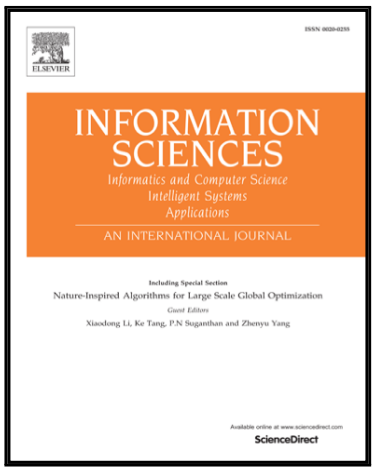
An angle dominance criterion for evolutionary many-objective optimization

Liu, Yuan; Zhu, Ningbo; Li, Kenli; Li, Miqing; Zheng, Jinhua; Li, Keqin

DOI:
10.1016/j.ins.2018.12.078

Document Version
Peer reviewed version

Citation for published version (Harvard):
Liu, Y, Zhu, N, Li, K, Li, M, Zheng, J & Li, K 2020, 'An angle dominance criterion for evolutionary many-objective optimization', *Information Sciences*, vol. 509, pp. 376-399. https://doi.org/10.1016/j.ins.2018.12.078



An Angle Dominance Criterion for Evolutionary Many-Objective Optimization

Yuan Liu, Ningbo Zhu, Kenli Li, Miqing Li, Jinhua Zheng, Keqin Li

PII: S0020-0255(18)31039-9
DOI: https://doi.org/10.1016/j.ins.2018.12.078
Reference: INS 14178

To appear in: *Information Sciences*

Received date: 20 December 2017
Revised date: 19 November 2018
Accepted date: 29 December 2018

Please cite this article as: Yuan Liu, Ningbo Zhu, Kenli Li, Miqing Li, Jinhua Zheng, Keqin Li, An Angle Dominance Criterion for Evolutionary Many-Objective Optimization, *Information Sciences* (2018), doi: https://doi.org/10.1016/j.ins.2018.12.078

An Angle Dominance Criterion for Evolutionary Many-Objective Optimization

Yuan Liu[a,b], Ningbo Zhu[a,b,*], Kenli Li[a,b], Miqing Li[c], Jinhua Zheng[d], Keqin Li[a,e]

[a] *College of Computer Science and Electronic Engineering, National Supercomputing Center in Changsha, Hunan University*
[b] *College of Computer Science and Electronic Engineering, Hunan University, Hunan, China*
[c] *CERCIA, School of Computer Science, University of Birmingham, Birmingham B15 2TT, U. K.*
[d] *School of Information Engineering, Xiangtan University, Hunan, China*
[e] *Department of Computer Science, State University of New York, New Paltz, NY 12561, USA*

Abstract

It is known that Pareto dominance encounters difficulties in many-objective optimization. This strict criterion could make most individuals of a population incomparable in a high-dimensional space. A straightforward approach to tackle this issue is modify the Pareto dominance criterion. This is typically done by relaxing the dominance region. However, this modification is often associated with one or more parameters of determining the relaxation degree, and the performance of the corresponding algorithm could be sensitive to such parameters. In this paper, we propose a new dominance criterion, angle dominance, to deal with many-objective optimization problems. This angle dominance criterion can provide sufficient selection pressure towards the Pareto front and be exempt from the parameter tuning. In addition, an interesting property of the proposed dominance criterion, in contrast to existing dominance criteria, lies in its capability to reflect an individual's extensity in the population. The angle dominance is integrated into NSGA-II

*Corresponding author
Email addresses: liu3yuan@hnu.edu.cn (Yuan Liu), quietwave@hnu.edu.cn (Ningbo Zhu), lkl@hnu.edu.cn (Kenli Li), limising@gmail.com (Miqing Li), jhzheng@xtu.edu.cn (Jinhua Zheng), lik@newpaltz.edu (Keqin Li)

(instead of Pareto dominance) and has demonstrated high competitiveness in many-objective optimization in comparison with a range of peer algorithms.

Keywords: Angle dominance criterion, Pareto dominance criterion, many-objective optimization, evolutionary algorithms.

1. Introduction

Many real-world optimization problems are composed of multiple conflicting objectives which need to be optimized simultaneously. For a multi-objective optimization problem (MOP), an improvement of the performance on one objective often leads to the deterioration on other objective(s). Therefore, Multi-objective optimization algorithms (MOEAs) can only search for a set of trade-off solutions to approximate Pareto optimal solutions.

In three-decades, MOEAs have attracted great attention for being able to solve a class of real-world optimization problems that have multiple criteria or objectives [30, 7, 43]. However, traditional Pareto-based MOEAs can only effectively solve two- or three-dimensional optimization problems. In real world, the number of considered objectives can be larger (i.e., over three), and these problems are known as many-objective optimization problems (MaOPs). When facing MaOPs, it's not easy for traditional Pareto-based MOEAs to converge into the Pareto front. The main reason is that the proportion of non-dominated solution increases rapidly with the number of objectives. Consequently, the density based second selection criterion in Pareto-based algorithms plays a leading role in the selection process of Pareto-based MOEAs [33]. However, the studies in [24] indicate that a diversity-based selection criterion has a detrimental impact on the population's convergence. This criterion prefers the dominance resistant solutions (DRSs) which have "good" diversity over the objective space but are far away from the desired Pareto front [33]. Therefore, balancing the convergence and diversity of the population for MaOPs has become a challenging research topic in the field of many-objective optimization.

To solve the problems above, many methods have been proposed, and can be divided into the following categories:

Loosening Pareto-dominance approach. With the increase of the number of objectives, the Pareto-dominance relationship is difficult to distinguish between solutions in terms of convergence [19]. By loosening the Pareto dominance relation (i.e., increasing solutions' dominance area). That the

solutions approach the Pareto front more rapidly will be identified. For example, the controlling dominance area (CDAS) approach [38] adjusts the dominance area of solutions by setting an appropriate parameter. The grid dominance [44] adds the selection pressure by adopting an adaptive grid construction. The ε-dominance based MOEA (ε-MOEA) [9] can obtain good convergence and uniformity performance by dividing the objective space into hyper-boxes, each of which is assigned at most one solution.

Decomposition-based approach. Using a set of uniformly distributed weight vectors, these approaches decompose a MOP into a number of single-objective sub-problems and then uses a search heuristic to optimise these sub-problems simultaneously and cooperatively. In contrast to the Pareto dominance criterion, decomposition-based approaches can rank the entire population and form a total order among the solutions and thus providing sufficient selection pressure in a high-dimensional objective space. Recently, some decomposition-based MOEAs, such as the MOEA with decomposition (MOEA/D) [45], multiple single-objective Pareto sampling (MSOPS) [22] and MOEA/D based on localized weighted sum (MOEA/D-LWS) [40], have been found to work well on MaOPs.

Ranking-based approach. Ranking-based algorithms can distinguish between solutions by defining a new sorting method. Similar to decomposition-based approaches, ranking-based approaches can form a total order among the solutions. In [13], the relation favour is to rank the solutions by comparing the numbers of their superior objectives. Therefore, the relation favour will prefer those solutions if most of their objectives are superior to that of others. Additionally, the average ranking (AR) [5] is another ranking method. Firstly, on each objective, it ranks all the non-dominated solutions on the basis of their objective values, and the number of ranking values of a solution equals to the number of the objectives of the problem. Then, AR sorts the solutions by means of their average ranking values.

Density estimation based approach. Recently, some researches have shown that some modification of diversity maintenance in Pareto-based algorithms can also promote the convergence of population, e.g., the diversity management operator (DMO) [1] and shift-based density estimation (SDE) [33]. DMO uses one indicator (i.e., the maximum extensity indicator) to evaluate the population's convergence. This approach relies on the true Pareto front of the problem. However, in practical applications, the Pareto front of most problems is unknown. SDE, as a diversity estimation, takes two aspects into account: convergence and diversity. In contrast to DMO, SDE has a high

usability, and does not need to know the Pareto front of the problem. Experimental results show that it can significantly improve the performance of Pareto-based algorithms.

To evaluate the performances of the above MOEAs, the diagnostic assessment framework [17, 37] is another research hotspot in the evolutionary computation community. It contains three important elements. The first one is multiple performance metrics. They mainly evaluate the effectiveness, reliability, efficiency, and controllability of an MOEA [37]. Effectiveness checks whether an MOEA achieves high level of performance. Reliability captures performance changes in the parametric process and random seed testing. Efficiency refers to achieving high levels of performance in the minimum number of function evaluations. Controllability measures the ease-of-use or sensitivity of MOEAs' to their parameterizations. Then, an adequate sample of problems is another element. A number of test problem have been developed to benchmark the performance of MOEAs, such as Deb-Thiele-Laumanns-Zitzler (DTLZ) [11], Walking Fish Group (WFG) [21] and multiline distance minimization problem (ML-DMP) [32]. Among them, the most widely-used element are DTLZ and WFG. And the last one is the ability to uncover pertinent parameter controls and dynamic search behavior within the algorithm.

Although the above studies clearly enhance the search ability of MOEAs and various methods were proposed to tackle MaOPs, the area of evolution many-objective optimization is far from being mature. Furthermore, the loosening Pareto-dominance approaches inevitably encounter difficulties in determining the degree of slack in the new dominance relation for difference problems, leading to the emergence of dynamic tuning methods. For the decomposition-based approaches, two critical issues need to be considered. One of which is that the specified weights' distribution needs to be consistent with a given problem's Pareto front. The other is that the configuration of weight vectors suffers the curse of dimensions in many-objective space. Due to the lack of a diversity maintenance strategy, the ranking-based approaches may lead the evolutionary population to converge into a small part of the Pareto front. As mentioned above, the true Pareto front of the problem affects the performance of the density estimation based approach. Although the SDE is off the hook, parts of the solution near the boundary are easily eliminated by it.

In this paper, we focus on the first approach and wish to propose a dominance relationship (named as angle dominance) that is insensitive to parameters. One interesting property of the angle dominance is its capability of

reflecting the convergence and extensity of solutions in the population. This is in contrast to existing dominance criteria, which typically only involve convergence (e.g., Pareto dominance) or both convergence and uniformity (e.g., ε-dominance). The basic idea of the angle dominance is simple. By substituting the objective vector of a solution with an angle vector, the angle dominance enlarges the dominance area of the solution. This not only increases the selection pressure towards the Pareto front but also is able to maintain boundary solutions very well. In addition, the angle dominance can be easily applied to any Pareto dominance based algorithms.

The rest of this paper is organized as follows. Section 2 briefly reviews the work related to dominance relationships and angle-based environmental selection. Section 3 is devoted to description of the proposed angle dominance criterion and introduces the framework of the angle dominance based NSGA-II which is denoted as NSAG-II+AD. Section 4 presents the algorithm settings, test functions, and performance metrics used for performance comparison. The experimental results and relevant analyses are presented in section 5. Finally, section 6 concludes the paper and gives our study priorities in the future work.

2. Related Works

2.1. Related Dominance Criteria

It has been demonstrated that traditional Pareto dominance generally fail to solve MaOPs. Therefore, in past decades, a number of loosening dominance criteria have been proposed, such as α-Dominance [23], ε-Dominance [9], CDAS [38], Cone ε-Dominance [4] and Grid-Dominance [44], etc. In this section, we will analyze the methods mentioned above in detail. For simplicity, we assume that the optimization problems mentioned are minimization ones throughout the paper.

A. Pareto Dominance

In 1896, Pareto proposed the concept of Pareto dominance, as shown in Fig. 1(a). Suppose that there are two solutions p and q, p dominates q (denoted as $p \prec q$) if the following conditions hold:

$$\begin{cases} f_i(p) \leq f_i(q), \ \forall i \in \{1, 2, \cdots, m\} \\ f_j(p) < f_j(q), \exists j \in \{1, 2, \cdots, m\} \end{cases}, \tag{1}$$

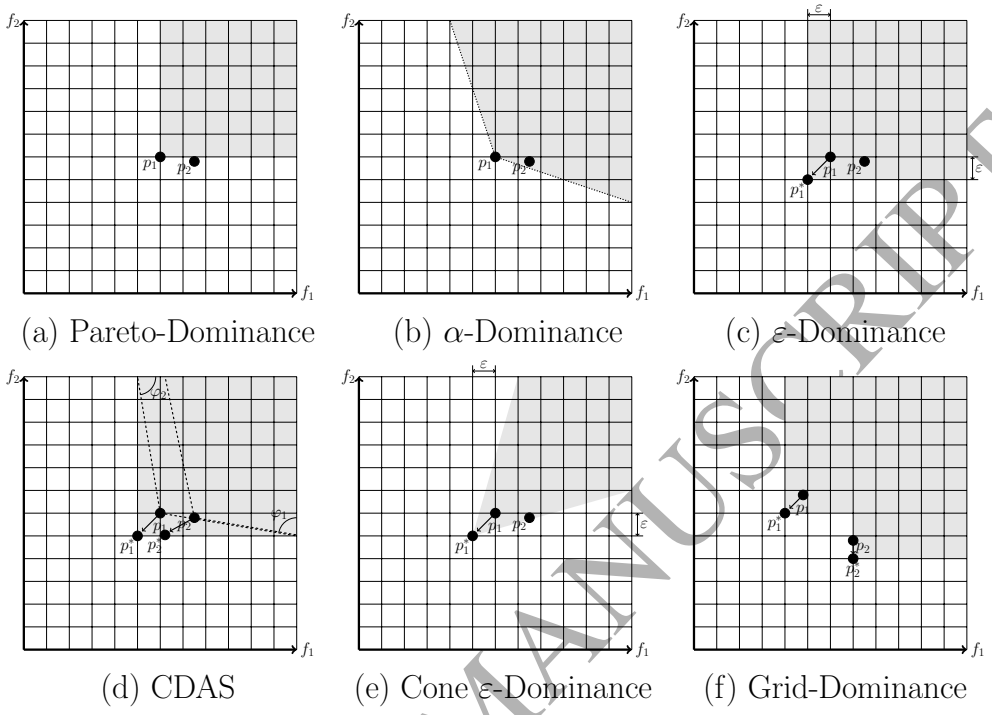


Figure 1: Illustration of six dominance relationships among points in a two-dimensional objective space: (a) Pareto dominance, where the grey area is the objective space where solutions are Pareto-dominated by p_1; (b) α-dominance, where the α-dominance area of p_1 is clearly larger than that in the Pareto dominance; (c) ε-dominance, where p_1^* is the shift point of p_1 by means of ε_1 and ε_2; (d) CDAS, where the shape of the grey area is quite similar to that in ε-dominance. The angles φ_1 and φ_2 are used to control the dominance area; (e) Cone ε-dominance, where the grey area is cone ε-dominated by p_1, and the shape of the dominance area is a cone; (f) Grid-dominance, where the grid is divided by the current population $\{p_1, p_2\}$. The shape of the grey area is similar to that in ε-dominance.

where m is the number of objectives. In other words, all objectives in p are not greater than the corresponding objectives of q, and at least one objective of p is less than that of q.

The Pareto dominance relationship can divide the original population into multiple sub-populations, then the sub-populations with high priorities will be preserved in the environmental selection. However, Pareto dominance is generally effective to handle two- and three-objective MOPs. When a MOP has more than three objectives, Pareto dominance will lose its effectiveness in most cases. This is because with the increase of the dimensionality, the

proportion of non-dominance individuals will grow exponentially. As can be seen in Fig. 2, when the number of objectives reaches 20, all the solutions are non-dominated and belong to the same sub-population. That is to say, Pareto dominance can not maintain the population's diversity. To solve that, the distribution mechanisms, such as the crowding distance [10] and archive truncation procedure [49], need be applied in the environmental selection.

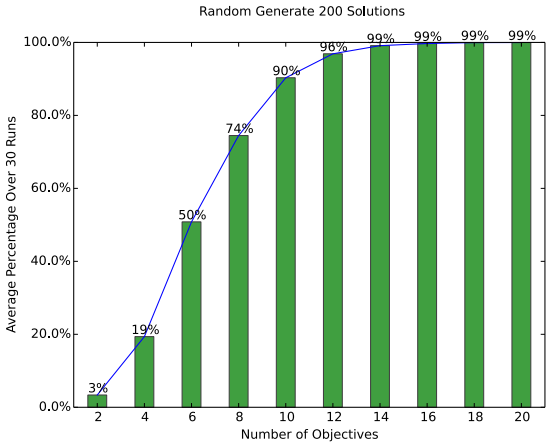


Figure 2: Percentage of the number of non-dominated solutions over that of objectives in a set of randomly generated 200 solutions. Where the abscissa represents the objective dimension and the ordinate represents the average percentage over 30 runs.

