is [26], possibly leading to unfair comparison results.

In addition to the aforementioned performance metrics, other comprehensive performance indicators are also frequently used, as follows.

- Modified Inverted Generational Distance (IGD$^+$) [27]: Let P^* be a set of points uniformly sampled over the true PF, and S be the set of solutions obtained by an MOEA. The only difference between the IGD(S, P^*) and IGD$^+$(S, P^*) lies in the distance calculation. In the minimization problems, the distance calculation for IGD$^+$ is

$$dist^+(y, S) = \sqrt{\sum_{i=1}^{m} (\max\{y_i - z_i, 0\})^2} \tag{6}$$

  where $dist^+(y, S)$ is the modified distance between a point $y \in P^*$ and its nearest neighbor $z \in S$. This distance modification ensures that IGD$^+$ is weakly Pareto compliant whereas the original IGD is Pareto non-compliant. Like IGD, the lower the IGD$^+$ value is, the better the quality of S is. Although IGD$^+$ possesses the weak Pareto compliance property, it still requires prior knowledge of PF as a reference set, as well as normalization before calculation.
- Δ_p [28] can be seen as an "averaged Hausdorff distance" between the obtained solution set S and the reference set P^*, which evaluates both convergence and diversity as follows [29], [30].

$$\begin{aligned} \Delta_p(S, P^*) &= \max(GD_p(S, P^*), IGD_p(S, P^*)) \\ &= \max((\frac{1}{|S|} \sum_{y \in S} dist(y, P^*)^p)^{\frac{1}{p}}, \\ &\quad (\frac{1}{|P^*|} \sum_{y \in P^*} dist(y, S)^p)^{\frac{1}{p}}) \end{aligned} \tag{7}$$

  where $dist(y, P^*)$ is the Euclidean distance between a solution $y \in S$ and its nearest neighbor in P^*, and $|S|$ is the cardinality of S, $dist(y, S)$ is the Euclidean distance between a solution $y \in P^*$ and its nearest neighbor in S, and $|P^*|$ is the cardinality of P^*. The lower the Δ_p value is, the better the quality of S for approximating the whole PF is. However, a reference set, as well as normalization is still needed before calculating Δ_p.
- p-metric [31] is a newly proposed indicator for high-dimensional approximations. With a set of uniform reference vectors, the objective space can be divided into subregions of hypercones. Given a solution set S, a solution $y \in S$ belongs to i-th subregion Φ_i if $i = argmax_{\lambda^i \in V} \frac{(\lambda^i)^T \cdot y}{\|\lambda^i\| \|y\|}$, where λ^i is the i-th reference vector. In each subregion Φ_i, the solution with the closest distance r_i to the origin point is used to compute p-metric, as follows.

$$p\text{-metric} = \sum_{i=1}^{M} \frac{1}{r_i} \tag{8}$$

  where M is the number of subregions and $\frac{1}{r} = 0$ indicates no solution exists in such a subregion. It can be observed from Eq.(8) that the diversity of a solution set, in terms of the p-metric, is measured by the number of reference vectors (subregions) that have been associated with solution(s). It should be noted that for p-metric, one subregion may contain more than one solution but one solution can be located in one and only one subregion.

As pointed out in [10], its accuracy cannot be improved by increasing the number of reference vectors, as N solutions can be at most located in N subregions. In addition, p-metric needs normalization before calculation. The existence of dominated solutions may also affect its comparison results.

- R2 [32] was first proposed to assess the relative quality of two solution sets. Assuming the standard weighted Tchebycheff function with a particular reference point z^*, the indicator can be used to assess the quality of a single individual set against z^*[33]. Given an approximation set S, a set of weight vectors W, and the standard Tchebycheff aggregation function, the R2 indicator can be defined as:

$$R2(S, W, z^*) = \frac{1}{|W|} \sum_{w \in W} \min_{y \in S} \{\max\{w_i(y_i - z_i^*)\}\} \quad (9)$$

where $i \in \{1, \ldots, m\}$, $w = (w_1, \ldots, w_m)$ is a weight vector, z^* is the ideal point. A lower R2 value indicates that a solution set S is closer to the reference point. The R2 indicator possesses a desirable property of weak Pareto compliance. However, it still needs a reference point and normalization before calculation.

- Hypervolume Ratio (HVR) [34] is calculated as the ratio of the hypervolume(HV) [13] of non-dominated solutions S to the hypervolume of Pareto optimal solutions(P^*).

$$HVR = \frac{HV(S)}{HV(P^*)} \quad (10)$$

HV is the volume of space occupied by the union of hypercubes constructed by the non-dominated solutions and the reference point. Thus, HVR quantifies both the convergence and the diversity of the non-dominated solutions. A higher HVR indicates that the non-dominated solutions are closer to the Pareto front and more diverse in the objective space.

- Hypervolume Difference (HVD) [35] measures the gap between the hypervolume of the obtained PF approximation S and that of the true PF, i.e., P*. HVD indicator can be defined as:

$$HVD = HV(S) - HV(P^*) \quad (11)$$

where $HV(S)$ is the hypervolume of a set S. By comparing the hypervolume of an observed nondominated set with that of the true Pareto set, a quantitative measure is obtained as to how much worse an observed non-dominated set is when compared to the true Pareto set. Unfortunately, like HV, both HVR and HVD require prior knowledge on a reference point. Their computational complexity is exponential to the number of objectives, which makes both of them difficult to use for high-dimensional PF approximations.

The major properties summarized in Section I of the aforementioned indicators and our proposed Grid-IGD are listed in Table I. It can be seen that Grid-IGD has almost all the desirable properties. For instance, unlike HV, HVR, HVD, IGD, IGD$^+$, Δ_p and R2, Grid-IGD does not need any prior knowledge about reference point or true PF, although it needs to specify one neighborhood size parameter T for reducing its computational complexity. Unlike HVR, HVD, IGD, IGD$^+$, Δ_p, p-metric and R2, it does not need to scale before using, not to mention that the dominated or duplicate solutions have no effect on it. Unlike indicators based on HV whose computational complexity grows exponentially with the increase of the objectives, the computational complexity of Grid-IGD grows linearly with the increase of the objectives and it is less than quadratic with regard to the population size. Thus it can work well for high-dimensional PF approximations. In contrast to the original IGD, Δ_p and p-metric that are Pareto non-compliant, Grid-IGD possesses the desirable weak Pareto compliance property.

It is worth noting that only the representative performance indicators that measure both convergence and diversity for multi/many-objective optimization have been discussed above. There exist more performance indicators that measures either one of the above two aspects or both of them in the literature. A more comprehensive survey can be found in [9].

III. GRID-IGD

The grid system has already been adopted in some multi/many-objective optimization algorithms. For example, in ϵ-MOEA [36], the solutions are selected based on the ϵ-dominance and distances to the grid corner points. Similarly, the grid system has been adopted in [37] for designing many-objective optimization algorithm, due to its inherent property of reflecting the information of neighborhood structures among the solutions. In [3], a constrained decomposition with a grid system is proposed for multiobjective optimization. As Grid-IGD is also computed in a grid system, the setup of a grid system is firstly introduced in this section, followed by the main framework of Grid-IGD.

A. The setup of a grid system

A grid system can be set up as follows. Each objective is divided into K equal intervals within the approximations of the ideal point z^* and the extended nadir point z^{e_nad}, which can be defined as follows.

$$z^{e\_nad} = z^{nad} + (z^{nad} - z^*)/K \quad (12)$$

The reason of adopting the extended nadir point instead of the original nadir point have been elaborated in Section.III-B2.

Then the grid interval vector D for m objectives can be calculated by:

$$D = (d_1, \ldots, d_m)^T \quad \text{where} \quad d_j = \left(z_j^{e\_nad} - z_j^*\right)/K \quad (13)$$

Definition 1: Grid Index: For a solution x, its grid index in the grid system can be defined as follows.

$$g_j(x) = \lfloor (f_j(x) - z_j^*)/d_j \rfloor \ , \ j \in \{1, \ldots, m\} \quad (14)$$

where $\lfloor . \rfloor$ denotes the floor function. The grid index of x is marked as $G(x) = (g_1(x), ...g_m(x)) \in \{0, \ldots, (K-1)\}^m$.

TABLE I: The comparison between the existing indicators and Grid-IGD

Indicator	HV	HVR	HVD	IGD	IGD$^+$	Δ_p	p-metric	R2	Grid-IGD
Prior knowledge needed	reference point	reference point	reference point	reference PF	reference PF	reference PF		reference point	one parameter T
Scaling before calculation needed		✓	✓	✓	✓	✓	✓	✓	
Effect of dominated/ duplicate solutions				✓		✓	✓		
Difficult for high dimensional PFs	✓	✓	✓						
Computational effect	exponential in m	exponential in m	exponential in m	quadratic	quadratic	quadratic	quadratic	quadratic	less than quadratic ($O(Nlog^{m-1}N)$)
Pareto Compliant	strictly	strictly	strictly	weakly				weakly	weakly

Definition 2: Grid corner point: For a solution x whose grid index is $G(x) = (g_1(x), \ldots, g_m(x))$, its grid corner point z^{gcp} can be defined as follows.

$$z^{gcp}(x) = (z_1^{gcp}, \ldots, z_m^{gcp})^T \quad \text{where} \quad z_j^{gcp} = z_j^* + d_j \times g_j \tag{15}$$

where d_j is the grid interval on the j-th objective.