B. α-Dominance

In multi-objective optimization, the solutions, which are far from the Pareto front but are hardly dominated, are defined as dominance resistant solutions (DRSs) [23]. These solutions may have a detrimental effect on the convergence of the population. In order to solve this issue, Ikeda *et al.* [23] proposed a relaxed form of the dominance relation, called α-dominance, as shown in Fig. 1(b).

In α-dominance, the upper $\alpha_{i,j}$ and lower $\frac{1}{\alpha_{i,j}}$ bounds of trade-off rates between two objectives f_i and f_j are pre-defined, and the trade-off rates between two objectives f_i and f_j of any two solutions are strictly controlled within the pre-defined bounds. Before judging the dominance relations between two individuals, the following definition is considered:

$$g_i(p,q) := f_i(p) - f_i(q) + \sum_{j \neq i}^{K} \alpha_{ij} \left(f_j(p) - f_j(q) \right), \ p, q \in \mathbf{P}, \tag{2}$$

where p and q are two solutions in population $\mathbf{P}$. The p dominates the q ($p \prec_\alpha q$) if and only if the following conditions hold:

$$\begin{cases} g_i(p,q) \leq 0, \forall i \in \{1,2,\cdots,m\} \\ g_j(p,q) < 0, \exists j \in \{1,2,\cdots,m\} \end{cases}, \tag{3}$$

where m is the number of objectives.

In Fig. 1(b), two Pareto non-dominated solutions p_1 and p_2 are regarded as p_1 α-dominating p_2. According to the definition of α-dominance criterion, when α is larger, the population can converge more easily to the Pareto front, but it is more prone to be trapped in a local optimum. On the contrary, when α is smaller, a wider Pareto front will be found, but more solutions far from the Pareto front will be preserved. Therefore, it is hard to find a well-distributed and well-converged trade-off solution set for α-dominance.

C. (1-k)-based Dominance

The (1-k)-based criterion [14] has been considered when addressing MaOPs. By comparing a solution to another and counting the number of objectives where it is better than, the same as, or worse than the other, this criterion uses these numbers to distinguish the relations of domination between individuals. Suppose that there are two solutions p and q, and n_b, n_e, n_w respectively represent the number of objectives where p is better than, equal to, or worse than q. p is said to k-dominate q if and only if:

$$n_e < m \quad \wedge \quad n_b \geq \frac{m - n_e}{k+1}, \tag{4}$$

where $0 \leq k \leq 1$, and m is the number of objectives.

Obviously, if $k = 0$, the $(1-k)$-dominance will be consistent with the traditional Pareto-dominance, and if k is larger, the $(1-k)$-dominance is looser. In addition, there are some extensions on $(1-k)$-dominance. For example, fuzzy numbers can be applied to compare the dominance relation

between solutions. In fuzzy dominance, n_b, n_e, n_w are defined as follows:

$$\begin{cases} n_b^F(p,q) = \sum_{i=1}^{M} u_b^{(i)}(f_i(p_i) - f_i(q_2)) \\ n_e^F(p,q) = \sum_{i=1}^{M} u_e^{(i)}(f_i(p_i) - f_i(q_2)) \\ n_w^F(p,q) = \sum_{i=1}^{M} u_w^{(i)}(f_i(p_i) - f_i(q_2)) \\ n_b^F + n_w^F + n_e^F = \sum_{i=1}^{M}(u_b^{(i)} + u_e^{(i)} + u_w^{(i)}) = m \end{cases}. \tag{5}$$

In terms of the above expressions, the $(1 - k_F)$-dominance considers the convergence of the population according to the above three situations. As a matter of fact, the $(1 - k_F)$-dominance can be seen as a dimensionality reduction strategy, so it is unavoidable to lose some target information.

D. ε-Dominance

Deb *et al.* proposed a steady-state MOEA, named as ε-MOEA. The ε-dominance is illustrated in Fig. 1(c). In the figure, within each box, only one non-Pareto dominated solution is preserved, with the following condition:

$$(1 - \varepsilon) \cdot f_i(p) \leq f_i(q), \ \forall i \in \{1, 2, \cdots, m\}, \tag{6}$$

where p and q are two non-Pareto dominance solutions and m is the number of objectives.

The ε-dominance criterion could maintain the uniformity of population. However, the extreme solutions are also easily ε-dominated by other solutions, which can affect the extensity of the population. In addition, different problems generally require distinct ε values.

E. Controlling Dominance Area of Solution (CDAS)

In 2007, Sato *et al.* proposed a new dominance, CDAS, which is defined as follows:

$$f_i'(x) = \frac{r \cdot sin(\omega_i + S_i \cdot \pi)}{sin(S_i \cdot \pi)}, \tag{7}$$

where r is the norm of $f(x)$ and ω_i is the declination angle between $f(x)$ and the coordinate axis. Fig. 1(d) shows the meaning of CDAS.

From Eq. (7), the dominance area of solution x can be controlled by the parameter S_i. It is obvious that if $S_i = 0.5$, then $f_i'(x) = f_i(x)$ which corresponds to the classical definition of Pareto dominance; if $S_i < 0.5$, then

$f_i'(x) > f_i(x)$; and if $S_i > 0.5$, then $f_i'(x) < f_i(x)$. Such objective modification changes the dominance area of solutions. Therefore, a suitable parameter S_i can effectively promote the convergence of the population. However, as the same as ε-dominance, the parameter of CDAS is also not easily determinable when facing different problems.

F. Volume Dominance

The volume dominance criterion was proposed by Khoi Le *et al.*. It uses the strength of solutions to analyze the dominance relations between solutions. For two solutions p and q, the dominated volume of each solution ($V(p)$ and $V(q)$) needs to be calculated as follows:

$$V(p) = \prod_{i=1}^{m} \left(f_i(p) - r_i\right), \tag{8}$$

where r is a reference point. And the shared dominated volume of p and q is defined as:

$$SV(p,q) = \prod_{i=1}^{m} \left(min(f_i(p), f_i(q)) - r_i\right). \tag{9}$$

It is said that p volume-dominates q ($p \prec_V q$) if either:

$$\begin{cases} V(q) = SV(q,p) \text{ and } V(p) < SV(p,q) \text{ or} \\ V(p) < V(q) < SV(p,q) \text{ and } \frac{V(p)-V(q)}{SV(p,q)<rSV} \end{cases}. \tag{10}$$

As described by Eq. (10), when $V(p)$ is smaller, the solution p is closer to the Pareto front, and vice versa. Therefore, the volume-dominance criterion can make the population quickly converge to the Pareto front. But the disadvantage of volume-dominance is that its performance relies heavily on the shape of the Pareto front.

G. Cone ε-Dominance

When dealing with MOPs, it has been found that ε-dominance may eliminate several viable solutions (See appendix A for explanation), which affects the convergence and extensity of population. Batista *et al.* proposed the cone ε-dominance. In Fig. 1(e), p_1 ε-dominates p_2 and both of them are non-cone ε-dominated. Cone ε-dominance introduces a parameter k ($k \in [0, 1)$), and k is applied to control the shape of dominance area of a solution. When $k \to 0$, the cone ε-dominance is consistent with the traditional Pareto-dominance.

When $k > 0$, the shape of the dominance area is a cone. As for two solutions p and q, p cone ε-dominates q (denoted as $p \prec_{cone\varepsilon} q$) if the following condition holds:

$$(p \prec q) \vee (\Psi\lambda = z|\lambda_i \geq 0, \forall i \in \{1, \cdots, m\}), \tag{11}$$

where Ψ is the cone ε-dominance matrix, $z = q - [p - \varepsilon]$, and $\varepsilon_i > 0$.

According to the above description, the cone ε-dominance could improve both convergence and uniformity of population (See appendix B for explanation). However, to improve the performance of an algorithm, except for the parameter ε, the cone ε-dominance has to add another parameter k into the algorithm. This parameter, together with ε, can limit the application of the algorithm.

H. Grid Dominance

Yang *et al.* modified the traditional ε-dominance and proposed a grid dominance criterion in the grid-based evolutionary algorithm (GrEA), as shown in Fig. 1(f). Inspired by the ideas in [27], GrEA adaptively constructs grids. Comparing with traditional grid-based approaches, GrEA adopts individually centered calculations of the grid by depicting the locations of solutions. This could determine the mutual relationship of solutions in a grid environment so as to increase the diversity. Nevertheless, the individually centered calculation of the grid suffers from potential deterioration of convergence since the adjacent well-converged solutions are eliminated.

Table 1: The properties of nine domination relations.

Dominance Criterion	Convergence	Uniformity	Extensity	Irreflexive	Asymmetric	Transitive	Strict Partial Order
Pareto-Dominance	√			√	√	√	√
α-Dominance	√			√	√	√	√
$(1-k)$-Dominance	√			√	√		
ε-Dominance	√	√				√	
CDAS	√			√	√	√	√
Volume-Dominance	√			√	√	√	√
Cone ε-Dominance	√	√				√	
Grid-Dominance	√	√		√	√	√	√
Angle Dominance	√		√	√	√	√	√

"√" in the cell indicates that the domination relation has a corresponding property.

Table 1 summaries the properties of all the dominance criteria. In this table, the diversity of solution set is subdivided into the uniformity and extensity of the solution set. In general, uniformity quantifies the distance between neighboring points in the solution set, and extensity refers to the range of the solution set. It worth mentioning that a uniformly-distributed

solution set does not necessarily mean that the solution set spread very well. As a complement to uniformity, extensity considers the spread of the solution set. This table shows that all the existing dominance criteria focus on the convergence of population, a few of which involve the uniformity of population, and no one of them involve the extensity of population. In this paper, we propose an angle dominance criterion that takes both the convergence and extensity into account. In this strategy, the angle vector is applied to replace the objective vector of a solution to reflect its position in the objective space. This will lead to two characteristics of the proposed criterion. The first one is that, it retains the basic information (like its objective position) of a solution. For each solution, the position of the angle vector corresponds to the position of the objective vector. And the other one is that, the angle dominance strategy flexibly enlarges the dominance area of a solution. Consequently, it can increase the selection pressure in terms of convergence so as to make the population move towards the Pareto front. As for the extensity, the angle dominance uses the angle vectors to determine the mutual relationships between solutions. When a solution is closer to the Pareto front, its angle becomes smaller, which makes it a higher fitness. In addition, angle dominance and Pareto dominance share some common properties, such as the irreflexive relation, asymmetric relation, transitive relation and strict partial order.

2.2. The Angle-based Environmental Selection

In recent years, the angle-based environmental selection is widely considered in EMO. For instance, in [48], MOEA/D-ADCP uses angle to determine the dominance relationship between two solutions. If the angle of the two solutions is greater than a given threshold, they are considered to be non-dominated by each other. In the decomposition-based approaches [26], the angle of solution and weight is used to judge their similarity in the search directions. In other words, the larger their angle, the smaller their similarity. The angle-based selection is also used to improve the diversity of Pareto-based approaches. In [42], VaEA first uses the maximum-vector-angle-first principle to guarantee the extensity and uniformity of the solutions, and then ensure the convergence of the solutions through the worse-elimination principle. In addition, in MOEA/VAN, the angle-based selection is exploited during mating and environmental selection by determining the neighborhood and the most crowded region in the objective space, respectively [12].

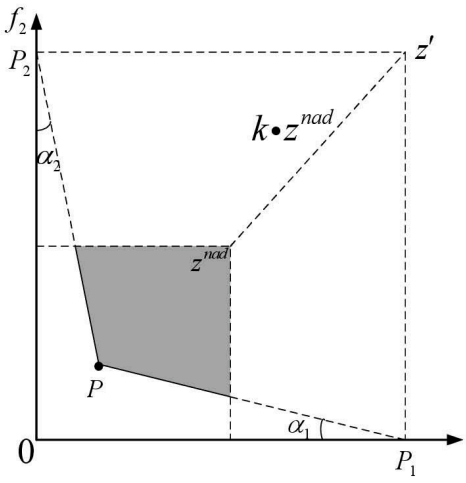


Figure 3: The angle dominance in a two-objective scenario. p is a solution, z^{nad} is the worst point, and the shaded region is the dominance area of p. α_1 and α_2 are the components of the angle vector of p. z' is constructed by $k \cdot z^{nad}$, where k is a preset parameter.

Although this paper also uses the angle-based selection criterion to select excellent individuals, this criterion is significantly different from the above methods. Firstly, the angle of this paper is composed of the solution and each axis node, and most of the angle-based selection criteria are to calculate the angle between each pair of solutions. Secondly, the purpose of our angle-based selection is to increase the convergence pressure of the solutions, while most of the other criteria based on angle selection are to improve the diversity of the solutions.

3. Proposed Angle Dominance Criterion

In this section, we first introduce the concept of angle dominance and analyze its properties, and then integrate the angle dominance criterion into NSGA-II.

3.1. Concept of Angle Dominance

In order to interpret the concept of the proposed angle dominance criterion, Fig. 3 presents two-objective scenario of the criterion. Solution p is assigned an identification angle vector $angle_p = (\alpha_1, \alpha_2)$, and the shaded region is the dominance area of p. More specifically, in Fig. 3, z^{nad} is the

worst point of the current population, defined as $z^{nad} = (z_1^{nad}, z_2^{nad})$, where $z_i^{nad} = \overset{n}{\underset{j=1}{max}} f_i(x_j)$ (n is the size of the population). The point z' is constructed by means of a preset parameter k and the worse point z^{nad}, and it is defined as $z' = (k \cdot z_1^{nad}, k \cdot z_2^{nad})$, where k is designed to control the dominance area of a solution. The smaller k is, the larger the dominance area of the solution is. For instance, if k is less than 1 and very close to 0, it means the z' is near to the origin of coordinate and dominates all the solutions. It will result in the solutions dominate each other. Inversely, with the k increasing, the dominance area of the solutions is gradually decrease. The detailed analysis is in Subsection 5.1. After that, $P_i = (a_1, a_2, \cdots, a_m)$ is determined as the components of the point z', where a_i is set to $k \cdot z_i^{nad}$ while other elements are set to 0. Finally, the angle vector $angle_p = (\alpha_1, \alpha_2)$ can be computed by Def. 1.

Definition 1. *By calculating the i-th node point $P_i = (0, \cdots, kz_i^{nad}, 0)$, α_i in the angle vector $angle_p = (\alpha_1, \alpha_2, \cdots, \alpha_m)$ is defined as follows:*

$$\alpha_i = arccos \frac{\overrightarrow{P_io} \cdot \overrightarrow{P_ip}}{\mid \overrightarrow{P_io} \mid \cdot \mid \overrightarrow{P_ip} \mid}, \tag{12}$$

where point o is the origin of coordinate or the ideal point of the current population that defined as $z^{ideal} = \left(z_1^{ideal}, z_2^{ideal}\right)$ where

$$z_i^{ideal} = \overset{n}{\underset{j=1}{min}} f_i(p_j). \tag{13}$$

Definition 2. *Assuming two solutions x and y, x is said to angle dominate y (denoted as $x \prec_{angle} y$) if the following condition holds:*

$$\forall i \in \{1, 2, \cdots M\} : \alpha_i^x \leq \alpha_i^y \wedge \exists i \in \{1, 2, \cdots, m\} : \alpha_i^x < \alpha_i^y. \tag{14}$$

The properties of the angle dominance criterion will be introduced, and the premise of these properties is that the parameter k has to be greater than 1. Suppose there are three different solutions p_1, p_2 and p_3. Their angle vectors are respectively denoted as follows: $angle_{p_1} = (\alpha_1, \alpha_2, \cdots, \alpha_m)$, $angle_{p_2} = (\beta_1, \beta_2, \cdots, \beta_m)$, and $angle_{p_3} = (\theta_1, \theta_2, \cdots, \theta_m)$.