Definition 3: Grid Distance: Let $u, v \in R^m$ be two points, the grid distance $GD(u, v)$ between u and v is defined as

$$GD(u, v) = \sum_{j=1}^{m} |g_j(u) - g_j(v)|. \tag{16}$$

Definition 4: Grid Neighbors: For a solution set S, the grid neighbors of a point $r \in R^m$ within distance T is defined as

$$GN(r, T, S) = \{x | GD(r, x) \leq T \qquad x \in S\}. \tag{17}$$

B. The main framework of computing Grid-IGD

For an m-objective optimization problem, given M approximations $S_1, \cdots, S_M$ ($|S_1| = \cdots = |S_M| = N$), Grid-IGD is used to estimate both the convergence and diversity among the M approximations. The main procedure of computing Grid-IGD is presented in Algorithm 1. The M approximations $S_1, \cdots, S_M$, the number of objectives m and the grid neighborhood size T are given as three inputs.

The whole framework of Grid-IGD can be further divided into three major steps: 1) initialization, 2) generating reference points and 3) computing Grid-IGD for each approximation. In the following sections, each step is specified in detail.

1) Initialization: In Step 1 (line 1-3 of Algorithm 1), an efficient Divide-and-Conquer Nondominated Sorting (DCNS-F) [38] is firstly employed for obtaining the nondominated set S_{non} of all the M approximations $S_1 \bigcup S_2 \cdots \bigcup S_M$. The ideal and extended nadir points are then initialized based on Eq. (2) and (12).

2) Generating reference points: In Step 2 of the Algorithm 1, a set of reference points R is generated by taking the grid corner points of all the nondominated solutions in S_{non}. However, such a method may cause that the distribution of the generated reference points is not as extensive as that of the nondominated set S_{non}, which leads to a biased measurement of Grid-IGD, as shown in Fig. 1a.

A simple and effective way of improving it is to increase the upper bounds of the grid system by extending from z^{nad} to z^{e_nad}, which ensures that the boundary solutions are located in the new grids based on Eq. (13). It can be observed in Fig. 1b that, by such a simple adjustment, the boundary solutions A and B are located at two new grids with much closer reference points (A' and B'); meanwhile the spread of the reference points have been significantly improved.

It can be easily proved that if S_{non} is distributed in $|R|$ grids, the number of generated reference points is $|R|$ as each grid may have one and only one grid corner point.

In Step 2, the grid system is set up based on the approximation of the ideal and extended nadir point, elaborated in Section III-A. The value of the interval parameter K in such a grid system is initialized as the maximum value of its lower bound

Algorithm 1: Grid-Based IGD (Grid-IGD)

Input : M approximations: $S_1, \cdots, S_M$;
The number of objectives: m;
The grid neighborhood size: T;
Output: $Grid\text{-}IGD_1, Grid\text{-}IGD_2, \cdots, Grid\text{-}IGD_M$;

```
Step 1: Initialization:
```
1 $S_{non} = \text{NONDOMINATED\_SELECT}(S_1 \bigcup S_2 \cdots \bigcup S_M)$;
2 Approximate the ideal point z^* based on Eq.(2) and the extended nadir point z^{e_nad} based on Eq. (12);
```
Step 2: Generating reference points:
```
3 $K = max\{K_0, 1\}$;
4 **do**
5 Set up the grid system based on Eq. (13) and (14);
6 $R_t = R$;
7 $R = \emptyset$;
8 **foreach** $y \in S_{non}$ **do**
9 $R = R \cup \{z^{gcp}(y)\}$;
10 **end**
11 $K++$;
12 **while** $|R| \geq |S_{non}|/2|$;
13 **if** $||R_t| - |S_{non}|/2|| < ||R| - |S_{non}|/2||$ **then**
14 $R = R_t$;
15 **end**
```
Step 3: Computing Grid-IGD:
```
16 **for** $i = 1 \rightarrow M$ **do**
17 $Grid\text{-}IGD_i = CGI(S_i, R, m, T, z^{e\_nad})$;
18 **end**
19 **return** $Grid\text{-}IGD_1, Grid\text{-}IGD_2, \cdots, Grid\text{-}IGD_M$;

K_0 (see Section III. C) and 1. Then, the K value is adaptively determined by desirably approximate $S_{non}/2$ reference points (Line 3-12). A set of reference points R can be eventually obtained (Line 13-15).

Algorithm 2: *Computing_Grid-IGD* (CGI)

Input : The approximation: S;
The reference points: R;
The number of objectives: m;
The grid neighborhood size: T;
The extended nadir point: z^{e_nad};
Output: $Grid\text{-}IGD$;

1 Initialize $d^* = (d_1^*, \ldots, d_{|R|}^*)$;
2 **for** $j = 1 \rightarrow |R|$ **do**
3 Compute d_j^* based on Eq. (18);
4 **end**
5 $Grid\text{-}IGD(S, R) = \frac{\sum_{j=1}^{|R|} d_j^*}{|R|}$;
6 **return** $Grid\text{-}IGD$;

3) Computing Grid-IGD: In Step 3 of Algorithm 1), Grid-IGD value of each approximation S is computed by calling Algorithm 2, which can be described as follows.

In Algorithm 2, for each reference point $r \in R$, its grid neighbors $GN(r, T, S)$ are obtained based on Eq. (17). After that, the distance d^* between each reference point r and its closest neighbor $s \in GN(r, T, S)$ is computed as follows.

$$d^* = \begin{cases} \min\limits_{s \in GN(r,T,S)} dist^*(r, s), & GN(r, T, S) \neq \emptyset \\ dist^*(r, z^{e\_nad}), & GN(r, T, S) = \emptyset\ . \end{cases} \tag{18}$$

where $dist^*$ is a modified Euclidean distance, adopted for the weak Pareto compliance property [27]

$$dist^*(r, s) = \sqrt{\sum_{i=1}^{m} (max\{(s_i - r_i), 0\})^2} \tag{19}$$

$dist^*(r, z^{e\_nad})$ is the penalty adopted when no solution in S exists within the grid neighborhood of reference point r. In this case, the extended nadir point z^{e_nad} is considered to be the closest solution in S to r.

The final Grid-IGD value is the average of the distance d^* for all the reference points (line 5 of Algorithm 2). Like IGD/IGD$^+$, the smaller the Grid-IGD value, the better a PF approximation is.

IGD/IGD$^+$ estimates the absolute quality of a nondominated set by assuming all the nondominated solutions as the reference PF, which may lead to a bias measurement. Nevertheless, it is rather difficult to obtain a representative PF approximation from all the nondominated solutions with low computational cost. Alternatively, unlike IGD/IGD$^+$, Grid-IGD evaluates the relative quality of different approximations, by adopting a set of well-distributed grid corner points out as a substitute of true Pareto optimal solutions. It is worth noting that such grid corner points (reference points) are generated regarding all the PF estimates (found by all the compared algorithms), as presented in Algorithm 1.

C. Discussion on the lower bound K_0 in Grid-IGD

In this section, we estimate the lower bound K_0 of the interval parameter K by the approximation size $|S|$. With the increase of K value, the number of grids grows exponentially and the approximation tends to distribute in more different grids. If each solution in the approximation is located in a different grid, an ideally diverse set of reference points can be generated.

For an m-objective optimization problem, the maximum number of reference points would be generated when the PF approximation is extremely convex or concave and the solutions are evenly distributed at most $K^m - (K-1)^m$ grids. The approximation size $|S|$ and the interval parameter K satisfy:

$$|S| \leq K^m - (K-1)^m \tag{20}$$

Eq.(20) is a unitary high order inequality whose solutions can be obtained based on the Abel's Theorem [39].

When $m = 2, 3$ and K satisfies:

$$\begin{cases} K \geq \lceil \frac{|S|+1}{2} \rceil, & m = 2 \\ K \geq \lceil \frac{1}{2} + \frac{\sqrt{3(4|S|-1)}}{6} \rceil, & m = 3 \end{cases} \tag{21}$$

When $m \geq 4$ and the range of K can be determined by using the number line method to solve the inequality as follows.

$$\begin{cases} \sum_{i=1}^{m} K_i = \frac{\binom{m+1}{2}}{m+1} \\ \sum_{i=1}^{m-1} \prod_{j=i+1}^{m} K_i K_j = \frac{\binom{m+1}{3}}{m+1} \\ \cdots \\ \prod_{i=1}^{m} K_i = (-1)^{m+1} \frac{(-1)^{m+1} + |S|}{m+1} \end{cases} \tag{22}$$

The lower bound K_0 for obtaining $|S|$ reference points can be estimated by Eqs. (20)- (22).

D. Computational Complexity of Grid-IGD

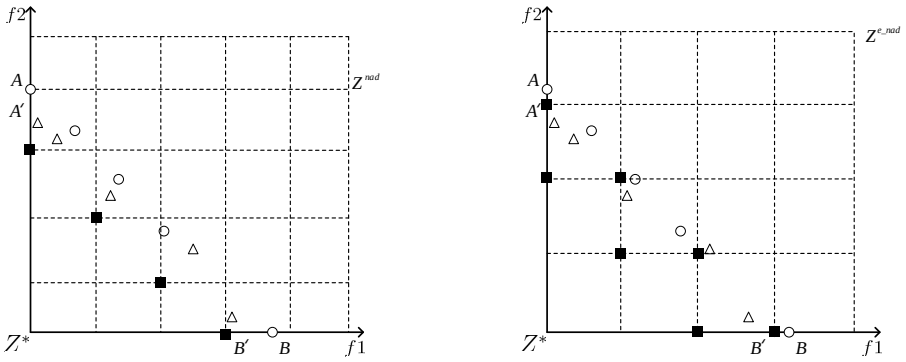
The average complexity of DCNS-F in Step 1 of Algorithm 1 is $O(Nlog^{m-1}N)$. Step 2 of Algorithm 1 at most requires $O(mKN/2)$ ($K \ll N$) to approximate the reference points. In Step 3, the average number of solutions in each subset is $\frac{N}{Q}$, where Q denotes the number of grids which contain solutions. Therefore, Step 3 on average requires $O(mTN)$ computations. In summary, the time complexity of Algorithm 1 is $\max\{O(Nlog^{m-1}N), O(mTN)\}$ ($T \ll N$), where N is the approximation size and m is the number of objectives. The computational complexity of Grid-IGD is obviously less than that of IGD, which needs quadratic time.

E. Weak Pareto compliance of Grid-IGD

Compared with IGD, one of the importantly desirable property of Grid-IGD is its weak Pareto compliance, as follows.

Theorem 1: Grid-IGD is a weakly Pareto Compliant indicator: if and only if $\forall A, B : A \preceq B \implies I(A) \leq I(B)$.

The proof of this theorem can be found in the supplementary material.



(a) The distribution of the generated reference points before extending the grid system.

(b) The distribution of the generated reference points after extending the grid system.

Fig. 1: An illustration of the reference points, generated by the approximations in a bi-objective optimization problem. The open circles represent one approximation solution set and the open triangles represent the another one, the solid squares are the reference points generated by the nondominated solutions of the two approximations.

IV. Experiments and Discussions

In this section, the following experiments are conducted for

- validating Grid-IGD on artificial PFs;
- validating Grid-IGD on real PFs;
- validating Grid-IGD on MaOPs with irregular PFs;
- investigating the effects of the interval parameter K on Grid-IGD;
- investigating the effects of maximum neighborhood size T on Grid-IGD;
- investigating the effects of the uniformity of reference points on IGD/IGD$^+$/Grid-IGD;

A. Experimental Setups

In Section IV. B, the size N of each artificial PF approximation ($m = 3$) is set to 120, as shown in Table II.

In Section IV. C-G, Grid-IGD is adopted to evaluate the real PF approximations, delivered by five many-objective optimizers (MOEA/D-DE [40], IBEA [41], PAES [42], GrEA [37] and NSGA-III [43]). The experimental setups for the PF approximations obtained by five many-objective optimizers are listed in Table II.

For some optimizers (e.g., MOEA/D-DE [40] and NSGA-III [43]), the population size is equal to the size of the reference vector set. They can be obtained by uniformly sampling on a unit simplex, which is usually called Das and Dennis's systematic approach [44]. In this case, the reference vector size is

$$N = \binom{m-1}{H+m-1} \tag{23}$$

where $H > 0$ is the number of divisions along each objective coordinate and m is the number of objectives. However, as pointed in [43], the direct use of Das and Dennis's approach may not be appropriate to use when $m > 6$. Instead, a two-layer direction vector generation method [43], [45] has been adopted for MaOPs with more than 6 objectives.

In all the experimental studies, each algorithm was run 30 times on each benchmark problem and the approximation with the median Grid-IGD value is used for comparisons. All the compared algorithms are implemented based on the open source MOEA platform, JMetal [46].

To verify the effectiveness of Grid-IGD, IGD [15], [16], (IGD$^+$) [27], IGD-NS [47] and HV [13] have also been adopted as the benchmark indicators for comparison in the experimental studies. For HV, the reference point is set to 1.1 times of the maximum values of PF approximations.

For Grid-IGD, when the maximum neighborhood size T is mK, where the complexity of Grid-IGD is the same as the original IGD. However, in the experiments, the value of T is set to 24 based on its sensitivity test in Section IV. F.

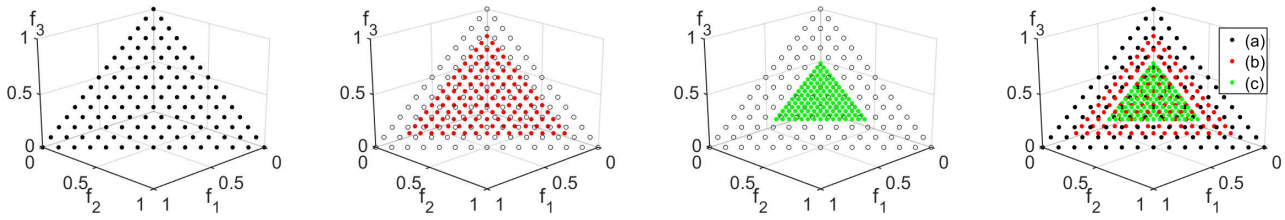
TABLE II: The approximation size with the different number of objectives in the experimental studies.

the number of objectives m	3	5	10
the approximation size N	120	126	276

B. Grid-IGD on Artificial PF Approximations

Three groups of artificial approximations are generated and presented in Fig. 2-Fig. 4. In the first group, the approximations are uniformly distributed though their spread is quite different from each other. As shown in Fig. 2a, the values of all the solutions for each objective range in $[0, 1]$. As shown in Fig. 2b, the values of all the solutions for each objective range in $[0.1, 0.8]$. As shown in Fig. 2c, the values of all the solutions for each objective range in $[0.2, 0.6]$. In addition, such three approximations are all located on the same hyperplane $f_1 + f_2 + f_3 = 1$, as shown in Fig. 2d. In the aforementioned figures, the wider distribution of an approximation, the smaller Grid-IGD value it has. The above observations indicate that Grid-IGD can accurately reflect the spread of PF approximations.

In the second group, three approximations are distributed over the entire PF, although their uniformities are different from each other. These three approximations are generated in the following way. First, the solutions are uniformly generated on the plane $f_1 + f_2 + f_3 = 1$, ranging in $[0, 1]$, as shown in Fig. 3a. After that, when 50% or 100% of solutions are replaced with the randomly generated solutions on the same plane, different solution sets can be obtained, as shown in Fig. 3b and 3c. As shown in Fig. 3a- 3c, the solution sets with



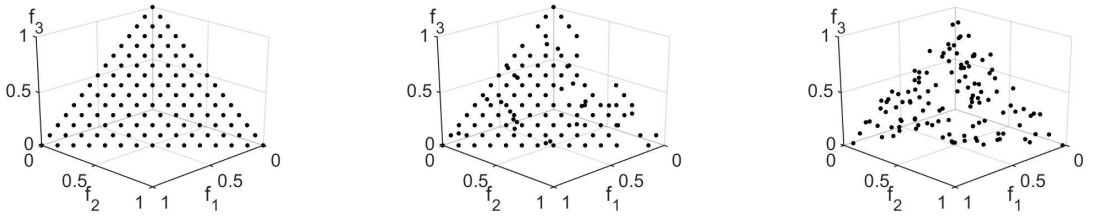
(a) Every objective value of the PF approximation ranges in $[0, 1]$ and its Grid-IGD or is 0.0408505

(b) Every objective value of the PF approximation ranges in $[0.1, 0.8]$ and its Grid-IGD is 0.0662420

(c) Every objective value of the PF approximation ranges in $[0.2, 0.6]$ and its Grid-IGD is 0.1407865

(d) The PF approximations in Fig 2a, 2b and 2c are located on the hyperplane $f_1 + f_2 + f_3 = 1$

Fig. 2: The artificial PF approximations with different spread, uniformly located on the hyperplane $f_1 + f_2 + f_3 = 1$

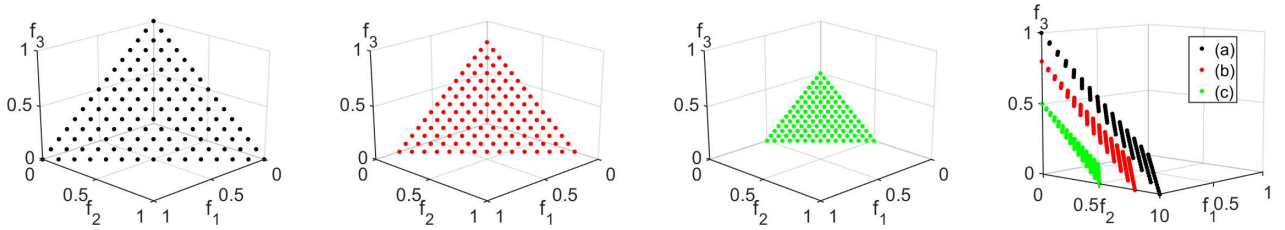


(a) Each objective value of all the PF approximations ranges in $[0, 1]$ and its Grid-IGD is 0.0394702

(b) With 50% of solutions replaced by randomly generated solutions on the same plane and its Grid-IGD is 0.0466158

(c) With 100% of solutions replaced by randomly generated solutions on the same plane and its Grid-IGD is 0.0579717

Fig. 3: The artificial approximations with various uniformity, located on a same hyperplane $f_1 + f_2 + f_3 = 1$



(a) PF approximation uniformly located on the hyperplane $f_1 + f_2 + f_3 = 1$ and its Grid-IGD is 0.2926927

(b) PF approximation uniformly located on the hyperplane $f_1 + f_2 + f_3 = 0.8$ and its Grid-IGD is 0.1771430

(c) PF approximation uniformly located on the hyperplane $f_1 + f_2 + f_3 = 0.5$ and its Grid-IGD is 0.0016108

(d) The PF approximations in Fig 4a, 4b and 4c are uniformly located on different hyperplanes

Fig. 4: The artificial approximations uniformly located on different hyperplanes

the worse uniformity have larger Grid-IGD values. Clearly, Grid-IGD is able to correctly reflect the uniformity of PF approximations.