Property 1. *The angle dominance is an irreflexive relation on the population.*

Proof. For any solution p_1, assuming its angle vector is $angle_{p_1} = (\alpha_1, \alpha_2, \cdots, \alpha_m)$, then $\forall i \in \{1, 2, \cdots, m\}, \alpha_i = \alpha_i$. Thus, the conditions for $p_1 \prec_{angle} p_1$ in Def. 2 do not hold. Hence, the angle dominance relation is irreflexive. □

Property 2. *The angle dominance is an asymmetric relation on the population.*

Proof. If $p_1 \prec_{angle} p_2$, then p_2 does not dominate p_1. From Def. 2, $p_1 \prec_{angle} p_2 \leftrightarrow \ \alpha_i \leq \beta_i, \ \forall i \in \{1, 2, \cdots, m\}$ and $\alpha_k < \beta_k, \ \exists k \in \{1, 2, \cdots, m\}$. Hence, if $p_1 \prec_{angle} p_2$, then p_2 does not dominate p_1. Therefore, the angle dominance is asymmetric. □

Property 3. *The angle dominance is a transitive relation on the population.*

Proof. If $p_1 \prec_{angle} p_2$ and $p_2 \prec_{angle} p_3$, from Def. 2, $\alpha_i \leq \beta_i, \beta_i \leq \theta_i, \ \forall i \in \{1, 2, \cdots, m\}$ and $\alpha_k < \beta_k, \beta_l < \theta_l, \ \exists k, l \in \{1, 2, \cdots, m\}$, so $p_1 \prec_{angle} p_3$. That is, the angle dominance is transitive. □

Property 4. *The angle dominance defines a strict partial order on the population.*

Proof. Since the angle dominance is an irreflexive, asymmetric and transitive relation on the population, it defines a strict partial order on the population. □

It is well-known that the traditional Pareto dominance criterion can not easily handle MaOPs, and the most effective way is to increase the dominance area of solutions. In terms of Def. 2, by amplifying the solutions' dominance area, the solutions that are far away from the Pareto front which are hardly eliminated by the Pareto dominance will be eliminated by the angle dominance. Therefore, angle dominance could effectively improve the convergence of population. Moreover, as with most dominance criteria, the angle dominance is Pareto compliant.

More specifically, in the angle dominance criterion, the solution closer to the Pareto front have higher priority and are first selected, and the solution far away from the Pareto front have larger dominance area. For example, in Fig. 4(a), it can be seen that p_1 and p_2 are non-Pareto dominated, and p_1' and p_2' are also non-Pareto dominated, while p_1 and p_2 Pareto dominate both p_1' and p_2'. However, in the sense of angle dominance, p_1 and p_2 angle dominate both p_1' and p_2', which verifies that the solutions having good convergence

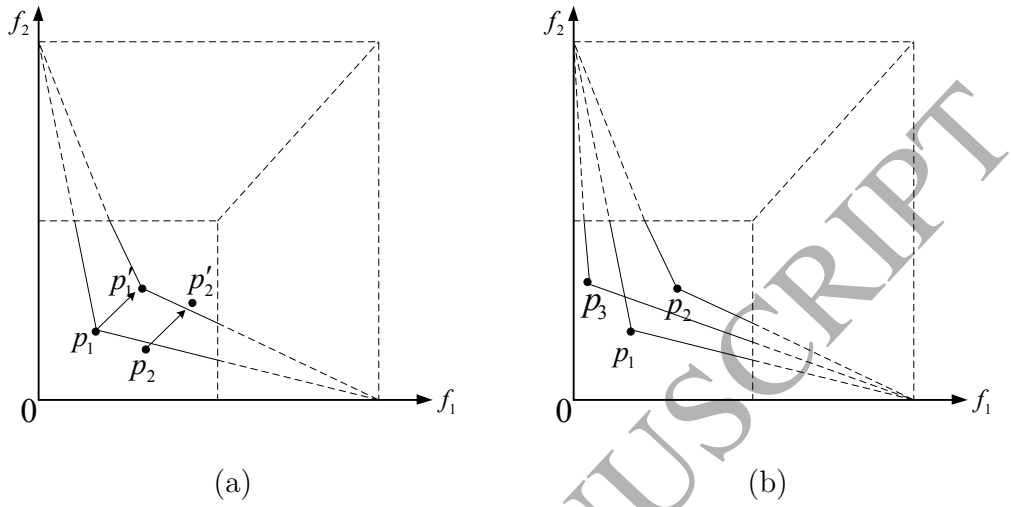
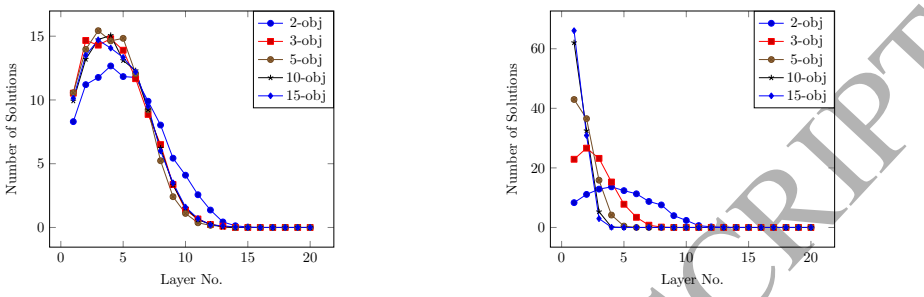


Figure 4: Properties of the angle dominance (a) An illustration of the ability of angle dominance to promote the convergence. There are four solutions $p_1 = (2, 2)$, $p_2 = (3, 1.7)$, $p_1' = (4, 4)$ and $p_2' = (5, 3.7)$. It is easy to know that $\overrightarrow{p_1p_1'} = \overrightarrow{p_2p_2'}$; (b) An illustration of the ability of angle dominance to promote extensity. There are three solutions $p_1 = (1.5, 1.5)$, $p_2 = (2.5, 2)$, $p_3 = (0.5, 3)$.

own higher priority for selection. Moreover, p_1 non-angle dominates p_2, but p_1' angle dominates p_2', which denotes that the solution far away from the Pareto front have larger dominance area, and vice versa. Consequently, in the fast sorting selection method [10], that of the non-angle dominated solutions in the layers will be decreased, which corresponds to the evolution of the population.

This work also investigated the number of solutions in each layer during the fast sorting selection between the angle dominance and Pareto dominance. Fig. 5 presents the experimental results on the DTLZ3 instances [11] with 2, 3, 5, 10, and 15 objectives. In Figs. 5(a) and (b), most solutions are crowded in the first layers. Specifically, from Fig. 5(a), with the increase of the number of objectives, the number of non-angle dominated solutions in each layer stays stable. But most non-Pareto dominated solutions are crowded in the first layer, as shown in Fig. 5(b). This phenomenon implies that the angle dominance can differentiate the solutions whereas Pareto dominance can not. The main reason is that the dominance area of the solution in the angle space is greater than it is in the objective space. It will result

in the angle dominance criterion having greater selection pressure than the Pareto dominance in the process of non-dominating sorting.



(a) Nondominated sorting of angle-dominance(b) Nondominated sorting of Pareto-dominance

Figure 5: The average number of solutions in all the nondominated layers under (a) the angle nondominated sorting and (b) the Pareto nondominated sorting, where the population size is 100, the number of runs is 30, and the test instance is DTLZ3.

Finally, except for the solutions close to the Pareto front, angle dominance also prefers the solutions close to the boundaries, which can promote the population's extensity. For example, there are three solutions located in different objective regions in Fig. 4(b), p_1 and p_3 are non-angle dominated but both of them dominate p_2. Notably, p_1 and p_3 are close to the Pareto front and the boundary respectively. Thus, the solutions close to the Pareto front and boundaries have high priority in the environmental selection.

3.2. Integrating Angle Dominance into NSGA-II

In this section, we take NSGA-II as an example to illustrate how to integrate the angle dominance criterion into its framework. This integration can also be applied to other Pareto dominance based MOEAs, e.g., SPEA2 [49].

The angle dominance based NSGA-II is denoted as NSGA-II+AD, whose framework is presented in Algorithm 1. Firstly, in line 2, the NSGA-II+AD randomly generates an initial population $\mathbf{P}$ with $\mid \mathbf{P} \mid = n$. For each iteration, the NSGA-II+AD applies some genetic operators, like mating selection [29], crossover [2] and mutation [6], as shown in lines 4-5, to generate the offspring population $\mathbf{Q}$. Then, NSGA-II+AD mixes $\mathbf{P}$ and $\mathbf{Q}$ into an interim population $\mathbf{R} = \mathbf{Q} \cup \mathbf{P}$, where $\mid \mathbf{R} \mid = 2n$. Finally, it uses the environmental selection from lines 7-20 to sort the mixed population and select the elite solutions into the next generation.

Algorithm 1 NSGA-II+AD Framework

Input: Number of objectives m, population size n, terminate condition $\mathcal{T}$
Output: The new population: $\mathbf{P}$;

1: $\mathbf{P} = \emptyset$, $i = 1$;
2: $\mathbf{P} = RandomInitiate(\mathbf{P})$;
3: **while** $\neg\mathcal{T}$ **do**
4: $\quad\mathbf{Q} = MatingSelection(\mathbf{P})$;
5: $\quad\mathbf{Q} = Variation(\mathbf{Q})$;
6: $\quad\mathbf{R} = \mathbf{P} \cup \mathbf{Q}$;
7: $\quad ComputeAngle(\mathbf{R})$;
8: $\quad(F_1, F_2, \cdots, F_l) = NonDominatedSort(\mathbf{R})$;
9: $\quad$**while** $\mid \mathbf{P} \mid + \mid F_j \mid \leq n$ **do**
10: $\quad\quad ComputeCrowdingDistance(F_i)$;
11: $\quad\quad\mathbf{P} = \mathbf{P} + F_j$ and $i = i + 1$;
12: $\quad$**end while**
13: $\quad$The last front to be included: $F_l = F_i$;
14: $\quad$**if** $\mid \mathbf{P} \mid = n$ **then**
15: $\quad\quad$**return** $\mathbf{P}$;
16: $\quad$**else**
17: $\quad\quad ComputeCrowdingDistance(F_l)$;
18: $\quad\quad$Calculate the number of points to be chosen from F_l: $k = n - \mid \mathbf{P} \mid$;
19: $\quad\quad$Choose k members one at a time from F_l to construct $\mathbf{P}$;
20: $\quad$**end if**
21: **end while**

In the environmental selection, firstly, we need to compute the angle vector for each solution according to Def. 1. Here, the nadir point of $\mathbf{R}$ is determined by identifying the maximum value (z_i^{nad}) of each objective, so the nadir point is $z^{nad} = \left(z_1^{nad}, z_2^{nad}, \cdots, z_m^{nad}\right)$, where $z_i^{nad} = \max_j^n f(x_i)$. Finally, from lines 8-20 in Algorithm 1, the elite solutions are selected into the next generation $\mathbf{P}$.

To achieve $\mathbf{P}$, first, the angle dominance criterion is used to sort the population $\mathbf{R}$ into different non-domination layers $(F_1, F_2, \cdots, F_l)$. Here, we assume that the last layer is F_l. Then, each non-domination layer is selected one at a time to construct a new population $\mathbf{P}$ and the crowding distance of each individual in F_i is computed, starting from F_1 and until the size of $\mathbf{P}$ is equal to n or is greater than n for the first time. In most situations, the last layer is only partially accepted. So, we should introduce a kind of distribution mechanism for selection of the last layer F_l. In lines 17-19, as in NSGA-II, we also apply the crowded-comparison approach, which computes the

crowding distance for every last layer member as the summation of objective-wise normalized distance between two neighboring solutions. Thereafter, the solutions with larger crowding distance values are chosen. When the termination condition is reached, the algorithm stops.

4. Experiment Design

This section is dedicated to the experimental design for investigating the performance of NSGA-II+AD. We first give the two well-defined test problem suites and performance metrics used in the experiment. Then, we briefly introduce eight MOEAs: NSGA-II [10], GrEA [44], MOEA/D [45], ε-MOEA [9], CDAS [38], PICEAg [39], MaOEARD [18], VaEA [42] and KnEA [47] which are used to validate the proposed criterion. Finally, the general experimental setting is provided for the comparative studies of these algorithms.

4.1. Test problems and Performance Metrics

Table 2: Properties of test problems and parameter setting in GrEA and ε-MOEA. The setting of *div* and ε corresponds to the different numbers of objectives of a problem. m denotes the number of tested objectives.

Problem	m	Properties	*div* in GrEA	ε in ε-MOEA	S in CDAS
DTLZ1	5,8,10,15,20	Linear, Multimodal	10,10,11,11,11	0.059,0.055,0.056,0.342,0.485	0.49,0.42,0.39,0.39,0.38
DTLZ2	5,8,10,15,20	Concave	10,10,11,11,11	0.192,0.290,0.308,0.320,0.320	0.49,0.40,0.38,0.32,0.30
DTLZ3	5,8,10,15,20	Concave, Multimodal	10,10,11,11,11	0.200,0.157,0.850,0.850,0.850	0.45,0.37,0.39,0.37,0.34
DTLZ4	5,8,10,15,20	Concave, Biased	10,10,11,11,11	0.193,0.290,0.308,0.380,0.382	0.49,0.45,0.45,0.45,0.45
DTLZ5	5,8,10,15,20	Concave, Degenerate	29,11,11,11,11	0.079,0.127,0.129,0.132,0.132	0.42,0.41,0.41,0.39,0.39
DTLZ6	5,8,10,15,20	Concave, Degenerate, Biased	24,50,50,50,50	0.355,1.150,1.450,1.800,1.800	0.42,0.40,0.39,0.38,0.38
DTLZ7	5,8,10,15,20	Mixed, Disconnected, Biased	9,8,5,4,4	0.158,0.225,0.560,0.565,0.569	0.49,0.48,0.48,0.48,0.48
WFG1	5,8,10,15,20	Mixed, Biased, Scaled	10,10,11,11,11	0.210, 0.322, 0.330, 0.353, 0.355	0.49,0.49,0.49,0.49,0.49
WFG2	5,8,10,15,20	Convex, Disconnected, Multimodal, Scaled	10,10,11,11,11	0.253,0.423,0.426,0.492,0.611	0.49,0.49,0.49,0.49,0.49
WFG3	5,8,10,15,20	Degenerate, Non-separable, Scaled	35,29,11,11,11	0.420,0.762,0.900,1.502,2	0.49,0.45,0.45,0.45,0.45
WFG4	5,8,10,15,20	Concave, Multimodal, Scaled	10,10,11,11,11	0.600,1.349,2.082,4.793,6.746	0.49,0.47,0.47,0.47,0.47
WFG5	5,8,10,15,20	Concave, Deceptive, Scaled	10,10,11,11,11	0.600,1.349,2.082,4.793,6.746	0.49,0.48,0.48,0.48,0.48
WFG6	5,8,10,15,20	Concave, Non-separable, Scaled	10,10,11,11,11	0.600,1.349,2.082,4.793,6.746	0.49,0.48,0.48,0.48,0.48
WFG7	5,8,10,15,20	Concave, Biased, Scaled	10,10,11,11,11	0.600,1.349,2.082,4.793,6.746	0.49,0.48,0.48,0.48,0.48
WFG8	5,8,10,15,20	Concave, Biased, Non-separable, Scaled	10,10,11,11,11	0.600,1.349,2.082,4.793,6.746	0.49,0.48,0.48,0.48,0.48
WFG9	5,8,10,15,20	Concave, Biased, Multimodal, Deceptive, Non-separable, Scaled	10,10,11,11,11	0.600,1.349,2.082,4.793,6.746	0.49,0.48,0.48,0.48,0.48

As a basis for the comparisons, two well-known test suites for many-objective optimization, DTLZ [11] and WFG [21], are selected in the experiments. All these problems can be scaled to any number of objectives and decision variables. For each problem, the number of objectives is set to 5, 8, 10, 15 and 20, respectively. As recommended in [21], the number of decision variables is set to $s = m + K - 1$, where m is the objective number, $K = 5$ for DTLZ1, $K = 10$ for DTLZ2 to DTLZ6 and $K = 20$ for DTLZ7. According to [21], the number of decision variables is set to $s = K + L$ for WFG test suite, where the position-related variable $K = 2 \cdot (m - 1)$, and the distance-related variable $L = 20$. The two test suites have been used to challenge different

abilities of algorithms, and the properties of the tests problems are shown in Table 2.

In order to compare the performance of the selected algorithms, two widely-used quality metrics, inverted generational distance (IGD) [50] and hypervolume (HV) [41], are introduced.