In the third group, all the approximations are uniformly distributed but with different convergence levels. Fig. 4a - 4c shows such three approximations, uniformly distributed on the hyperplanes $f_1 + f_2 + f_3 = 1$, $f_1 + f_2 + f_3 = 0.8$ and $f_1 + f_2 + f_3 = 0.5$, respectively. It can be observed that Grid-IGD values decrease with better convergence levels for the approximations. Obviously, Grid-IGD can correctly reflect approximations' convergence performance.

C. Grid-IGD on real PF approximations

DTLZ2 [48], whose PF is regular, has been selected as a representative benchmark problem for verifying the effectiveness of Grid-IGD. For DTLZ2, it is relatively easy for all the optimizers to converge to its PF. This characteristic can help test the diversity performance of approximations obtained by the different algorithms.

The performance of PF approximations obtained by five optimizers on m-objective DTLZ2, in terms of IGD/IGD$^+$/Grid-IGD/IGD-NS/HV values and their corresponding ranks, are given in Table III. Due to the page limit, PF approximations or parallel coordinate plots on DTLZ2 are put in the supplementary material.

It can be observed from Table III that the ranks of the approximations obtained by the five compared algorithms, in terms of Grid-IGD values, are always consistent with that in terms of IGD$^+$ values. This indicates that Grid-IGD is able to correctly reflect the relative performance of different algorithms on problems with regular PFs. Another interesting observation is that the ranks of the approximations in terms of IGD/IGD-NS are different from that of IGD$^+$/Grid-IGD. This can be explained by the fact that IGD/IGD-NS is not Pareto compliant. The ranks of approximations in terms of HV indicator are the same as that of IGD$^+$/Grid-IGD on tri- and 5-objective DTLZ2, but different on 10-objective DTLZ2. The rank differences are produced by the PF approximations

TABLE III: IGD, IGD$^+$, Grid-IGD, IGD-NS and HV values for PF approximations obtained by five algorithms on m-objective DTLZ2. Their ranks in terms of IGD/IGD$^+$/Grid-IGD /IGD-NS/HV are presented in the parenthesis.

instances (m-obj)	indicator	MOEA/D-DE	IBEA	PAES	GrEA	NSGA-III
DTLZ2 (3-obj)	IGD	$6.355E-02$ (2)	$8.259E-02$ (4)	$2.972E-01$ (5)	$7.718E-02$ (3)	$4.679E-02$ (1)
	IGD$^+$	$3.199E-02$ (4)	$1.872E-02$ (1)	$1.961E-01$ (5)	$3.016E-02$ (3)	$1.964E-02$ (2)
	Grid-IGD	$6.868E-02$ (4)	$5.263E-02$ (1)	$2.116E-01$ (5)	$6.138E-02$ (3)	$5.521E-02$ (2)
	IGD-NS	$6.272E+02$ (2)	$8.152E+02$ (4)	$2.933E+03$ (5)	$7.618E+02$ (3)	$4.618E+02$ (1)
	HV	$7.143E-01$ (4)	$7.572E-01$ (1)	$4.741E-01$ (5)	$7.248E-01$ (3)	$7.530E-01$ (2)
DTLZ2 (5-obj)	IGD	$3.251E-01$ (4)	$2.359E-01$ (3)	$6.748E-01$ (5)	$1.991E-01$ (2)	$1.949E-01$ (1)
	IGD$^+$	$1.190E-01$ (4)	$6.325E-02$ (1)	$4.552E-01$ (5)	$7.677E-02$ (3)	$7.192E-02$ (2)
	Grid-IGD	$1.978E-01$ (4)	$1.550E-01$ (1)	$5.229E-01$ (5)	$1.681E-01$ (3)	$1.613E-01$ (2)
	IGD-NS	$2.880E+03$ (4)	$2.089E+03$ (3)	$5.977E+03$ (5)	$1.763E+03$ (2)	$1.726E+03$ (1)
	HV	$1.138E-00$ (4)	$1.301E-00$ (1)	$4.308E-01$ (5)	$1.275E-00$ (3)	$1.280E-00$ (2)
DTLZ2 (10-obj)	IGD	$4.701E-01$ (2)	$7.900E-01$ (5)	$6.806E-01$ (4)	$5.001E-01$ (3)	$4.222E-01$ (1)
	IGD$^+$	$2.794E-01$ (2)	$3.642E-01$ (4)	$4.552E-01$ (5)	$2.950E-01$ (3)	$1.794E-01$ (1)
	Grid-IGD	$3.618E-01$ (2)	$4.424E-01$ (4)	$5.487E-01$ (5)	$3.849E-01$ (3)	$2.310E-01$ (1)
	IGD-NS	$3.300E+03$ (2)	$5.537E+03$ (5)	$4.784E+03$ (4)	$3.511E+03$ (3)	$2.958E+03$ (1)
	HV	$2.177E-00$ (2)	$2.089E-00$ (3)	$1.294E-00$ (5)	$1.903E-00$ (4)	$2.511E-00$ (1)

of IBEA and GrEA. It can be observed from the parallel coordinate plots on 10-objective DTLZ2 (Fig. 3 in the supplementary material) that IBEA has better performance in terms of convergence but worse performance in terms of diversity, than GrEA. Apparently, although IGD$^+$/Grid-IGD and HV are all Pareto compliant, different from HV, IGD$^+$/Grid-IGD has more preference towards diversity.

D. Grid-IGD on MaOPs with irregular PFs

The effectiveness of Grid-IGD on MaOPs with the irregular PFs is further verified in this section. DTLZ7 [48] contains a typical irregular PF, consisting of 2^{m-1} disconnected segments, which can be either convex or concave. UF9 [49] is a tri-objective problem with linear and discontinuous PF. WFG2 [50] has a disconnected PF, which is scalable in the number of objectives.

The performance of PF approximations obtained by five optimizers on various problems, in terms of IGD/IGD$^+$/Grid-IGD/IGD-NS/HV values and their corresponding ranks, are given in Table IV. Due to the page limit, their PF approximations or parallel coordinate plots are put in the supplementary material.

It can be observed from Table IV that the ranks of the approximations obtained by the five compared algorithms, in terms of Grid-IGD values, are always consistent with that in terms of IGD$^+$ values. This indicates that Grid-IGD is able to correctly reflect the relative performance of different algorithms on problems with irregular PFs. Another interesting observation is that the ranks of the approximations in terms of IGD/IGD-NS are different from that of IGD$^+$/Grid-IGD on most problems. This can be explained by the fact that neither IGD nor IGD-NS is Pareto compliant. In addition, the ranks of the approximations in terms of HV indicator are the same as that of IGD$^+$/Grid-IGD on 3-objective problems, but different on 5 and 10-objective problems, due to the fact that HV and IGD$^+$/Grid-IGD have different preferences towards convergence or diversity.

E. Effects of interval parameter K on Grid-IGD

In this section, we investigate the effects of interval parameter K on Grid-IGD. With the increase of the K values, the number of grids grows exponentially and the solutions of an approximation tend to distribute in more different grids, leading to an increasing number of reference points. Apparently, the setting of parameter K can be determined by the final number of reference points. Eqs. (20)- (22) only estimates the lower bounds K_0 of K values for keeping $|S|$ reference points, given m-dimensional PF approximations.

Fig. 6 shows the Grid-IGD values for different approximations with different K values in the range of $[2, 100]$ for various problems. It further verifies the effects of K on Grid-IGD as follows.

When K values are too small, Grid-IGD generates too few reference points to conduct fair measurements for different PF approximations. However, an extremely large K value also seems unnecessary as each solution of all the approximations will be located in a different grid and the number of reference points is equal to the number of solutions for all the approximations. Under this circumstance, Grid-IGD is equivalent to IGD$^+$ [27] using all the nondominated solutions as a reference set. It can be observed from Fig. 6 that K is indeed a problem-dependent parameter. However, when K is set within a certain value range, a representative set of enough diversely-populated reference points can be generated by Grid-IGD for correctly ranking different PF approximations, as shown in Fig. 6, Table III and Table IV.