The former is a metrics to measure the distance of a solution set from the Pareto front, and is used to analyze the influence of different values k on the performance of NSGA-II+AD. However, it is almost always the case that the true Pareto front is unknown in real-world applications. In this case, a common practice is to specify a set of reference point based on the Pareto dominance as an approximation of the Pareto front, and then calculate the average distance from each reference point to the nearest solution in the solution set. Mathematically, let P^* be a reference set representing the Pareto front, and P be an obtained solution set. Then, the IGD value of the obtained solution set P is defined as follows:

$$IGD\left(P\right) = \frac{\sum_{x^* \in P^*} d(x*, P)}{|P^*|}, \tag{15}$$

where $|P^*|$ denotes the size of P^* (the number of points in P^*) and $d\left(x^*, P\right)$ is the minimum Euclidean distance from x^* to P ($d\left(x^*, P\right) = min_{x \in P} \| f(x^*) - f(x) \|$). A low IGD value is preferable, which indicates that the obtained solution set is close to the Pareto front and that it has good diversity.

The last one is used to evaluate the performance of algorithms on DTLZ and WFG, which measures the volume of the objective space enclosed by a Pareto front approximation and a reference point in the objective space. The Pareto front approximation with larger HV values is better. Then, the HV metrics can be described as the Lebesgue measure Λ of the union hypercubes h_i defined by a solution p_i in the approximation and the reference point x_{ref} as follows:

$$HV = \Lambda\left(\{\bigcup_i h_i \mid p_i \in P\}\right) = \Lambda\left(\bigcup_{p_i \in P} \{x \mid p_i \prec x \prec x_{ref}\}\right). \tag{16}$$

Following the recommendation in [25], that reference point x_{ref} is slightly larger than z_i^{nad} is suitable since the balance between convergence and diversity of the solution set is well emphasized. In our experiments, We first normalize the objective value of obtained solution according to the range of the problem's PF, then set the reference point to 1.1 times of the z_i^{nad}.

Otherwise, considering that the exact calculation of the HV measure is computationally highly demanding, and current algorithms are exponential in the number of objectives. we suggest a methodology based on Monte Carlo sampling [1] to estimate the HV result of a solution set [3], where the number of sampling points is set to 1,000,000.

4.2. Nine Other Algorithms in Comparison

To verify the performance of NSGA-II+AD, the following nine peer algorithms are considered:

NSGA-II [10]: It is the most well-known and frequently-used MOEA in the literature. In NSGA-II, the population is sorted based on non-domination into each front. The first front being completely non-dominated set and the second front being dominated by the individuals in the first front and the front goes so on. Moreover, the secondary ranking criterion for solutions on the same front is called crowding distance. Large average crowding distance will result in better diversity in the population. It is worth pointing that NSGA-II as a well-established algorithm in the area, its performance has been outperformed by many other algorithms, as shown in [17].

GrEA [44]: It adopts the adaptive construction of grids to strengthen the selection pressure toward the PF while maintaining an extensive and uniform distribution among solutions. To this end, two concepts (i.e., grid dominance and grid difference), three gird-based criteria (i.e., grid ranking, grid crowding distance, and coordinate point distance), and a fitness adjustment strategy are incorporated into GrEA.

MOEA/D [45]: It is a representative decomposition-based algorithm using a decomposition method to decompose the MOP into a number of scalar optimization problems. In this paper, considering that penalty-based boundary (PBI) [45] is more effective than other decomposition methods for many-objective optimization in a recent study [46], we select it as the aggregation function for MOEA/D.

ε-MOEA [9]: It is a steady-state algorithm using the ε-dominance criterion. The objective space is divided into hyperboxes, whose size can be

[1] The main idea is that not the actual indicator values are important, but rather the rankings of solutions induced by the hypervolume indicator. To this end, samples of objective vectors are randomly drawn and the proportion of objective vectors that are solely dominated by a specific individual represents an estimate for the hypervolume contribution of this individual.

adjusted by the choice of ε. Each hyperbox is assigned at most a single point by the ε-dominance and the distance from solutions to the utopia point in the hyperbox. ε-MOEA has been verified to perform well for many-objective optimization problems in a recent study [36].

CDAS [38]: It can control the degree of expansion or contraction of the dominance area of solutions using a user-defined parameter S. Modifying the dominance area of solutions changes their dominance relation inducing a ranking of solutions that is different to conventional dominance. In this paper, we integrated CDAS into NSGA-II for comparison experiments with the proposed algorithms.

PICEAg [39]: It introduces a new concept of preference-based coevolutionary algorithm (PICEA), which coevolves a family of decision-maker preferences together with a population of candidate solutions. In PICEAg, the preferences gain higher fitness by being satisfied by fewer candidate solutions, and the candidate solutions gain fitness by meeting as many preferences as possible.

MaOEARD [18]: It includes two stages: 1) search for the target points around the true Pareto area and constrain the objective search space and 2) a diversity improvement strategy is then applied to facilitate the extent and distribution of the population while constantly updating target points to ensure convergence. In MaOEARD, the performance improvement is gained directly by overcoming two fundamental challenges existing in MaOPs: 1) extremely large objective space and 2) ineffectiveness of Pareto-dominance.

VaEA [42]: It is a vector angle based evolutionary algorithm and uses two principles: the maximum-vector-angle-first principle is used in the environmental selection to guarantee the extensity and the uniformity of solution set; the worse-elimination principle replaces worse solutions in terms of the convergence.

KnEA [47]: It is knee point driven evolutionary algorithm, and the basic ideal of it is that knee points are naturally most preferred among non-dominated solutions if no explicit user preferences are given. The preference of knee points can be seen as a bias towards larger hypervolume which assists in achieving good convergence and diversity.

4.3. General Experimental Setting

In this section, the general parameter settings for all conducted experiments are given as follows:

Table 3: Terminate condition (the number of evaluations allowed).

Problem	DTLZ(1,3,6)	DTLZ(2,4,5,7)	WFG1-WFG9
Evaluations	100,000	30,000	30,000

Table 4: The settings of the reference point of MOEA/D based on the normal-boundary intersection (NBI) method and its two-layered version (NBI2).

Number-of-objectives (m)	5	8	10	15	20
Partitions (p)	5	3	2,2	2	2
Number-of-reference-point	126	120	55+55=100	120	210
Population-size	128	120	112	120	210

1. *Number of runs and termination Criterion*: All algorithms are run 30 times independently for each test instance. The number of termination criterion of an algorithm is a predefined number of evaluations. As shown in Table 3, for the DTLZ1, DTLZ3 and DTLZ6, it is set to 100,000, and for the other test problems (DTLZ2, DTLZ4, DTLZ5, DTLZ7 and WFG1-WFG9), it is set to 30,000.
2. *Parameters for crossover and mutation*: A crossover probability $p_c = 1$ and a mutation probability $p_m = \frac{1}{n}$ (where n denotes the number of decision variables) are used. The distribution index is set as $\eta = 20$ for both the SBX and PM operators.
3. *Population and archive size and parameter setting in selected algorithms*: For all selected algorithms, the population size is set to 100, and the archive is also maintained with the same size if required. In ε-MOEA, the size of the archive set is determined by a parameter ε. To guarantee a fair comparison, this paper set ε as shown in Table 2, and the archive size is approximately the same as that of the other algorithms. GrEA requires a grid division parameter *div*, and the settings of *div* are shown in Table 2. Meanwhile, Table 2 also gives the parameter S of CDAS. The number of goals is set to $m \cdot 100$ in PICEA-g [39]. For MOEA/D [45], a preset of weight vectors are needed to maintain the diversity of population. The normal-boundary intersection (NBI) method [35] and the two-layered version (NBI2) [8] are used in MOEA/D. In consideration of the combinatorial nature of uniformly distributed weight vectors, the population size should be as close as possible to the number of weight vectors. As shown in Table 4, the number of objectives and the division parameter of NBI are m and p, respectively.

5. Results and Discussions

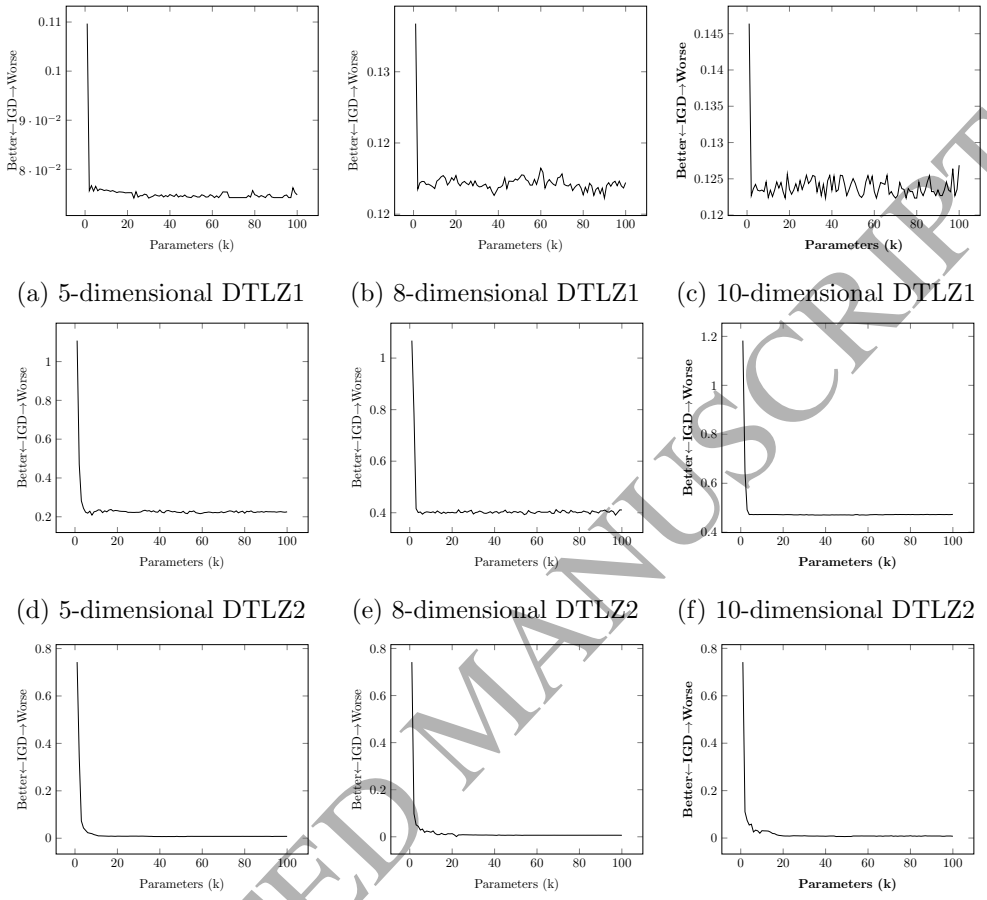
5.1. Parameter Sensitivity Analysis

In most of the improved dominance criteria, a series of parameters must be set. But, these parameters bring some difficulties to the application of algorithms. For example, ε-dominance needs the user to set difference parameters for varying MOPs with different dimensions, and so does the CDAS. In angle dominance criterion, there is a specific parameter k which is the magnification factor of the nadir point. In this section, we investigate the effect of parameter k on the performance of the angle dominance criterion.

In the experiments, angle dominance was combined with NSGA-II, which is denoted as NSGA-II+AD (presented in Section 3 in detail). The algorithm was run for 30 times independently with varying values of k, where $k \in [1, 100]$, on the set of DTLZ instances, respectively. Fig. 6 presents the experimental results regarding the IGD values on 5-, 8- and 10-objective DTLZ1, DTLZ2 and DTLZ6 with linear Pareto front, spherical Pareto front and one-dimensional linear manifold, respectively. The population size was set to 100, and the maximum number of generations was set to 1000 on DTLZ1 and DTLZ6 and 300 on DTLZ2, respectively.

From Fig. 6, the trend of the change of IGD values is similar on varying problems. Note that in the figure the IGD values are displayed in logarithm. The IGD values first decrease sharply, and then begin to level out and fluctuate in a very small range. Specifically, on DTLZ1 problems with 5, 8 and 10 objectives, the IGD values are stable about 0.221, 0.400 and 0.470 respectively with the increase of the k from 2 to 100. This shows that when k is greater than 2, the performance of the algorithm on the DTLZ1 test problem is relatively stable. When it comes to DTLZ2 in 5, 8 and 10 objectives, the IGD values level out and fluctuate around 0.065, 0.025 and 0.47 respectively when $k > 10$. In addition, the IGD values on DTLZ2 tend to be more stable than that on DTLZ1. This shows that the sensitivity of parameter k on DTLZ2 is smaller than that of parameter k on DTLZ1. Finally, when $k > 10$, the IGD vales level out around 0.0071, 0.0077 and 0.0082 on DTLZ6 with 5, 8 and 10 objectives, respectively.

According to the above parametric analysis, it can be found that, when $k > 10$, the performance of the algorithm is stable and the change of parameter k has very little effect on the performance of the angle dominance. In fact, when $k = 1$, an individual has the largest angle dominated space. In this case, there may be a total order relationship between solutions. As



(a) 5-dimensional DTLZ1 (b) 8-dimensional DTLZ1 (c) 10-dimensional DTLZ1

(d) 5-dimensional DTLZ2 (e) 8-dimensional DTLZ2 (f) 10-dimensional DTLZ2

(g) 5-dimensional DTLZ6 (h) 8-dimensional DTLZ6 (j) 10-dimensional DTLZ6

Figure 6: Study of different settings of the k parameter. Where the abscissa represents the k parameter and the ordinate indicates the IGD value. The smaller the IGD value, the better the performance.

parameter k increases, the solution's angle dominant space will gradually decrease, and the angle dominance progressively approximates Pareto dominance. Thus the k cannot be set too small or too large. Given that when k is between $[10, 100]$, the performance of angle dominance on different problems is stable, we consider a value (50) near the media position of $[10, 100]$ as the magnification factor of the angle dominated space.

Table 5: IGD results (mean and SD) of the NSGA-II and NSGA-II+AD on the DTLZ series of problems. The best mean of the algorithms for each problem instance is shown with a gray background and the values in parentheses are the SDs

Problem	5-objective		8-objective		10-objective	
	NSGA-II	NSGA-II+AD	NSGA-II	NSGA-II+AD	NSGA-II	NSGA-II+AD
DTLZ1	2.2573e-1 (2.75e-1) -	7.5862e-2 (1.78e-3)	6.8256e+0 (1.32e+1) -	1.1755e-1 (1.56e-3)	9.8175e+0 (1.05e+1) -	1.2408e-1 (1.08e-3)
DTLZ2	2.6194e-1 (3.60e-2) -	2.2136e-1 (4.08e-3)	1.9706e+0 (3.80e-1) -	4.0175e-1 (3.67e-3)	1.8670e+0 (4.14e-1) -	4.7152e-1 (4.13e-3)
DTLZ3	9.8325e+0 (2.24e+1) -	1.3318e+0 (3.49e+0)	7.6080e+2 (2.59e+2) -	4.0164e-1 (5.35e-3)	1.0940e+3 (3.81e+2) -	4.7375e-1 (5.09e-3)
DTLZ4	2.5532e-1 (1.32e-2) -	2.2062e-1 (3.19e-3)	1.7507e+0 (2.72e-1) -	3.9981e-1 (2.55e-3)	1.7775e+0 (3.34e-1) -	4.7425e-1 (5.44e-3)
DTLZ5	1.0950e-1 (3.71e-2) -	7.0084e-3 (1.74e-4)	2.4631e-1 (9.11e-2) -	7.5029e-3 (6.14e-4)	4.1743e-1 (1.87e-1) -	8.0859e-3 (7.19e-4)
DTLZ6	3.1606e+0 (9.20e-1) -	7.1727e-3 (4.83e-4)	6.0747e+0 (9.88e-1) -	7.7340e-3 (8.06e-4)	6.2883e+0 (7.80e-1) -	8.2108e-3 (1.07e-3)
DTLZ7	4.9868e-1 (1.27e-1) ≈	4.6150e-1 (9.97e-2)	2.5557e+0 (1.82e+0) -	8.6303e-1 (1.62e-1)	5.8761e+0 (3.57e+0) -	1.1052e+0 (2.17e-1)
+/ − / ≈	0/6/1		0/7/0		0/7/0	

"+", " -" and "≈" indicate that the result is significantly better, significantly worse and statistically similar to that obtained by NSGA-II+AD, respectively.