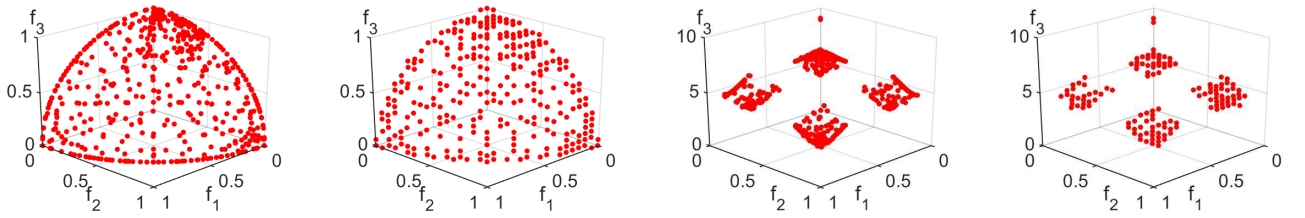
A simple method for tuning the appropriate K value for different problems is based on the number of reference points generated by Grid-IGD, as presented in Algorithm 1. Table V demonstrates the estimated lower bound K_0 obtained by Eq. (20) - (22), the actual number of reference points and the corresponding K values obtained by Grid-IGD on different problems with m-objectives in the experimental studies.

F. Sensitivity test of T on Grid-IGD

In this section, the sensitivity tests on the neighborhood size T are conducted. Fig. 7 shows the grid-IGD values of the approximations obtained by five MOEAs with different T values on various problems. In the experiments, T values range in $[1, mK]$, where K is the number of intervals on each objective and m is the number of objectives. The maximum value of T is set to mK, as mK is the Manhattan distance

TABLE IV: IGD, IGD$^+$, IGD-NS, HV and Grid-IGD values for PF approximations obtained by five algorithms on problems with irregular PFs with different number of objectives m. Their ranks in terms of IGD/IGD$^+$/IGD-NS/HV/Grid-IGD are presented in the parenthesis.

instance (m-obj)	indicator	MOEA/D-DE	IBEA	PAES	GrEA	NSGA-III
DTLZ7 (3-obj)	IGD	$2.147E-01$ (4)	$1.100E-01$ (3)	$1.150E+00$ (5)	$1.032E-01$ (2)	$6.640E-02$ (1)
	IGD$^+$	$9.515E-02$ (4)	$2.584E-02$ (1)	$1.013E+00$ (5)	$4.233E-02$ (3)	$3.631E-02$ (2)
	Grid-IGD	$1.116E-01$ (4)	$5.428E-02$ (1)	$1.242E+00$ (5)	$7.642E-02$ (3)	$6.335E-02$ (2)
	IGD-NS	$2.187E+03$ (4)	$1.100E+03$ (3)	$1.150E+04$ (5)	$1.032E+03$ (2)	$6.640E+02$ (1)
	HV	$1.189E+00$ (4)	$1.632E+00$ (1)	$0.865E+00$ (5)	$1.545E+00$ (3)	$1.623E+00$ (2)
DTLZ7 (5-obj)	IGD	$3.192E+00$ (5)	$1.479E+00$ (3)	$2.205E+00$ (4)	$2.776E-01$ (1)	$3.231E-01$ (2)
	IGD$^+$	$8.905E-01$ (3)	$1.187E+00$ (4)	$2.000E+00$ (5)	$1.116E-01$ (1)	$1.878E-01$ (2)
	Grid-IGD	$1.033E+00$ (3)	$1.303E+00$ (4)	$1.931E+00$ (5)	$1.961E-01$ (1)	$2.763E-01$ (2)
	IGD-NS	$3.192E+04$ (5)	$1.480E+04$ (3)	$2.206E+04$ (4)	$2.776E+03$ (1)	$3.232E+03$ (2)
	HV	$9.910E-03$ (5)	$1.516E+00$ (3)	$1.363E+00$ (4)	$2.256E+00$ (1)	$2.054E+00$ (2)
DTLZ7 (10-obj)	IGD	$2.606E+00$ (3)	$5.928E+00$ (5)	$4.054E+00$ (4)	$8.550E-01$ (1)	$1.285E+00$ (2)
	IGD$^+$	$1.492E+00$ (3)	$5.747E+00$ (5)	$3.860E+00$ (4)	$6.501E-01$ (1)	$9.882E-01$ (2)
	Grid-IGD	$1.564E+00$ (3)	$5.474E+00$ (5)	$3.798E+00$ (4)	$8.759E-01$ (1)	$1.514E+00$ (2)
	IGD-NS	$5.142E+04$ (3)	$1.167E+05$ (5)	$7.982E+04$ (4)	$1.683E+04$ (1)	$2.537E+04$ (2)
	HV	$1.815E-05$ (5)	$1.552E+00$ (3)	$2.208E+00$ (2)	$1.229E+00$ (4)	$2.303E+00$ (1)
UF9 (3-obj)	IGD	$3.123E-01$ (1)	$3.936E-01$ (3)	$6.836E-01$ (5)	$3.950E-01$ (4)	$3.818E-01$ (2)
	IGD$^+$	$3.099E-01$ (1)	$3.423E-01$ (2)	$6.780E-01$ (5)	$3.744E-01$ (3)	$3.764E-01$ (4)
	Grid-IGD	$8.922E-02$ (1)	$1.850E-01$ (2)	$4.949E-01$ (5)	$2.139E-01$ (3)	$2.300E-01$ (4)
	IGD-NS	$1.617E+03$ (3)	$2.012E+03$ (2)	$3.500E+03$ (5)	$1.998E+03$ (1)	$1.964E+03$ (4)
	HV	$5.877E-01$ (1)	$5.061E-01$ (2)	$1.226E-01$ (5)	$4.717E-01$ (3)	$4.514E-01$ (4)
WFG2 (3-obj)	IGD	$5.161E-01$ (5)	$2.396E-01$ (3)	$2.598E-01$ (4)	$2.101E-01$ (2)	$1.737E-01$ (1)
	IGD$^+$	$1.956E-01$ (4)	$3.840E-02$ (1)	$2.003E-01$ (5)	$5.902E-02$ (2)	$8.630E-02$ (3)
	Grid-IGD	$2.335E-01$ (4)	$1.269E-02$ (1)	$2.802E-01$ (5)	$1.479E-02$ (2)	$1.705E-01$ (3)
	IGD-NS	$1.344E+03$ (5)	$6.231E+02$ (3)	$6.795E+02$ (4)	$5.463E+02$ (2)	$4.517E+02$ (1)
	HV	$5.466E+02$ (4)	$5.912E+02$ (1)	$5.414E+02$ (5)	$5.864E+02$ (2)	$5.827E+02$ (3)
WFG2 (5-obj)	IGD	$1.577E-00$ (5)	$3.398E-01$ (4)	$8.015E-01$ (2)	$1.073E-00$ (3)	$7.357E-01$ (1)
	IGD$^+$	$5.461E-01$ (4)	$1.332E-01$ (1)	$5.666E-01$ (5)	$1.425E-01$ (2)	$1.655E-01$ (3)
	Grid-IGD	$6.274E-01$ (4)	$2.659E-01$ (1)	$7.706E-01$ (5)	$3.351E-01$ (2)	$3.358E-01$ (3)
	IGD-NS	$9.498E+03$ (5)	$5.662E+03$ (3)	$4.865E+03$ (2)	$6.460E+03$ (4)	$4.444E+03$ (1)
	HV	$5.570E+03$ (4)	$5.969E+03$ (1)	$4.839E+03$ (5)	$5.788E+03$ (3)	$5.840E+03$ (2)
WFG2 (10-obj)	IGD	$3.514E-00$ (4)	$5.294E-00$ (5)	$2.567E-00$ (1)	$2.864E-00$ (2)	$3.202E-00$ (3)
	IGD$^+$	$3.675E-01$ (4)	$1.560E-01$ (1)	$7.099E-01$ (5)	$3.253E-01$ (3)	$3.127E-01$ (2)
	Grid-IGD	$1.028E+00$ (4)	$3.586E-01$ (1)	$1.796E+00$ (5)	$5.888E-01$ (3)	$5.424E-01$ (2)
	IGD-NS	$6.971E+04$ (4)	$1.048E+05$ (5)	$5.165E+04$ (1)	$5.721E+04$ (2)	$6.381E+04$ (3)
	HV	$9.204E+09$ (2)	$9.381E+09$ (1)	$6.644E+09$ (5)	$8.925E+09$ (3)	$8.891E+09$ (4)



(a) All the non-dominated solutions as a reference set on DTLZ2.

(b) All the grid corner points in a grid system as a reference set on DTLZ2.

(c) All the non-dominated solutions as a reference set on DTLZ7.

(d) All the grid corner points in a grid system as a reference set on DTLZ7.

Fig. 5: Grid corner points (reference sets) obtained by Grid-IGD on tri-objective DTLZ2 or DTLZ7.

TABLE V: The estimated lower bounds K_0 based on Eq. (20) - (22), the number of reference points obtained by Grid-IGD and the corresponding K values on different problems with m-objectives in the experimental studies.

	DTLZ2			DTLZ7			UF9	WFG2		
the number of objectives m	3	5	10	3	5	10	3	3	5	10
the lower bound K_0 (Eq. 20- 22)	7	3	2	7	3	2	7	7	3	2
obtained K values in the experiments	17	7	5	40	16	9	26	20	12	12
the number of reference points	276	306	603	275	278	683	110	153	221	492
the size of nondominated set S_{non}	554	592	1073	533	571	1297	228	281	423	938

TABLE VI: IGD, IGD$^+$ and Grid-IGD values for PF approximations obtained by five algorithms with three different sets of reference points on m-objective DTLZ2. Their ranks in terms of IGD/IGD$^+$/Grid-IGD are presented in the parenthesis.

indicator (m-obj)	reference set	MOEA/D-DE	IBEA	PAES	GrEA	NSGA-III
IGD (3-obj)	uniformly sampled true PF	$6.355E-02$ (2)	$8.259E-02$ (4)	$2.972E-01$ (5)	$7.718E-02$ (3)	$4.679E-02$ (1)
	all the nondominated solutions (Fig. 5a)	$5.878E-02$ (4)	$5.569E-02$ (3)	$2.467E-01$ (5)	$5.319E-02$ (2)	$3.216E-02$ (1)
	grid corner points (Fig. 5b)	$8.168E-02$ (3)	$8.244E-02$ (4)	$2.855E-01$ (5)	$8.120E-02$ (2)	$6.233E-02$ (1)
IGD$^+$ (3-obj)	uniformly sampled true PF	$3.199E-02$ (4)	$1.872E-02$ (1)	$1.961E-01$ (5)	$3.016E-02$ (3)	$1.964E-02$ (2)
	all the nondominated solutions (Fig. 5a)	$3.182E-02$ (4)	$1.266E-02$ (2)	$1.632E-01$ (5)	$1.974E-02$ (3)	$1.468E-02$ (2)
	grid corner points (Fig. 5b)	$6.868E-02$ (4)	$5.263E-02$ (1)	$2.116E-01$ (5)	$6.138E-02$ (3)	$5.521E-02$ (2)
Grid-IGD (3-obj)	grid corner points (Fig. 5b)	$6.868E-02$ (4)	$5.263E-02$ (1)	$2.116E-01$ (5)	$6.138E-02$ (3)	$5.521E-02$ (2)
IGD (5-obj)	uniformly sampled true PF	$3.251E-01$ (4)	$2.359E-01$ (3)	$6.748E-01$ (5)	$1.991E-01$ (2)	$1.949E-01$ (1)
	all the nondominated solutions	$2.210E-02$ (4)	$1.365E-02$ (2)	$5.931E-01$ (5)	$1.196E-02$ (1)	$1.269E-02$ (3)
	grid corner points	$2.785E-01$ (4)	$2.091E-01$ (3)	$6.648E-01$ (5)	$1.984E-01$ (2)	$1.840E-01$ (1)
IGD$^+$ (5-obj)	uniformly sampled true PF	$1.190E-01$ (4)	$6.325E-02$ (1)	$4.552E-01$ (5)	$7.677E-02$ (3)	$7.192E-02$ (2)
	all the nondominated solutions	$8.222E-02$ (4)	$3.762E-02$ (1)	$4.167E-01$ (5)	$5.399E-02$ (2)	$5.623E-02$ (3)
	grid corner points	$1.978E-01$ (4)	$1.550E-01$ (1)	$5.229E-01$ (5)	$1.681E-01$ (3)	$1.613E-01$ (2)
Grid-IGD (5-obj)	grid corner points	$1.978E-01$ (4)	$1.550E-01$ (1)	$5.229E-01$ (5)	$1.681E-01$ (3)	$1.613E-01$ (2)
IGD (10-obj)	uniformly sampled true PF	$4.701E-01$ (2)	$7.900E-01$ (5)	$6.806E-01$ (4)	$5.001E-01$ (3)	$4.222E-01$ (1)
	all the nondominated solutions	$2.608E-01$ (2)	$6.354E-01$ (5)	$5.502E-01$ (4)	$4.227E-01$ (3)	$2.335E-01$ (1)
	grid corner points	$4.280E-01$ (2)	$7.296E-01$ (5)	$7.087E-01$ (4)	$5.134E-01$ (3)	$3.038E-01$ (1)
IGD$^+$ (10-obj)	uniformly sampled true PF	$2.794E-01$ (2)	$3.642E-01$ (4)	$4.552E-01$ (5)	$2.950E-01$ (3)	$1.794E-01$ (1)
	all the nondominated solutions	$1.579E-01$ (2)	$2.228E-01$ (4)	$3.295E-01$ (5)	$2.140E-01$ (3)	$5.990E-02$ (1)
	grid corner points	$3.618E-01$ (2)	$4.424E-01$ (4)	$5.487E-01$ (5)	$3.849E-01$ (3)	$2.310E-01$ (1)
Grid-IGD (10-obj)	grid corner points	$3.618E-01$ (2)	$4.424E-01$ (4)	$5.487E-01$ (5)	$3.849E-01$ (3)	$2.310E-01$ (1)

TABLE VII: IGD, IGD$^+$ and Grid-IGD values for PF approximations obtained by five algorithms with three different sets of reference points on m-objective DTLZ7. Their ranks in terms of IGD/IGD$^+$/Grid-IGD are presented in the parenthesis.

indicator (m-obj)	reference set	MOEA/D-DE	IBEA	PAES	GrEA	NSGA-III
IGD (3-obj)	uniformly sampled true PF	$2.147E-01$ (4)	$1.100E-01$ (3)	$1.150E+00$ (5)	$1.032E-01$ (2)	$6.640E-02$ (1)
	all the nondominated solutions (Fig. 5c)	$2.779E-01$ (4)	$8.744E-02$ (2)	$1.146E+00$ (5)	$9.262E-01$ (3)	$6.915E-02$ (1)
	grid corner points (Fig. 5d)	$2.031E-01$ (4)	$1.152E-01$ (2)	$1.380E+00$ (5)	$1.288E-01$ (3)	$9.650E-02$ (1)
IGD$^+$ (3-obj)	uniformly sampled true PF	$9.515E-02$ (4)	$2.584E-02$ (1)	$1.013E+00$ (5)	$4.233E-02$ (3)	$3.631E-02$ (2)
	all the nondominated solutions (Fig. 5c)	$1.099E-01$ (4)	$1.760E-02$ (1)	$1.012E+00$ (5)	$3.089E-02$ (3)	$2.810E-02$ (2)
	grid corner points (Fig. 5d)	$1.116E-01$ (4)	$5.428E-02$ (1)	$1.242E+00$ (5)	$7.642E-02$ (3)	$6.335E-02$ (2)
Grid-IGD (3-obj)	grid corner points (Fig. 5d)	$1.116E-01$ (4)	$5.428E-02$ (1)	$1.242E+00$ (5)	$7.642E-02$ (3)	$6.335E-02$ (2)
IGD (5-obj)	uniformly sampled true PF	$3.192E+00$ (5)	$1.479E+00$ (3)	$2.205E+00$ (4)	$2.776E-01$ (1)	$3.231E-01$ (2)
	all the nondominated solutions	$3.753E+00$ (5)	$1.346E+00$ (3)	$1.833E+00$ (4)	$1.970E-01$ (1)	$2.220E-01$ (2)
	grid corner points	$3.397E+00$ (5)	$1.486E+00$ (3)	$2.070E+00$ (4)	$2.529E-01$ (1)	$3.104E-01$ (2)
IGD$^+$ (5-obj)	uniformly sampled true PF	$8.905E-01$ (3)	$1.187E+00$ (4)	$2.000E+00$ (5)	$1.116E-01$ (1)	$1.878E-01$ (2)
	all the nondominated solutions	$1.021E+00$ (3)	$1.183E+00$ (4)	$1.689E+00$ (5)	$1.189E-01$ (1)	$1.826E-01$ (2)
	grid corner points	$1.033E+00$ (3)	$1.303E+00$ (4)	$1.931E+00$ (5)	$1.961E-01$ (1)	$2.763E-01$ (2)
Grid-IGD (5-obj)	grid corner points	$1.033E+00$ (3)	$1.303E+00$ (4)	$1.931E+00$ (5)	$1.961E-01$ (1)	$2.763E-01$ (2)
IGD (10-obj)	uniformly sampled true PF	$2.606E+00$ (3)	$5.928E+00$ (5)	$4.054E+00$ (4)	$8.550E-01$ (1)	$1.285E+00$ (2)
	all the nondominated solutions	$4.000E-00$ (3)	$5.318E-00$ (5)	$4.126E+00$ (4)	$1.130E-00$ (1)	$2.142E-00$ (2)
	grid corner points	$3.275E+00$ (3)	$5.593E+00$ (5)	$3.917E+00$ (4)	$1.020E+00$ (1)	$1.657E+00$ (2)
IGD$^+$ (10-obj)	uniformly sampled true PF	$1.492E+00$ (3)	$5.747E+00$ (5)	$3.860E+00$ (4)	$6.501E-01$ (1)	$9.882E-01$ (2)
	all the nondominated solutions	$1.355E+00$ (2)	$5.173E+00$ (5)	$3.995E+00$ (4)	$9.411E-01$ (1)	$1.959E+00$ (3)
	grid corner points	$1.564E+00$ (3)	$5.474E+00$ (5)	$3.798E+00$ (4)	$8.759E-01$ (1)	$1.514E+00$ (2)
Grid-IGD (10-obj)	grid corner points	$1.564E+00$ (3)	$5.474E+00$ (5)	$3.798E+00$ (4)	$8.759E-01$ (1)	$1.514E+00$ (2)