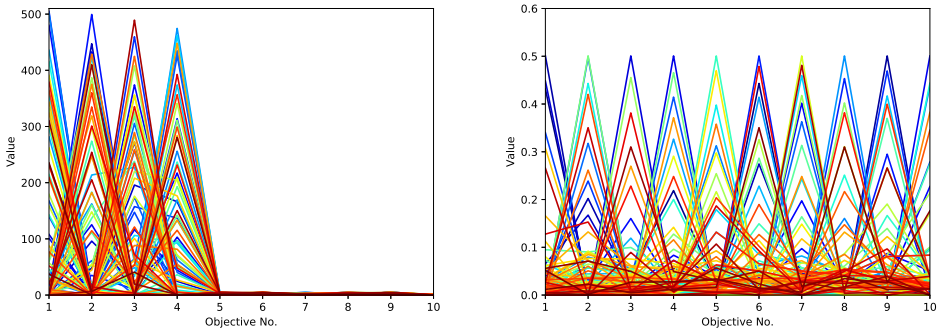
5.2. NSGA-II vs. NSGA-II+AD

Table 5 shows the results of the two algorithms on the DTLZ test suite regarding the mean and standard deviation (SD) values, where IGD was used for DTLZ problem. The better result regarding the mean for each problem is highlighted. Moreover, in order to have statistically sound conclusions, Wilcoxon's rank sum test [16] at a significance level of 0.05 was conducted on the experimental results by two competing algorithms, where the symbols "+", "−" and "≈" indicate that the result by NSGA-II is significantly better, significantly worse and statistically similar to the obtained by NSGA-II+AD, respectively.

As can be seen from Table 5, the performance of NSGA-II has a clear improvement when the angle dominance is applied to the algorithm, achieving a better value for all the 21 test instances. Also, for most of the problems on which NSGA-II+AD outperforms NSGA-II, the results have statistical significance (20 out of the 21 problems). Fig. 7 shows the final solutions of a single run of NSGA-II and NSGA-II+AD on the 10-objective DTLZ1 by parallel coordinates [2] [34]. The global optimal front of this problem is a linear hyper-plane satisfying $\sum_{i=1}^{M} f_i = 0.5$ in the range $f_i \in [0, 0.5]$. This particular run is associated with the result which is the closest to the mean IGD value. It is clear from the figure that the convergence performance of NSGA-II is significantly improved when the angle dominance is applied to the algorithm.

[2]Parallel coordinates display multidimensional data (a set of vectors) in a two-dimensional graph, with each dimension of the original data being translated onto a vertical axis in the graph, and a vector is represented as a polyline with vertices on the axes.

Through the above analysis, we can show that integrating the angle dominance criterion into NSGA-II is going to be effective.



(a) NSGA-II (b) NSGA-II+AD

Figure 7: The final solution set of the two algorithms on the ten-objective DTLZ1, shown by parallel coordinates. Where the abscissa represents the objective dimension and the ordinate indicates the objective value.

5.3. Comparison with well-established algorithms

In this section, NSGA-II+AD is applied to compare with four well-established MOEAs to demonstrate whether the NSGA-II+AD is competitive when addressing MaOPs. The compared MOEAs are GrEA, MOEA/D, ε-MOEA and CDAS, respectively.

Table 6 shows the comparative results of the investigated algorithms on the DTLZ problem suite regarding the mean and standard deviation (SD) of the HV values, where the gray background and bold represent the best and second best results, respectively. On the whole, NSGA-II+AD is the best performing algorithm and followed closely by CDAS. The former gains 8 best results and 25 second-best results, and the latter owns 14 best results and 7 second-best results. Although the number of best results obtained by NSGA-II+AD is lightly less than that obtained by CDAS, the sum of the best and second-best results of the former far exceeds that of the latter. In particular, NSGA-II+AD only performs worse than GrEA and MOEA/D on DTLZ2 but performs better on other problems, while CDAS does best on DTLZ3 and degradation problems (e.g., DTLZ5 and DTLZ6). GrEA is the third-best algorithm, and it performs better on DTLZ2 and DTLZ4 compared to other

four algorithms. It was followed by MOEA/D, which produced three and one of the best results on DTLZ1 and DTLZ3, respectively. Compared with the above four algorithms, ε-MOEA is inferior in DTLZ test problems. The sum of its best and second-best results is zero. As for the statistical significance analysis, it can be observed that the difference between NSGA-II+AD and the peer algorithms is significant on most of the test instances. Specifically, the proportion of the test instances where NSGA-II+AD outperforms GrEA, MOEA/D, ε-MOEA and CDAS with statistical significance is 23/35, 28/35, 35/35 and 14/35, respectively. Conversely, the proportion of the instances where NSGA-II+AD performs worse than GrEA, MOEA/D, ε-MOEA and CDAS with statistical significance is 9/35, 7/35, 0/35 and 7/35, respectively.

Table 6: HV results (mean and SD) of the five algorithms on the DTLZ test suite. The best and the second mean among the algorithms for each problem instance are highlighted in gray background and bold, respectively.

Problem	M	D	NSGA-II+AD	GrEA	MOEA/D	ε-MOEA	CDAS
DTLZ1	5	9	**4.5874e-2 (2.34e-4)**	4.3054e-2 (5.20e-3) ≈	4.8845e-2 (1.33e-5) +	4.0278e-2 (2.44e-3) −	4.4901e-2 (1.76e-4) −
	8	12	**8.1246e-3 (1.90e-5)**	5.0269e-3 (2.10e-3) −	8.1721e-3 (5.71e-5) +	6.7463e-3 (2.35e-4) −	8.0541e-3 (2.41e-5) −
	10	14	**2.4945e-3 (5.31e-6)**	9.8809e-4 (4.37e-4) −	2.5203e-3 (2.16e-6) +	6.8710e-4 (8.30e-4) −	2.4814e-3 (6.04e-6) −
	15	24	1.2705e-4 (7.42e-8)	7.7048e-6 (1.39e-5) −	1.1155e-4 (9.71e-7) −	1.0507e-5 (1.12e-5) −	1.2691e-4 (8.21e-8) −
	20	29	6.4082e-6 (3.72e-9)	1.0412e-6 (2.08e-6) −	5.3096e-6 (5.08e-8) −	2.2039e-8 (2.62e-8) −	6.3263e-6 (2.77e-8) −
DTLZ2	5	14	1.1529e+0 (1.14e-2)	1.2636e+0 (2.71e-3) +	1.2453e+0 (8.43e-4) +	8.0239e-1 (6.70e-2) −	**1.1545e+0 (1.07e-2)** ≈
	8	17	1.7675e+0 (1.64e-2)	1.9544e+0 (2.37e-3) +	1.8574e+0 (8.56e-3) +	1.1385e+0 (4.94e-2) −	**1.7699e+0 (1.66e-2)** ≈
	10	19	**2.2489e+0 (2.09e-2)**	2.4650e+0 (4.66e-3) +	2.3893e+0 (1.86e-2) +	1.1831e+0 (6.06e-2) −	2.2471e+0 (1.77e-2) ≈
	15	24	3.8630e+0 (3.09e-2)	3.9967e+0 (4.44e-2) +	7.7519e-1 (3.10e-1) −	1.4541e+0 (3.25e-1) −	**3.9005e+0 (2.03e-2)** +
	20	29	6.3982e+0 (4.21e-2)	6.3005e+0 (3.93e-2) −	4.0909e+0 (1.57e-1) −	2.4917e+0 (6.33e-1) −	**6.3836e+0 (4.88e-2)** ≈
DTLZ3	5	14	**1.1393e+0 (1.61e-2)**	4.6010e-1 (1.93e-1) −	1.2343e+0 (1.24e-2) +	9.9742e-1 (6.30e-2) −	1.1379e+0 (1.61e-2) ≈
	8	17	**1.7496e+0 (3.05e-2)**	1.6729e-1 (1.64e-1) −	1.1804e+0 (6.74e-1) ≈	6.4395e-1 (1.12e+0) −	1.7500e+0 (3.04e-2) ≈
	10	19	**2.2231e+0 (3.68e-2)**	0.0000e+0 (0.00e+0) −	1.6018e+0 (1.02e+0) ≈	0.0000e+0 (00e+0) −	2.2293e+0 (3.31e-2) ≈
	15	24	**3.8375e+0 (6.73e-2)**	0.0000e+0 (0.00e+0) −	3.7533e-1 (4.32e-3) −	0.0000e+0 (00e+0) −	3.8456e+0 (8.05e-2) ≈
	20	29	**6.3332e+0 (1.34e-1)**	0.0000e+0 (0.00e+0) −	2.9665e+0 (1.77e+0) −	0.0000e+0 (00e+0) −	6.3747e+0 (6.35e-2) ≈
DTLZ4	5	14	**1.1774e+0 (9.44e-3)**	1.2484e+0 (4.85e-2) +	7.8540e-1 (3.52e-1) −	8.2801e-1 (1.75e-1) −	1.1728e+0 (9.42e-3) −
	8	17	**1.8037e+0 (1.28e-2)**	1.9578e+0 (1.48e-3) +	1.5218e+0 (2.00e-1) −	1.5929e+0 (1.40e-1) −	1.7952e+0 (1.55e-2) −
	10	19	**2.2857e+0 (1.91e-2)**	2.4632e+0 (1.87e-2) +	1.5752e+0 (3.72e-1) −	2.2299e+0 (3.56e-2) −	2.2768e+0 (1.56e-2) −
	15	24	**3.9393e+0 (2.37e-2)**	4.0581e+0 (2.89e-3) +	1.7627e+0 (7.44e-1) −	3.5952e+0 (4.42e-2) −	3.9342e+0 (1.59e-2) ≈
	20	29	6.6009e+0 (1.03e-2)	6.4635e+0 (9.78e-2) ≈	3.0669e+0 (1.15e+0) −	6.5034e+0 (8.81e-2) −	**6.5939e+0 (1.14e-2)** −
DTLZ5	5	14	**9.1838e-3 (2.29e-5)**	8.6099e-3 (1.83e-4) −	8.8915e-3 (1.98e-5) −	5.0025e-3 (1.58e-3) −	9.2124e-3 (2.31e-5) +
	8	17	**1.9553e-5 (5.01e-8)**	2.0733e-6 (2.51e-6) −	1.8214e-5 (2.31e-6) −	1.2127e-6 (1.58e-6) −	1.9610e-5 (5.72e-8) +
	10	19	**6.1910e-8 (1.57e-10)**	7.3977e-10 (1.92e-9) −	6.0942e-8 (1.83e-10) −	4.5008e-9 (6.94e-9) −	6.2169e-8 (1.93e-10) +
	15	24	**8.7763e-17 (3.28e-19)**	1.1816e-20 (3.54e-20) −	8.5604e-17 (3.79e-19) −	5.1899e-18 (8.98e-18) −	8.8195e-17 (2.87e-19) +
	20	29	**2.2299e-29 (6.74e-32)**	0.0000e+0 (0.00e+0) −	2.2018e-29 (5.15e-32) −	8.7874e-30 (4.54e-30) −	2.2397e-29 (7.58e-32) ≈
DTLZ6	5	14	**9.1955e-3 (2.67e-5)**	7.4988e-3 (4.50e-4) −	8.8448e-3 (4.35e-5) −	6.3572e-3 (2.03e-4) −	9.2175e-3 (2.96e-5) ≈
	8	17	**1.9550e-5 (4.72e-8)**	0.0000e+0 (0.00e+0) −	1.8264e-5 (1.96e-6) −	1.1121e-6 (1.93e-6) −	1.9593e-5 (6.81e-8) ≈
	10	19	**6.2084e-8 (1.39e-10)**	0.0000e+0 (0.00e+0) −	6.0896e-8 (1.16e-10) −	0.0000e+0 (0.00e+0) −	6.2115e-8 (1.64e-10) +
	15	24	**8.7977e-17 (3.34e-19)**	0.0000e+0 (0.00e+0) −	8.5313e-17 (2.74e-19) −	0.0000e+0 (0.00e+0) −	8.8216e-17 (2.34e-19) +
	20	29	**2.2346e-29 (9.71e-32)**	0.0000e+0 (0.00e+0) −	2.1928e-29 (2.63e-31) −	0.0000e+0 (0.00e+0) −	2.2388e-29 (6.46e-32) ≈
DTLZ7	5	24	**2.0167e+0 (3.51e-2)**	2.2262e+0 (1.99e-2) +	1.3316e-1 (1.26e-1) −	1.6548e+0 (1.71e-2) −	1.9838e+0 (7.22e-2) ≈
	8	27	2.3221e+0 (1.61e-2)	1.7276e+0 (1.08e-1) −	2.7269e-2 (4.43e-2) −	1.2244e-0 (2.62e-1) −	**2.2132e+0 (2.91e-2)** −
	10	29	2.4041e+0 (3.75e-2)	9.0524e-1 (1.67e-1) −	7.1790e-2 (1.73e-1) −	1.5370e-1 (1.49e-1) −	**2.1785e+0 (8.57e-2)** −
	15	24	2.2840e+0 (1.72e-1)	**1.9208e+0 (1.06e-1)** ≈	3.6223e-1 (9.43e-2) −	4.3057e-2 (4.34e-2) −	1.3168e+0 (3.57e-1) −
	20	29	1.7842+0 (1.98e-1)	**1.6866e+0 (6.00e-2)** −	1.0922e-5 (1.96e-5) −	2.4452e-3 (1.96e-3) −	3.5167e-1 (1.14e-1) −
$+/-/\approx$				9/23/3	7/28/2	0/35/0	7/14/14

"+", "-" and "≈" indicate that the result is significantly better, significantly worse and statistically similar to that obtained by NSGA-II+AD, respectively.

Table 7 gives the comparative results of the four well-established algorithms on the WFG problems with 5, 8, 10, 15 and 20 objectives. As shown, the best-performing algorithm is GrEA, which has a clear advantage over the other 4 algorithms on more of the test instances. In addition, NSGA-II+AD is second only to GrEA. Specifically, GrEA obtains the best and second-best

HV results on 26 and 5 out of the 45 instances respectively, and NSGA-II+AD on 17 and 14 respectively. In fact, GrEA, despite being brought up several years ago, has shown to be very competitive against state of the arts on this particular type of problems (i.e., WFG4-WFG9) [31]. NSGA-II+AD is less than GrEA in terms of the number of best results, but the sum of the best results and the second-best results obtained by the two algorithms is the same. For other three algorithms, they are inferior to the above two algorithms in WFG test problems. CDAS and ε-MOEA only have the second-best solutions, while MOEA/D performs poorly on all instances. Concerning the statistical results, in addition to GrEA which is significantly better than our method on 25 instances and worse than our method on 20 instances, our proposed method outperforms other algorithms in most of the test instances. For example, NSGA-II+AD performs significantly better than MOEA/D, ε-MOEA and CDAS, respectively on 41, 36 and 30 test instances over all 45 problems.

5.4. Comparison with state-of-the-art algorithms

In this section, we select four state-of-the-art algorithms for comparison tests with NSGA-II+AD. The four algorithms are PICEAg, MaOEARD, VaEA and KnEA, respectively.

As can be observed from Table 8, NSGA-II+AD performs best, presenting a clear advantage over the other four algorithms on the majority of the test instances. More specifically, It obtains the best and second best HV results on 24 and 7 out of the 35 test instances respectively. KnEA works well for some relatively simple problems (e.g., DTLZ2 and DTLZ4). For multi-mode problems (e.g., DTLZ1 and DTLZ3), the performance of KnEA is far inferior to other algorithms. The main reason is that it is difficult to converge on this problem. For other three algorithms, they obtain worse HV values on almost all the test instances, except for five-objective DTLZ1. Statistically, NSGA-II+AD shows significant improvement over other algorithms on most of the test instances. Specifically, the proportion of the test instances where NSGA-II+AD performs better than PICEAg, MaOEARD, VaEA and KnEA is 30/35, 32/35, 30/35 and 25/35, respectively. Conversely, the proportion that NSGA-II+AD is defeated by the peer algorithms only is 1/35, 1/35, 4/35 and 10/35 for PICEAg, MaOEARD, VaEA and KnEA, respectively.

Different from the performance on the DTLZ test problem, NSGA-II+AD is not as good as the other algorithms on the WFG test problem. As shown

Table 7: HV results (mean and SD) of the five algorithms on the WFG test suite. The best and the second mean among the algorithms for each problem instance are highlighted in gray background and bold, respectively.

Problem	M	D	NSGA-II+AD	GrEA	MOEA/D	ε-MOEA	CDAS
WFG1	5	28	5.8425e+3 (1.20e+2)	4.9280e+3 (4.65e+2) −	4.5726e+3 (4.43e+2) −	2.1211e+3 (4.41e+1) −	**5.7515e+3 (1.68e+2)** −
	8	34	2.0632e+7 (9.97e+3)	1.5161e+7 (2.31e+6) −	1.3706e+7 (1.94e+6) −	7.2539e+6 (9.04e+5) −	**2.0264e+7 (7.46e+5)** −
	10	38	8.6472e+9 (3.12e+6)	6.3780e+9 (6.78e+8) −	3.1680e+9 (5.92e+8) −	2.7023e+9 (3.06e+8) −	**8.3454e+9 (4.99e+8)** −
	15	48	1.3899e+17 (3.42e+13)	**1.3785e+17 (9.59e+14)** −	4.8927e+16 (7.88e+15) −	3.4102e+16 (1.49e+15) −	1.2142e+17 (1.37e+16) −
	20	58	1.0837e+25 (5.48e+23)	1.0029e+25 (6.47e+23) −	2.4862e+24 (2.81e+23) −	2.2694e+24 (3.39e+23) −	**1.0589e+25 (5.90e+23)** −
WFG2	5	28	6.0775e+3 (5.10e+0)	5.9312e+3 (3.17e+1) −	5.6273e+3 (9.99e+1) −	5.7445e+3 (7.87e+1) −	**6.0452e+3 (1.04e+1)** −
	8	34	2.1991e+7 (2.89e+4)	2.1401e+7 (1.05e+5) −	1.8859e+7 (1.18e+6) −	1.9385e+7 (3.69e+6) −	**2.1580e+7 (1.08e+5)** −
	10	38	9.5767e+9 (1.91e+7)	9.3151e+9 (6.73e+7) −	8.3644e+9 (3.66e+8) −	8.6920e+9 (1.04e+9) −	**9.3715e+9 (4.84e+7)** −
	15	48	1.7696e+17 (2.46e+14)	1.7341e+17 (1.08e+15) −	1.5288e+17 (7.28e+15) −	**1.7499e+17 (2.00e+15)** −	1.7040e+17 (1.60e+15) −
	20	58	1.6937e+25 (7.42e+22)	**1.6559e+25 (1.19e+23)** −	1.2975e+25 (8.31e+23) −	1.5137e+25 (1.22e+24) −	1.6273e+25 (1.69e+23) −
WFG3	5	28	2.8919e+0 (7.87e-2)	1.8652e+0 (3.52e-1) −	3.8335e-1 (3.99e-1) −	1.2746e-1 (2.21e-1) −	**2.7400e+0 (1.18e-1)** −
	8	34	**1.2908e-2 (3.55e-3)**	1.8515e-2 (6.70e-3) +	0.0000e+0 (0.00e+0) −	0.0000e+0 (0.00e+0) −	5.6306e-3 (3.97e-3) −
	10	38	**1.0244e-5 (1.23e-5)**	1.5611e-5 (8.00e-6) +	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	5.5192e-7 (2.65e-6) −
	15	48	0.0000e+0 (0.00e+0)	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈
	20	58	0.0000e+0 (0.00e+0)	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈
WFG4	5	28	4.0058e+3 (6.91e+0)	4.6981e+3 (1.80e+1) +	3.4758e+3 (1.51e+2) −	**4.1242e+3 (8.31e+1)** +	3.9922e+3 (8.48e+0) −
	8	34	**1.6917e+7 (5.65e+4)**	1.9374e+7 (1.84e+5) +	6.5549e+6 (8.80e+5) −	1.0100e+7 (1.40e+6) −	1.6843e+7 (7.21e+4) −
	10	38	7.8040e+9 (2.22e+7)	7.6770e+9 (3.02e+8) −	3.0285e+9 (4.91e+8) −	5.7940e+9 (4.73e+8) −	**7.7590e+9 (4.99e+7)** −
	15	48	1.6394e+17 (9.19e+14)	**1.6035e+17 (1.21e+15)** −	2.2562e+16 (8.85e+15) −	8.1631e+16 (2.81e+16) −	1.5473e+17 (1.48e+15) −
	20	58	1.5522e+25 (2.91e+23)	**1.5416e+25 (3.26e+23)** ≈	8.9956e+24 (1.98e+24) −	1.0987e+25 (1.20e+24) −	1.4951e+25 (2.72e+23) −
WFG5	5	28	3.7015e+3 (6.11e+0)	4.4965e+3 (1.29e+1) +	3.4749e+3 (1.48e+2) −	**3.9569e+3 (8.11e+1)** +	3.7072e+3 (7.60e+0) +
	8	34	1.5620e+7 (5.19e+4)	1.8484e+7 (7.66e+4) +	7.1593e+6 (6.89e+5) −	9.6277e+6 (9.83e+5) −	**1.5651e+7 (4.34e+4)** ≈
	10	38	7.1934e+9 (2.36e+7)	8.1664e+9 (2.28e+8) +	2.4829e+9 (2.14e+8) −	4.8907e+9 (4.87e+8) −	**7.2052e+9 (2.71e+7)** ≈
	15	48	1.4525e+17 (4.62e+14)	1.4429e+17 (2.17e+15) ≈	2.8521e+16 (6.35e+15) −	1.0243e+17 (1.22e+16) −	**1.4457e+17 (7.84e+14)** ≈
	20	58	1.4464e+25 (6.16e+22)	1.3774e+25 (1.76e+23) −	7.2635e+24 (1.24e+24) −	8.2017e+24 (2.35e+24) −	**1.4045e+25 (1.54e+23)** −
WFG6	5	28	3.5749e+3 (1.15e+2)	4.4569e+3 (7.48e+1) +	2.6684e+3 (2.38e+2) −	**3.6444e+3 (1.39e+2)** +	3.5874e+3 (9.99e+1) ≈
	8	34	**1.5341e+7 (5.06e+5)**	1.8063e+7 (4.72e+5) +	2.3162e+6 (2.89e+5) −	8.6665e+6 (9.89e+5) −	1.4928e+7 (6.07e+5) −
	10	38	**6.8342e+9 (1.91e+8)**	7.8934e+9 (2.02e+8) +	9.2613e+8 (3.31e+8) −	6.6464e+8 (4.00e+8) −	6.8186e+9 (2.58e+8) ≈
	15	48	1.3419e+17 (1.02e+16)	1.4817e+17 (4.01e+15) +	9.0731e+15 (2.09e+15) −	2.6149e+15 (6.03e+14) −	**1.3571e+17 (5.85e+15)** ≈
	20	58	1.4053e+25 (5.95e+23)	**1.3987e+25 (4.67e+23)** −	9.9388e+24 (9.12e+23) −	2.8717e+23 (6.25e+22) −	1.3619e+25 (5.53e+23) ≈
WFG7	5	28	**3.9858e+3 (9.53e+0)**	4.8225e+3 (8.82e+0) +	3.1898e+3 (2.21e+2) −	4.1663e+3 (4.16e+1) +	3.9854e+3 (9.32e+0) ≈
	8	34	1.6659e+7 (5.30e+4)	1.9781e+7 (9.28e+4) +	5.5036e+6 (1.55e+6) −	9.0292e+6 (2.49e+5) −	**1.6864e+7 (4.28e+4)** +
	10	38	7.7112e+9 (3.00e+7)	8.7758e+9 (2.08e+8) +	1.4770e+9 (6.95e+8) −	4.1598e+9 (1.93e+9) −	**7.7684e+9 (7.23e+7)** +
	15	48	**1.5784e+17 (4.49e+14)**	1.6123e+17 (3.55e+15) +	2.5056e+16 (1.27e+16) −	5.7846e+16 (4.64e+16) −	1.5749e+17 (1.25e+15) ≈
	20	58	**1.5078e+25 (3.64e+23)**	1.5554e+25 (2.09e+23) +	3.4001e+24 (1.53e+24) −	6.4296e+24 (1.07e+24) −	1.3944e+25 (3.61e+23) −
WFG8	5	28	2.7521e+3 (2.14e+1)	4.0357e+3 (2.43e+1) +	1.8859e+3 (8.95e+2) ≈	**3.5402e+3 (1.43e+1)** +	2.7845e+3 (3.91e+1) ≈
	8	34	1.1255e+7 (3.14e+5)	1.6637e+7 (2.79e+5) +	7.6165e+5 (1.29e+6) −	1.1786e+7 (7.96e+5) −	**1.1603e+7 (6.36e+5)** ≈
	10	38	5.3731e+9 (4.30e+8)	7.6851e+9 (1.18e+8) +	8.9075e+8 (2.30e+9) −	3.3513e+9 (1.03e+9) −	**5.4055e+9 (2.73e+8)** ≈
	15	48	**1.2511e+17 (9.49e+15)**	1.5210e+17 (1.31e+15) +	0.0000e+0 (0.00e+0) −	5.6088e+16 (4.75e+16) −	1.1852e+17 (8.70e+15) −
	20	58	**1.3141e+25 (1.13e+24)**	1.4651e+25 (8.57e+22) +	1.4114e+22 (2.82e+22) −	3.4150e+24 (1.86e+24) −	1.2479e+25 (1.01e+24) −
WFG9	5	28	3.9126e+3 (1.63e+1)	4.5151e+3 (1.50e+1) +	2.9949e+3 (3.85e+2) −	**4.0479e+3 (4.40e+1)** +	3.8707e+3 (5.86e+1) −
	8	34	**1.6211e+7 (2.65e+5)**	1.7957e+7 (2.68e+5) +	5.8459e+6 (1.72e+6) −	1.0749e+7 (9.21e+5) −	1.5595e+7 (1.09e+6) −
	10	38	**7.3481e+9 (2.07e+8)**	8.0696e+9 (2.48e+8) +	8.3015e+8 (9.01e+8) −	3.8171e+9 (2.17e+9) −	7.0520e+9 (4.97e+8) −
	15	48	**1.3562e+17 (8.85e+15)**	1.4881e+17 (1.48e+15) +	2.6361e+16 (1.67e+16) −	1.1232e+17 (3.25e+15) −	1.2886e+17 (7.97e+15) −
	20	58	**1.3019e+25 (5.80e+23)**	1.4093e+25 (2.43e+23) +	4.4725e+24 (1.60e+24) −	8.1349e+24 (1.18e+24) −	1.2517e+25 (7.66e+23) −
+/−/≈				25/20/4	0/41/4	6/36/3	3/30/12

"+", "-" and "≈" indicate that the result is significantly better, significantly worse and statistically similar to that obtained by NSGA-II+AD, respectively.

Table 8: HV results (mean and SD) of the five algorithms on the DTLZ test suite. The best and the second mean among the algorithms for each problem instance are highlighted in gray background and bold, respectively.

Problem	M	D	NSGA-II+AD	PICEAg	MaOEARD	VaEA	KnEA
DTLZ1	5	9	**4.5874e-2 (2.34e-4)**	1.6019e-2 (3.99e-3) −	4.7362e-2 (4.78e-4) +	3.3939e-2 (1.84e-2) −	3.4511e-2 (7.40e-3) −
	8	12	8.1246e-3 (1.90e-5)	1.9150e-3 (3.82e-4) −	**7.3150e-3 (4.86e-4)** −	6.6560e-3 (2.12e-3) −	6.8324e-3 (1.88e-3) −
	10	14	2.4945e-3 (5.31e-6)	4.9511e-4 (5.95e-5) −	1.1518e-3 (8.73e-4) −	**1.7113e-3 (8.65e-4)** −	0.0000e+0 (0.00e+0) −
	15	24	1.2705e-4 (7.42e-8)	1.8075e-5 (7.01e-6) −	**1.0258e-4 (1.07e-5)** −	1.0154e-4 (3.77e-5) −	0.0000e+0 (0.00e+0) −
	20	29	6.4082e-6 (3.72e-9)	1.1205e-6 (7.62e-7) −	**4.9711e-6 (4.29e-7)** −	2.8170e-6 (2.15e-6) −	0.0000e+0 (0.00e+0) −
DTLZ2	5	14	1.1529e+0 (1.14e-2)	1.1677e+0 (2.38e-2) +	1.0446e+0 (2.30e-2) −	**1.2575e+0 (5.68e-3)** +	1.2657e+0 (1.15e-2) +
	8	17	1.7675e+0 (1.64e-2)	1.5222e+0 (8.39e-2) −	1.2337e+0 (7.42e-2) −	**1.8721e+0 (2.46e-2)** +	1.8947e+0 (2.23e-2) +
	10	19	**2.2489e+0 (2.09e-2)**	1.8171e+0 (1.10e-1) −	1.1735e+0 (1.43e-1) −	1.2929e+0 (3.83e-1) −	2.3431e+0 (8.29e-2) +
	15	24	**3.8630e+0 (3.09e-2)**	2.7683e+0 (1.73e-1) −	1.5435e+0 (3.44e-1) −	3.3761e+0 (2.06e-1) −	4.1027e+0 (1.16e-2) +
	20	29	6.3982e+0 (4.21e-2)	4.3959e+0 (2.82e-1) −	2.2676e+0 (3.14e-1) −	4.4919e+0 (4.31e-1) −	6.3468e+0 (9.68e-1) −
DTLZ3	5	14	1.1393e+0 (1.61e-2)	3.8450e-1 (1.13e-1) −	**1.1067e+0 (3.13e-2)** ≈	9.7596e-1 (1.47e-1) −	6.6070e-1 (2.39e-1) −
	8	17	1.7496e+0 (3.05e-2)	3.4042e-1 (6.62e-2) −	**8.6162e-1 (6.59e-1)** −	3.2048e-2 (1.52e-1) −	0.0000e+0 (0.00e+0) −
	10	19	2.2231e+0 (3.68e-2)	3.7802e-1 (6.46e-2) −	**7.0236e-1 (6.31e-1)** −	0.0000e+0 (0.00e+0) −	0.0000e+0 (0.00e+0) −
	15	24	3.8375e+0 (6.73e-2)	3.2125e-1 (3.13e-2) −	**1.0456e+0 (1.05e+0)** −	0.0000e+0 (0.00e+0) −	0.0000e+0 (0.00e+0) −
	20	29	6.3332e+0 (1.34e-1)	4.6535e-1 (6.18e-2) −	**6.5421e-1 (1.02e+0)** −	0.0000e+0 (0.00e+0) −	0.0000e+0 (0.00e+0) −
DTLZ4	5	14	1.1774e+0 (9.44e-3)	1.0650e+0 (2.06e-1) ≈	1.0790e+0 (5.25e-2) ≈	**1.2550e+0 (8.07e-3)** +	1.2773e+0 (1.24e-2) +
	8	17	1.8037e+0 (1.28e-2)	1.7904e+0 (5.91e-2) ≈	1.5641e+0 (9.10e-2) −	**1.8404e+0 (2.96e-2)** +	1.9371e+0 (1.86e-2) +
	10	19	**2.2857e+0 (1.91e-2)**	2.1585e+0 (1.71e-1) ≈	1.9399e+0 (7.49e-2) −	2.0004e+0 (1.43e-1) −	2.4627e+0 (1.96e-2) +
	15	24	**3.9393e+0 (2.37e-2)**	3.7050e+0 (8.66e-2) −	3.3898e+0 (1.33e-1) −	3.6023e+0 (1.50e-1) −	4.0854e+0 (1.92e-2) +
	20	29	**6.6009e+0 (1.03e-2)**	5.9393e+0 (1.23e-1) −	5.7553e+0 (2.14e-1) −	5.1063e+0 (3.37e-1) −	6.6437e+0 (1.86e-2) +
DTLZ5	5	14	9.1838e-3 (2.29e-5)	6.4987e-3 (8.63e-5) −	6.4704e-3 (4.09e-8) −	**6.9137e-3 (3.76e-4)** −	3.9186e-3 (2.02e-3) −
	8	17	1.9553e-5 (5.01e-8)	**1.7218e-5 (4.60e-7)** −	1.6815e-5 (5.37e-9) −	1.6259e-5 (8.47e-7) −	6.2703e-6 (5.35e-6) −
	10	19	6.1910e-8 (1.57e-10)	**5.6908e-8 (1.15e-9)** −	5.6188e-8 (2.56e-11) −	5.2471e-8 (4.54e-9) −	1.7660e-8 (1.72e-8) −
	15	24	8.7763e-17 (3.28e-19)	**8.4304e-17 (1.72e-20)** −	8.3949e-17 (5.44e-19) −	8.3122e-17 (1.76e-18) −	1.2456e-17 (2.07e-17) −
	20	29	2.2299e-29 (6.74e-32)	**2.1826e-29 (5.56e-33)** −	2.1818e-29 (2.71e-32) −	1.6586e-29 (3.10e-30) −	2.6653e-30 (6.27e-30) −
DTLZ6	5	14	9.1915e-3 (2.78e-5)	**7.2682e-3 (5.42e-4)** −	6.4705e-3 (0.00e+0) −	5.4077e-3 (2.15e-3) −	4.7887e-3 (2.87e-3) −
	8	17	1.9550e-5 (4.72e-8)	**1.6831e-5 (5.89e-13)** −	1.6821e-5 (0.00e+0) −	1.6779e-6 (5.12e-6) −	2.2198e-6 (5.61e-6) −
	10	19	6.2084e-8 (1.39e-10)	**5.6231e-8 (6.94e-24)** −	5.6218e-8 (0.00e+0) −	3.7559e-9 (1.43e-8) −	2.4005e-9 (1.06e-8) −
	15	24	8.7977e-17 (3.34e-19)	**8.4335e-17 (1.29e-32)** −	8.4321e-17 (1.31e-32) −	2.8109e-17 (4.04e-17) −	2.5103e-18 (1.37e-17) −
	20	29	2.2346e-29 (9.71e-32)	**2.1862e-29 (5.88e-45)** −	2.1834e-29 (2.97e-45) −	2.9152e-30 (7.56e-30) −	0.0000e+0 (0.00e+0) −
DTLZ7	5	24	**2.0167e+0 (3.51e-2)**	1.6028e+0 (3.33e-1) −	1.5284e+0 (5.80e-2) −	1.9821e+0 (3.94e-2) ≈	2.2666e+0 (3.10e-2) +
	8	27	2.3221e+0 (1.61e-2)	1.4317e+0 (6.02e-2) −	2.6157e-1 (1.58e-1) −	1.3323e+0 (1.77e-1) −	**1.5039e+0 (1.93e-1)** −
	10	29	2.4041e+0 (3.75e-2)	**1.4387e+0 (5.79e-2)** −	5.4562e-2 (5.82e-2) −	1.4074e-1 (3.05e-1) −	9.3581e-1 (2.80e-1) −
	15	24	2.2840e+0 (1.72e-1)	**1.4484e+0 (5.74e-2)** −	6.4608e-4 (1.12e-3) −	7.0452e-1 (2.14e-1) −	3.1193e-4 (1.09e-3) −
	20	29	1.7842+0 (1.98e-1)	**1.2521e+0 (2.98e-2)** ≈	4.7367e-5 (1.09e-4) −	1.5099e-2 (3.36e-2) −	0.0000e+0 (0.00e+0) −
+/ − / ≈				1/30/4	1/32/2	4/30/1	10/25/0