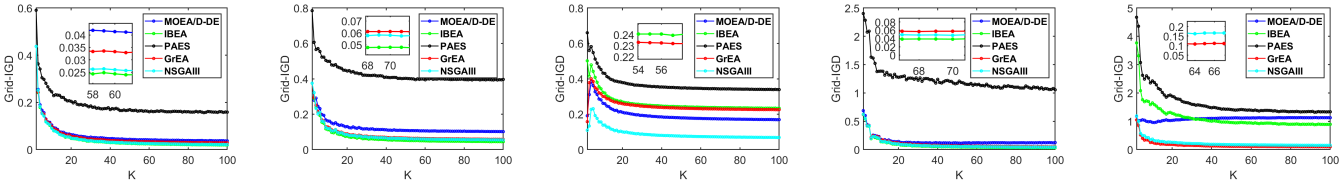
between the extended nadir point $z^{e\text{-}nad}$ and the ideal point z^*, which is the maximum possible distance between a reference point and a solution.

It can be observed from Fig. 7 that the absolute Grid-IGD values of the approximations obtained by different multiobjective optimizers reach their maximum values after certain large threshold values of T. However, the ranks of different optimizers in terms of the Grid-IGD values remain the same after certain threshold values of T on different problems. It can be observed in Fig. 7 that the Grid-IGD remains stable when $T = m$ for most problems and $T = 24$ for all the tested problems.

G. Effects of uniformity of reference points on IGD/IGD$^+$ /Grid-IGD

One major difference between IGD/IGD$^+$ and Grid-IGD is the distribution of the reference points. For IGD/IGD$^+$, the nondominated solutions of all the PF approximations are considered as a set of reference points when true PFs are unknown *a priori*, which is a very common scenario for the real-world multi/many-objective optimization problems. For Grid-IGD, the reference set is more uniformly distributed due to the use of the grid system. In this section, Grid-IGD is compared with IGD/IGD$^+$ and the effects of the distributions of the reference set on them are also discussed.

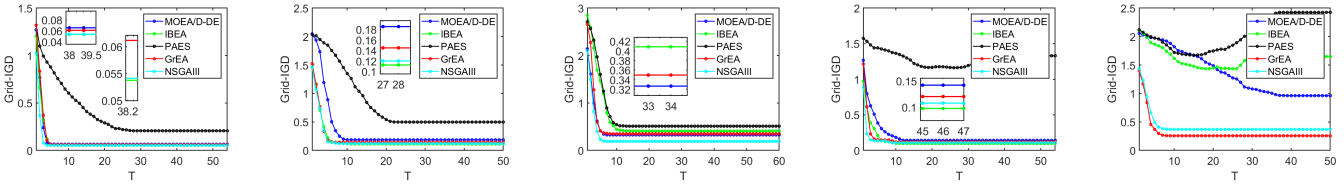
The nondominated solutions of all the PF approximations on tri-objective DTLZ2 are considered as a set of reference points, as shown in Fig. 5a. It can be observed that the obtained



(a) 3-objective DTLZ2 (b) 5-objective DTLZ2 (c) 10-objective DTLZ2 (d) 3-objective DTLZ7 (e) 5-objective DTLZ7

(f) 10-objective DTLZ7 (g) 3-objective WFG2 (h) 5-objective WFG2 (i) 10-objective WFG2 (j) 3-objective UF9

Fig. 6: Grid-IGD values of the approximations obtained by five MOEAs with different K values.



(a) 3-objective DTLZ2 (b) 5-objective DTLZ2 (c) 10-objective DTLZ2 (d) 3-objective DTLZ7 (e) 5-objective DTLZ7

(f) 10-objective DTLZ7 (g) 3-objective WFG2 (h) 5-objective WFG2 (i) 10-objective WFG2 (j) 3-objective UF9

Fig. 7: Grid-IGD values of the approximations obtained by five MOEAs with different T values.

reference set is not uniformly distributed over the entire PF. On the contrary, the reference set obtained by the grid corner points in Grid-IGD ($K = 18$) is more uniformly distributed, as shown in Fig. 5b. Similarly, all the nondominated solutions obtained by five algorithms on tri-objective DTLZ7 are presented in Fig. 5c and its grid corner points are given in Fig. 5d.

To investigate the effects of the distributions of the reference sets, IGD/IGD$^+$ using all the nondominated solutions as a reference set, IGD/IGD$^+$ using the uniformly sampled true PFs as a reference set, IGD/IGD$^+$ using the grid corner points as a reference set and the proposed Grid-IGD are compared in terms of the ranks of five algorithms on DTLZ2 and DTLZ7, as shown in Table VI and Table VII, respectively.

It can be observed that the ranks of five compared algorithms in terms of IGD$^+$ using the uniformly sampled true PFs as a reference set, IGD$^+$ using the grid corner points as a reference set and Grid-IGD are always consistent with each other. However, five algorithms may have different ranks in terms of IGD or IGD$^+$ using all the nondominated solutions as a reference set.

We can have the following observations based on Table VI and Table VII.

- Using all the nondominated solutions as a reference set is not a good strategy for IGD-based indicators, as the uniformity of a reference set is of great importance for fair performance evaluations of PF approximations.
- Without knowing the true PF *a priori*, Grid-IGD can still provide a set of uniformly distributed reference points by using the grid corner points, which successfully addresses the above issue.
- Grid-IGD can better reflect the comprehensive performance of the approximations delivered by multi/many-objective optimizers, compared with IGD.

V. Conclusion

In this paper, a Grid-IGD is proposed for improving the performance evaluations of the original IGD on PF approximations delivered by multi/many-objective optimizers. Different from the original IGD, Grid-IGD does not require the information about the true PFs *a priori*; it can more accurately reflect the comprehensive quality of PF approximations and it has lower computational complexity, which is very desirable for evaluating high dimensional PF approximations. Experimental results show the effectiveness of Grid-IGD for measuring the comprehensive quality for both artificial approximations and

the ones obtained by five MOEAs. The proof of Grid-IGD as a weakly Pareto compliant indicator is also given. Further work includes designing multi/many-objective optimization algorithms based on Grid-IGD.

ACKNOWLEDGEMENT

We would like to thank Dr. Mustafa Misir for proofreading the manuscript.

REFERENCES

[1] K. Miettinen, *Nonlinear multiobjective optimization*. Springer Science & Business Media, 2012, vol. 12.

[2] X. Cai, Y. Li, Z. Fan, and Q. Zhang, "An external archive guided multiobjective evolutionary algorithm based on decomposition for combinatorial optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 4, pp. 508–523, 2015.

[3] X. Cai, Z. Mei, Z. Fan, and Q. Zhang, "A constrained decomposition approach with grids for evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 4, pp. 564–577, 2018.

[4] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.

[5] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on evolutionary computation*, vol. 11, no. 6, pp. 712–731, 2007.

[6] K. Deb and K. Miettinen, *Multiobjective optimization: Interactive and evolutionary approaches*. Springer Science & Business Media, 2008, vol. 5252.

[7] C. A. C. Coello, G. B. Lamont, and D. A. Van Veldhuizen, *Evolutionary algorithms for solving multi-objective problems*. Springer Science & Business Media, 2007.

[8] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. D. Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, April 2003.

[9] M. Li and X. Y, "Quality evaluation of solution sets in multiobjective optimisation: A survey," *ACM Comput. Surv.*, vol. 52, no. 2, pp. 26:1–26:38, 2019.

[10] X. Cai, H. Sun, and Z. Fan, "A diversity indicator based on reference vectors for many-objective optimization," *Information Sciences*, vol. 430, pp. 467–486, 2018.

[11] M. Li, S. Yang, and X. Liu, "Diversity comparison of Pareto front approximations in many-objective optimization," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2568–2584, 2014.

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

[13] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.

[14] J. Wu and S. Azarm, "Metrics for quality assessment of a multiobjective design optimization solution set," *Journal of Mechanical Design*, vol. 123, no. 1, pp. 18–25, 2001.

[15] P. A. Bosman and D. Thierens, "The balance between proximity and diversity in multiobjective evolutionary algorithms," *IEEE transactions on evolutionary computation*, vol. 7, no. 2, pp. 174–188, 2003.

[16] C. A. C. Coello and M. R. Sierra, "A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm," in *MICAI 2004: Advances in Artificial Intelligence, Third Mexican International Conference on Artificial Intelligence, Mexico City, Mexico, April 26-30, 2004, Proceedings*, 2004, pp. 688–697.