"+", "-" and "≈" indicate that the result is significantly better, significantly worse and statistically similar to that obtained by NSGA-II+AD, respectively.

in Table 9, the HV results of WFG test instances are given. The best performing algorithm is KnEA, which gets 25 best results and 5 second-best results, respectively. It was closely followed by the VaEA. Although the number of VaEA's best results is far less than KnEA's, its second-best solutions account for the vast majority. For NSGA-II+AD, what we can observe is that for problems with complicated fronts (such as WFG1 and WFG2), and with higher dimensional objective space (such as 15-objective WFG4 and WFG9, 20-objective WFG4-WFG8, etc.), NSGA-II+AD achieve better performance. For PICEAg, it performs best on the WFG3 with 8-objective and 10-objective but struggles on other problems with high dimension. For MaOEARD, it obtains worse HV values on almost all the test instances. According to statistical test, NSGA-II+AD is significantly superior to PICEAg, MaOEARD, VaEA and KnEA on 24, 39, 15 and 12 test instances while it is inferior to PICEAg, MaOEARD, VaEA and KnEA on 12, 0, 25 and 29 test instances.

5.5. Visualization of Experimental Results

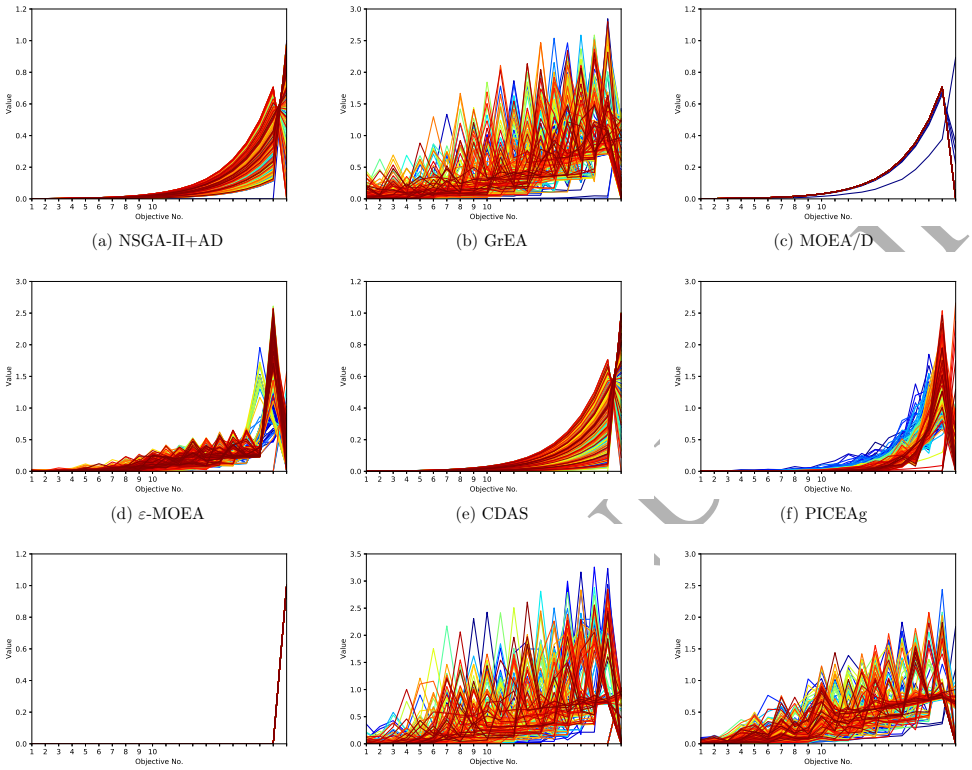
To intuitively illustrate the results in terms of convergence and diversity, we chose 20-objective DTLZ5 and 10-objective WFG1 as examples. As shown in Fig. 8, this figure plots the final solutions of one run with respect to the 20-objective DTLZ5 by the parallel coordinate [34]. This test problem has a degenerate Pareto front. This particular run is associated with the result which is the closest to the mean HV value. It is clear from Fig. 8 that the solutions of GrEA, ε-MOEA, PICEAg, VaEA and KnEA fail to converge into the optimal front. Although MOEA/D and MaOEAD can converge, they struggle to maintain diversity, with their solutions converging into the local area of the Pareto front. NSGA-II+AD and CDAS perform similarly. The only difference is that the solutions obtained by the latter is slightly more distributive than that of the former.

For the 10-objective WFG1, the final solutions obtained by all the algorithms are shown in Fig. 9. It is clear from this figure that solutions of NSGA-II+AD are with highest quality in terms of both the convergence and extensity and the upper and lower bounds of obtained objective i are 0 and $2 \cdot i$, respectively. However, what needs to be pointed out is that some middle regions of the Pareto front have not been covered by NSGA-II+AD. For GrEA, MOEA/D and MaOEARD, their solutions converge to the local area of Pareto front. The solutions obtained by PICEAg are better than the above algorithms in terms of diversity, but fail to reach some regions of the

Table 9: HV results (mean and SD) of the five algorithms on the DTLZ test suite. The best and the second mean among the algorithms for each problem instance are highlighted in gray background and bold, respectively.

Problem	M	D	NSGA-II+AD	PICEAg	MaOEARD	VaEA	KnEA
WFG1	5	28	5.8425e+3 (1.20e+2)	**5.5932e+3 (2.10e+2)** ≈	2.9457e+3 (9.42e+2) −	3.6534e+3 (3.27e+2) −	5.3553e+3 (2.17e+2) −
	8	34	2.0632e+7 (9.97e+3)	**2.0556e+7 (2.04e+5)** ≈	8.7541e+6 (2.84e+6) −	1.0082e+7 (6.91e+5) −	1.7777e+7 (1.07e+6) −
	10	38	8.6472e+9 (3.12e+6)	**8.6147e+9 (1.66e+7)** −	3.0679e+9 (5.13e+8) −	3.8692e+9 (5.29e+8) −	7.4835e+9 (9.82e+8) −
	15	48	1.3899e+17 (3.42e+13)	**1.3780e+17 (3.51e+14)** ≈	3.4310e+16 (5.27e+15) −	5.2126e+16 (8.78e+15) −	9.5304e+16 (3.10e+16) −
	20	58	1.0837e+25 (5.48e+23)	**9.3357e+24 (9.63e+23)** −	2.2579e+24 (4.35e+23) −	2.2954e+24 (1.93e+23) −	5.5775e+24 (1.79e+24) −
WFG2	5	28	**6.0775e+3 (5.10e+0)**	5.8561e+3 (1.18e+2) −	5.8616e+3 (7.37e+1) −	6.0942e+3 (1.27e+1) +	6.1076e+3 (9.39e+0) +
	8	34	2.1991e+7 (2.89e+4)	2.0986e+7 (3.30e+5) −	2.0996e+7 (4.39e+5) −	**2.1871e+7 (7.70e+4)** −	2.1869e+7 (4.82e+4) −
	10	38	9.5767e+9 (1.91e+7)	8.9992e+9 (7.87e+7) −	9.1639e+9 (9.48e+7) −	**9.5260e+9 (3.12e+7)** −	9.5188e+9 (2.85e+7) −
	15	48	1.7696e+17 (2.46e+14)	1.6198e+17 (2.88e+15) −	1.6755e+17 (2.71e+15) −	**1.7685e+17 (5.82e+14)** −	1.7321e+17 (4.33e+15) −
	20	58	**1.6937e+25 (7.42e+22)**	1.4574e+25 (1.96e+24) −	1.5722e+25 (1.39e+23) −	1.6994e+25 (8.08e+22) +	1.5911e+25 (5.28e+23) −
WFG3	5	28	2.8919e+0 (7.87e-2)	**2.1822e+0 (2.36e-1)** −	4.1796e-1 (5.60e-1) −	1.2350e+0 (1.88e-1) −	6.6087e-1 (2.44e-1) −
	8	34	**1.2908e-2 (3.55e-3)**	2.0237e-2 (1.40e-3) +	0.0000e+0 (0.00e+0) −	5.5873e-3 (3.21e-3) −	0.0000e+0 (0.00e+0) −
	10	38	**1.0244e-5 (1.23e-5)**	5.4966e-5 (4.30e-6) +	0.0000e+0 (0.00e+0) −	3.4609e-6 (5.45e-6) ≈	0.0000e+0 (0.00e+0) −
	15	48	0.0000e+0 (0.00e+0)	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈
	20	58	0.0000e+0 (0.00e+0)	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈	0.0000e+0 (0.00e+0) ≈
WFG4	5	28	4.0058e+3 (6.91e+0)	**4.6777e+3 (5.79e+1)** +	3.6140e+3 (2.85e+2) ≈	4.5928e+3 (2.96e+1) +	4.7239e+3 (2.01e+1) +
	8	34	1.6917e+7 (5.65e+4)	1.5379e+7 (1.59e+6) ≈	1.1883e+7 (1.77e+6) −	**1.8841e+7 (2.03e+5)** +	1.9570e+7 (4.31e+1) +
	10	38	7.8040e+9 (2.22e+7)	6.3983e+9 (9.73e+8) −	5.4068e+9 (6.07e+8) −	**8.4124e+9 (7.01e+7)** +	9.0679e+9 (2.10e+7) +
	15	48	**1.6394e+17 (9.19e+14)**	1.0426e+17 (1.76e+16) −	9.7469e+16 (8.64e+15) −	1.5849e+17 (1.90e+15) −	1.7427e+17 (5.68e+14) +
	20	58	**1.5522e+25 (2.91e+23)**	2.8922e+24 (1.48e+24) −	8.7864e+24 (7.92e+23) −	1.4978e+25 (2.10e+23) −	1.5953e+25 (1.18e+24) +
WFG5	5	28	3.7015e+3 (6.11e+0)	**4.3438e+3 (7.25e+1)** +	3.2694e+3 (9.48e+1) −	4.4000e+3 (3.11e+1) +	4.4746e+3 (2.37e+1) +
	8	34	1.5620e+7 (5.19e+4)	1.3892e+7 (1.45e+6) −	1.1198e+7 (6.74e+5) −	**1.7980e+7 (1.28e+5)** +	1.8264e+7 (2.14e+5) +
	10	38	7.1934e+9 (2.36e+7)	5.3548e+9 (4.91e+8) −	4.2400e+9 (5.70e+8) −	**7.9657e+9 (5.48e+7)** +	8.5046e+9 (1.84e+7) +
	15	48	1.4525e+17 (4.62e+14)	8.0010e+16 (1.40e+15) −	7.3247e+16 (6.61e+15) −	**1.4826e+17 (1.42e+15)** +	1.6209e+17 (5.19e+14) +
	20	58	**1.4464e+25 (6.16e+22)**	1.3207e+24 (5.01e+23) −	7.2920e+24 (1.51e+24) −	1.3868e+25 (1.53e+23) −	1.5656e+25 (5.44e+22) +
WFG6	5	28	3.5749e+3 (1.15e+2)	4.1717e+3 (9.65e+1) +	3.1614e+3 (1.42e+2) −	**4.2615e+3 (1.27e+2)** +	4.3131e+3 (9.29e+1) +
	8	34	1.5341e+7 (5.06e+5)	1.5467e+7 (3.95e+5) +	1.0588e+7 (2.09e+6) −	1.7666e+7 (3.05e+5) +	**1.7313e+7 (5.42e+5)** +
	10	38	6.8342e+9 (1.91e+8)	6.4936e+9 (4.02e+8) ≈	3.8789e+9 (8.58e+8) −	**7.8409e+9 (1.90e+8)** +	8.2036e+9 (2.90e+8) +
	15	48	1.3419e+17 (1.02e+16)	9.8257e+16 (1.25e+16) −	7.5573e+16 (1.35e+16) −	**1.5069e+17 (6.68e+15)** +	1.5742e+17 (5.62e+15) +
	20	58	**1.4053e+25 (5.95e+23)**	1.1473e+24 (5.75e+23) −	5.9903e+24 (1.03e+24) −	1.3898e+25 (4.71e+23) ≈	1.4872e+25 (3.97e+23) +
WFG7	5	28	3.9858e+3 (9.53e+0)	4.5915e+3 (8.07e+1) +	3.0065e+3 (1.93e+2) −	**4.6946e+3 (3.23e+1)** +	4.7836e+3 (2.67e+1) +
	8	34	1.6659e+7 (5.30e+4)	1.8293e+7 (9.37e+5) +	8.7424e+6 (1.72e+6) −	1.9519e+7 (1.92e+5) +	**1.9401e+7 (1.76e+5)** +
	10	38	7.7112e+9 (3.00e+7)	7.3057e+9 (6.83e+8) ≈	2.6565e+9 (1.22e+9) −	**8.7427e+9 (7.63e+7)** +	9.0369e+9 (9.53e+7) +
	15	48	1.5784e+17 (4.49e+14)	1.2212e+17 (1.60e+16) −	4.5887e+16 (1.48e+16) −	**1.6464e+17 (3.06e+15)** +	1.7526e+17 (7.36e+14) +
	20	58	1.5078e+25 (3.64e+23)	1.1376e+25 (1.36e+24) −	4.2140e+24 (1.74e+24) −	1.5866e+25 (1.99e+23) +	**1.5743e+25 (8.52e+23)** +
WFG8	5	28	2.7521e+3 (2.14e+1)	3.7763e+3 (9.54e+1) +	2.7978e+3 (2.50e+2) ≈	**3.8525e+3 (7.94e+1)** +	3.9084e+3 (7.31e+1) +
	8	34	1.1255e+7 (3.14e+5)	1.5377e+7 (7.81e+5) +	1.0218e+7 (2.75e+6) ≈	**1.5410e+7 (2.48e+5)** +	1.5959e+7 (4.45e+5) +
	10	38	5.3731e+9 (4.30e+8)	6.8649e+9 (3.97e+8) +	4.6810e+9 (9.20e+8) ≈	7.2276e+9 (1.09e+8) +	**7.1055e+9 (6.84e+8)** +
	15	48	1.2511e+17 (9.49e+15)	1.1882e+17 (1.53e+16) −	8.1001e+16 (1.14e+16) −	1.4664e+17 (4.41e+15) +	**1.3008e+17 (2.29e+16)** ≈
	20	58	**1.3141e+25 (1.13e+24)**	1.1328e+25 (1.73e+24) −	7.0935e+24 (1.22e+24) −	1.5167e+25 (1.84e+23) +	1.1869e+25 (2.86e+24) ≈
WFG9	5	28	3.9126e+3 (1.63e+1)	4.4570e+3 (6.68e+1) +	2.9265e+3 (1.58e+2) −	**4.2752e+3 (1.20e+2)** +	4.5746e+3 (2.80e+1) +
	8	34	1.6211e+7 (2.65e+5)	1.5357e+7 (1.72e+6) ≈	8.0088e+6 (8.92e+5) −	**1.6527e+7 (1.20e+6)** +	1.8495e+7 (8.91e+5) +
	10	38	**7.3481e+9 (2.07e+8)**	5.7145e+9 (6.97e+8) −	3.4804e+9 (6.69e+8) −	6.9523e+9 (5.79e+8) −	8.1776e+9 (6.50e+8) +
	15	48	**1.3562e+17 (8.85e+15)**	8.0756e+16 (1.63e+16) −	5.9948e+16 (1.15e+16) −	1.2948e+17 (1.13e+16) −	1.5123e+17 (1.32e+16) +
	20	58	1.3019e+25 (5.80e+23)	8.2074e+24 (9.28e+23) −	5.5214e+24 (1.42e+24) −	**1.3294e+25 (5.74e+23)** +	1.3761e+25 (1.11e+24) +
+/ − / ≈				12/24/9	0/39/6	25/16/4	29/12/4