[17] J. Bader and E. Zitzler, "HypE: An algorithm for fast Hypervolume-based many-objective optimization," *Evolutionary computation*, vol. 19, no. 1, pp. 45–76, 2011.

[18] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, "How to specify a reference point in Hypervolume calculation for fair performance comparison," *Evolutionary Computation*, vol. 26, no. 3, 2018.

[19] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima, "Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 2, pp. 169–190, 2017.

[20] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, "Reference point specification in inverted generational distance for triangular linear Pareto front," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 6, pp. 961–975, 2018.

[21] M. Asafuddoula, T. Ray, and R. Sarker, "A decomposition-based evolutionary algorithm for many objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 3, pp. 445–460, 2015.

[22] E. Zitzler, D. Brockhoff, and L. Thiele, "The Hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2007, pp. 862–876.

[23] L. S. Russo and A. P. Francisco, "Quick Hypervolume," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 481–502, 2014.

[24] T. Watanabe, T. Tatsukawa, and A. Oyama, "On the fast Hypervolume calculation method," in *Evolutionary Computation*, 2010, pp. 965–969.

[25] L. While, L. Bradstreet, and L. Barone, "A fast way of calculating exact Hypervolumes," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pp. 86–95, 2012.

[26] M. Li, S. Yang, and X. Liu, "A performance comparison indicator for Pareto front approximations in many-objective optimization," in *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2015, Madrid, Spain, July 11-15, 2015*, 2015, pp. 703–710.

[27] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, "Modified distance calculation in generational distance and inverted generational distance," in *Evolutionary Multi-Criterion Optimization - 8th International Conference, EMO 2015, Guimarães, Portugal, March 29 -April 1, 2015. Proceedings, Part II*, 2015, pp. 110–125.

[28] O. Schütze, X. Esquivel, A. Lara, and C. A. C. Coello, "Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 4, pp. 504–522, 2012.

[29] G. Rudolph, H. Trautmann, S. Sengupta, and O. Schütze, "Evenly spaced Pareto front approximations for tricriteria problems based on triangulation," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2013, pp. 443–458.

[30] A. Menchaca-Mendez and C. A. C. Coello, "GD-MOEA: A new multi-objective evolutionary algorithm based on the generational distance indicator," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2015, pp. 156–170.

[31] Z. He and G. G. Yen, "Visualization and performance metric in many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 3, pp. 386–402, 2016.

[32] M. P. Hansen and A. Jaszkiewicz, *Evaluating the quality of approximations to the non-dominated set*. IMM, Department of Mathematical Modelling, Technical Universityof Denmark, 1994.

[33] D. Brockhoff, T. Wagner, and H. Trautmann, "On the properties of the R2 indicator," in *Proceedings of the 14th annual conference on Genetic and evolutionary computation*. ACM, 2012, pp. 465–472.

[34] D. A. V. Veldhuizen and G. B. Lamont, "Multiobjective evolutionary algorithm test suites," in *ACM Symposium on Applied Computing*, vol. 99, 1999, pp. 351–357.

[35] S. Jiang and S. Yang, "A steady-state and generational evolutionary algorithm for dynamic multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 1, pp. 65–82, 2017.

[36] K. Deb, M. Mohan, and S. Mishra, "Evaluating the ϵ-domination based multi-objective evolutionary algorithm for a quick computation of pareto-optimal solutions," *Evolutionary computation*, vol. 13, no. 4, pp. 501–525, 2005.

[37] S. Yang, M. Li, X. Liu, and J. Zheng, "A grid-based evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 5, pp. 721–736, 2013.

[38] F. Fortin, S. Grenier, and M. Parizeau, "Generalizing the improved run-time complexity algorithm for non-dominated sorting," in *Proceedings of the 15th annual conference on Genetic and evolutionary computation*. ACM, 2013, pp. 615–622.

[39] S. S. Chern and P. Griffiths, "Abel's theorem and webs," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 80, no. 1-2, pp. 13–110, 1978.

[40] H. Li and Q. Zhang, "Multiobjective Optimization Problems With Complicated Pareto Sets, MOEA/D and NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 284–302, 2009.

[41] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2004, pp. 832–842.

[42] J. Knowles and D. Corne, "The Pareto archived evolution strategy: A new baseline algorithm for Pareto multiobjective optimisation," in

Congress on Evolutionary Computation (CEC99), vol. 1, 1999, pp. 98–105.

[43] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.

[44] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.

[45] K. Li, K. Deb, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition." *IEEE Trans. Evolutionary Computation*, vol. 19, no. 5, pp. 694–716, 2015.

[46] J. J. Durillo and A. J. Nebro, "jmetal: A java framework for multi-objective optimization," *Advances in Engineering Software*, vol. 42, no. 10, pp. 760–771, 2011.

[47] Y. Tian, X. Zhang, R. Cheng, and Y. Jin, "A multi-objective evolutionary algorithm based on an enhanced inverted generational distance metric," in *IEEE Congress on Evolutionary Computation, CEC 2016, Vancouver, BC, Canada, July 24-29, 2016*. IEEE, 2016, pp. 5222–5229.

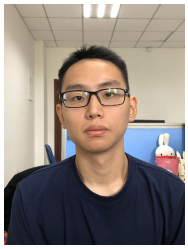
[48] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary multiobjective optimization*. Springer, 2005, pp. 105–145.

[49] Q. Zhang, A. Zhou, S. Zhao, S. P., W. Liu, and S. Tiwari, "Multiobjective optimization test instances for the CEC 2009 special session and competition," University of Essex, Colchester, UK and Nanyang technological University, Singapore, Tech. Rep., 01 2008.

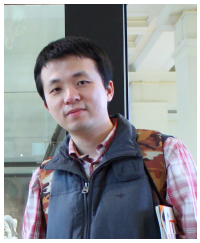
[50] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 5, pp. 477–506, 2006.

Xinye Cai (M'10) received the BSc in information engineering from Huazhong University of Science and Technology, China in 2004, the MSc in electronic engineering from University of York, UK, in 2006 and the PhD in electrical engineering from Kansas State University, U.S, in 2009, respectively.

He is an Associate Professor at the Department of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics (NUAA), China. His current research interests include optimization, machine learning and their applications. He is currently leading an Intelligent Optimization Research Group in NUAA, who has won the Evolutionary Many-Objective Optimization Competition at the Congress of Evolutionary Computation 2017. Dr. Cai is an Associate Editor of Swarm and Evolutionary Computation.



Yushun Xiao is currently pursuing the Master degree in Software Engineering at the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China. His research interests include multiobjective optimization, data mining and kernel method.



Miqing Li received the Ph.D. degree in computer science from the Department of Computer Science, Brunel University London, U.K. in 2015.

Dr Miqing Li is currently a lecturer with the University of Birmingham, U.K. His research is principally on multi-objective optimisation, where he works on developing population-based randomised algorithms (e.g. evolutionary algorithms) for both general challenging problems (e.g. many-objective optimisation, constrained optimisation, multi-modal optimisation, robust optimisation, expensive optimisation) and specific challenging problems in other fields (e.g. software engineering, system engineering, product disassembly, post-disaster response, neural architecture search, and reinforcement learning). He is the founding chair of IEEE CIS Task Force on Many-Objective Optimisation.

Han Hu received his Master degree in Computer Science at the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China, in 2019. Currently, he is working as an engineer in VIVO. His research interests include multiobjective optimization.

Hisao Ishibuchi (M'93?SM'10?F'14) received the B.S. and M.S. degrees in precision mechanics from Kyoto University, Kyoto, Japan, in 1985 and 1987, respectively, and the Ph.D. degree in computer science from Osaka Prefecture University, Sakai, Osaka, Japan, in 1992. He was with Osaka Prefecture University in 1987-2017. Since 2017, he is a Chair Professor at Southern University of Science and Technology, China. His research interests include fuzzy rule-based classifiers, evolutionary multi-objective and many-objective optimization, memetic algorithms, and evolutionary games.

Dr. Ishibuchi was the IEEE CIS Vice-President for Technical Activities in 2010-2013, an AdCom member of the IEEE CIS in 2014-2019, and the Editor-in-Chief of the IEEE Computational Intelligence Magazine in 2014-2019.

Xiaoping Li (M'09-SM'12) received his B.Sc. and M.Sc. degrees in Applied Computer Science from the Harbin University of Science and Technology, Harbin, China, in 1993 and 1999, respectively, and the Ph.D. degree in Applied Computer Science from the Harbin Institute of Technology, Harbin, China, in 2002. He is currently a distinguished professor at the School of Computer Science and Engineering, Southeast University, Nanjing, China. He is the author or co-author over more than 100 academic papers, some of which have been published in international journals such as *IEEE Transactions on Computers*; *IEEE Transactions on Parallel and Distributed Systems*; *IEEE Transactions on Services Computing*; *IEEE Transactions on Cybernetics*; *IEEE Transactions on Automation Science and Engineering*; *IEEE Transactions on Cloud Computing*; *IEEE Transactions on Systems, Man and Cybernetics: Systems*; *Information Sciences*; *Omega*, *European Journal of Operational Research*. His research interests include Scheduling in Cloud Computing, Scheduling in Cloud Manufacturing, Service Computing, Big Data and Machine Learning.