"+", "-" and "≈" indicate that the result is significantly better, significantly worse and statistically similar to that obtained by NSGA-II+AD, respectively.



(a) NSGA-II+AD (b) GrEA (c) MOEA/D

(d) ε-MOEA (e) CDAS (f) PICEAg

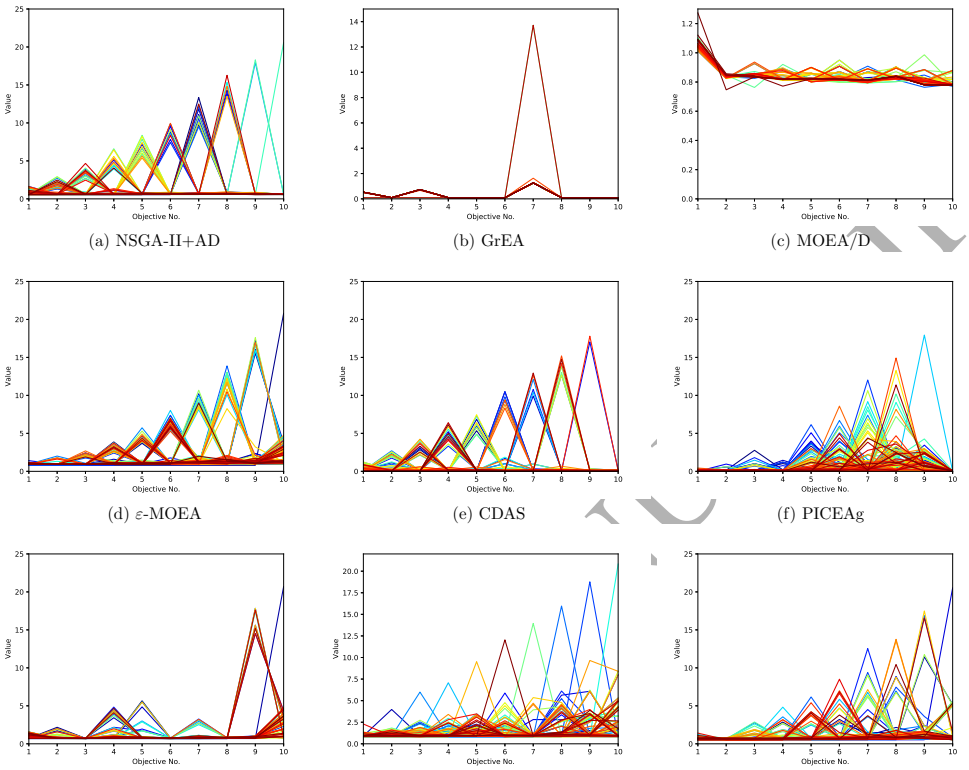
(g) MaOEARD (h) VaEA (j) KnEA

Figure 8: The final solution set of the nine algorithms on the 20-objective DTLZ5, shown by parallel coordinates. Where the abscissa represents the objective dimension and the ordinate indicates the objective value.

Pareto front. For VaEA and KnEA, the solutions obtained by the former is more extensive than the latter, but the latter is more uniform than the former. The distribution of solutions obtained by CDAS is similar to that of NSGA-II+AD, while our algorithm is better than CDAS in extensity.

In addition, this section extends the study on the selection pressure of angle dominance in the objective space. Here we consider eight instances based on the Pareto front of the problems [3] to demonstrate the average

[3]In the test instances selected in this paper, DTLZ2,DTLZ3 and DTLZ4 have the same Pareto front, DTLZ5 and DTLZ6 have the same Pareto front, and the Pareto front of WFG4-WFG9 are consistent.



(a) NSGA-II+AD (b) GrEA (c) MOEA/D

(d) ε-MOEA (e) CDAS (f) PICEAg

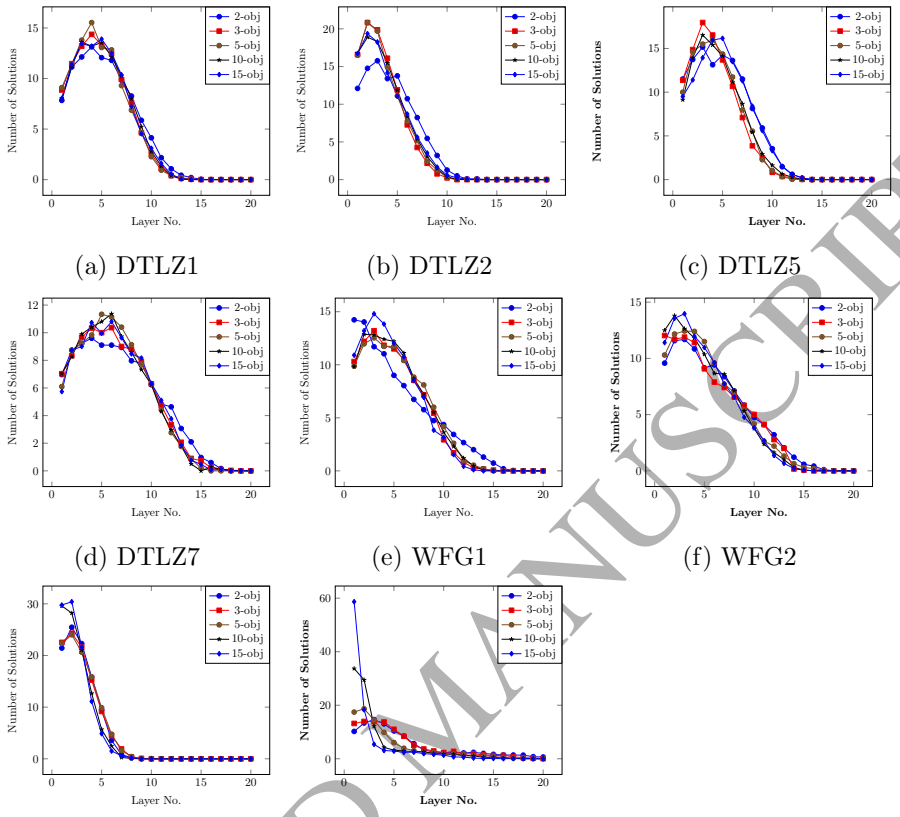
(g) MaOEARD (h) VaEA (j) KnEA

Figure 9: The final solution set of the six algorithms on the ten-objective WFG1, shown by parallel coordinates. Where the abscissa represents the objective dimension and the ordinate indicates the objective value.

number of solutions for all the nondominated layers. As shown in Fig. 10, the distribution curve of nondominated solutions on all problems but WFG4 is consistent with that in Fig. 3(a). The phenomenon shown in the figure is highly correlated with the results of HV. They all showed that NSGA-AD is better than other algorithms in DTLZ1-7 and WFG1-WFG3, but its performance is slightly inferior to the compared algorithms when facing the same Pareto front as WFG4.

5.6. Discussions

The impressive results of NSGA-II+AD motivate us to deeply explore the shortcomings of NSGA-II.



(a) DTLZ1 (b) DTLZ2 (c) DTLZ5

(d) DTLZ7 (e) WFG1 (f) WFG2

(g) WFG3 (h) WFG4

Figure 10: The average number of solutions in all the nondominated layers under the angle nondominated sorting, where the population size is 100, the number of runs is 30.

It is well known that as the number of objective increases, an increasingly larger fraction of a population becomes non-dominated. Since NSGA-II distinguishes between solutions mainly by the Pareto dominance criterion, in many-objective optimization there are not much room for identifying promising solutions. This slows down the search process and therefore NSGA-II could become ineffective.

In addition, the crowding distance of NSGA-II does not provide good diversity of solutions in many-objective optimization, as the considered crowding distance fails to accurately evaluate the crowding degree of solutions when the number of the objectives is greater than two. For the case $M = 2$, crowding is simply measured as the Manhattan distance between neighboring

solutions p_{i-1} and p_{i+1}. For $M \geq 3$, however, the analogy does not hold, because objectives are processed independently and nearest neighbor solutions may change for each objective.

In this paper, our main work is to alleviate the first difficulty mentioned above by using angle dominance that appropriately expands the solution's dominance area. For the second difficulty mentioned above, the improvement measures can be found in [15].

6. Conclusions

Many-objective optimization brings enormous challenges to the EMO community because the Pareto dominance criterion is ineffective in a high-dimensional space. In this paper, we have proposed an angle dominance approach to deal with MaOPs. Converting many objectives of a given problem into angle vectors, angle dominance can achieve a good balance between convergence and extensity.

Systematic experiments have been carried out based on the two test suites (e.g., DTLZ and WFG). From the comparative results, it has been shown that after the implementation of angle dominance, NSGA-II achieves an improvement of performance with varying dimensions. Moreover, five state-of-the-art MOEAs (i.e., GrEA, PICEAg, MaOEARD, MOEAD and ε-MOEA) have been used to compare with the angle dominance based NSGA-II (denoted NSGA-II+AD). The experimental results show that NSGA-II+AD is very competitive against the peer algorithms in terms of providing a good balance between convergence and diversity.

As a new dominance approach in EMO, angle dominance takes account of convergence and extensity. Angle dominance divides the population into many smaller sub-populations by comparing angle vectors, where the convergence and extensity of solutions are kept in the first layers.

Despite the high competitiveness of angle dominance shown in our first attempt, more work is needed to further investigate its benefits and limitations. Although angle dominance is not sensitive to the parameter, an adaptive technology on the parameter would be better when addressing varying problems. Taking uniformity into consideration will also be a focus of our subsequent study.

Acknowledgments

This work was supported in part by the National Key R&D Program of China under Grant 2016YFB0200200, the National Natural Science Foundation of China under Grant 61572177, the National Outstanding Youth Science Program of National Natural Science Foundation of China under Grant 61625202, the Key Program of Natural Science Foundation of China under Grant 61432005 and the International (Regional) Cooperation and Exchange Program of National Natural Science Foundation of China under Grant 61661146006.

References

[1] Adra, S. F. and Fleming, P. J. (2009). A diversity management operator for evolutionary many-objective optimisation. In Ehrgott, Matthias, F. C. M. G. X. H. J.-K. and Sevaux, M., editors, *Evolutionary Multi-Criterion Optimization*, pages 81–94, Berlin, Heidelberg. Springer Berlin Heidelberg.

[2] Agrawal, R. B., Deb, K., and Agrawal, R. (1995). Simulated binary crossover for continuous search space. *Complex systems*, 9(2):115–148.

[3] Bader, J. and Zitzler, E. (2011). Hype: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary computation*, 19(1):45–76.

[4] Batista, Lucas S., C. F. G. F. G. and Ramírez, J. A. (2011). Pareto cone ϵ-dominance: Improving convergence and diversity in multiobjective evolutionary algorithms. In Takahashi, Ricardo H. C., D. K. W. E. F. and Greco, S., editors, *Evolutionary Multi-Criterion Optimization*, pages 76–90, Berlin, Heidelberg. Springer Berlin Heidelberg.

[5] Bentley, P. J., W. J. P. (1998). Finding acceptable solutions in the pareto-optimal range using multiobjective genetic algorithms. In Chawdhry, P. K., R. R. and Pant, R. K., editors, *Soft Computing in Engineering Design and Manufacturing*, pages 231–240, London. Springer London.

[6] Chen, H., Cheng, R., Wen, J., Li, H., and Weng, J. (2018a). Solving large-scale many-objective optimization problems by covariance matrix adaptation evolution strategy with scalable small subpopulations. *Information Sciences*.

[7] Chen, H., Zhu, X., Liu, G., and Pedrycz, W. (2018b). Uncertainty-aware online scheduling for real-time workflows in cloud service environment. *IEEE Transactions on Services Computing*.

[8] Deb, K. and Jain, H. (2014). 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*, 18(4):577–601.

[9] Deb, K., Mohan, M., and Mishra, S. (2005). Evaluating the -domination based multi-objective evolutionary algorithm for a quick computation of pareto-optimal solutions. *Evolutionary Computation*, 13(4):501–525.

[10] Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Trans. Evol. Comp*, 6(2):182–197.

[11] Deb, Kalyanmoy, T. L. L. M. and Zitzler, E. (2005). *Scalable Test Problems for Evolutionary Multiobjective Optimization*, pages 105–145. Springer London, London.

[12] Denysiuk, R. and Gaspar-Cunha, A. (2017). Multiobjective evolutionary algorithm based on vector angle neighborhood. *Swarm and Evolutionary Computation*, 37:45–57.

[13] Drechsler, N., Drechsler, R., and Becker, B. (2001). Multi-objective optimisation based on relation favour. In *Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization*, EMO '01, pages 154–166, London, UK, UK. Springer–Verlag.

[14] Farina, M. and Amato, P. (2002). On the optimal solution definition for many-criteria optimization problems. In *2002 Annual Meeting of the North American Fuzzy Information Processing Society Proceedings. NAFIPS-FLINT 2002 (Cat. No. 02TH8622)*, pages 233–238.

[15] Fortin, F.-A. and Parizeau, M. (2013). Revisiting the nsga-ii crowding-distance computation. In *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation*, GECCO '13, pages 623–630, New York, NY, USA. ACM.

[16] Guohua Wu, Xin Shen, H. L. H. C. A. L. and Suganthan, P. (2018). Ensemble of differential evolution variants. *Information Sciences*, 423:172–186.

[17] Hadka, D., Reed, P. M., and Simpson, T. W. (2012). Diagnostic assessment of the borg moea for many-objective product family design problems. In *2012 IEEE Congress on Evolutionary Computation*, pages 1–10.

[18] He, Z. and Yen, G. G. (2016). Many-objective evolutionary algorithm: Objective space reduction and diversity improvement. *IEEE Transactions on Evolutionary Computation*, 20(1):145–160.

[19] He, Z. and Yen, G. G. (2017). Many-objective evolutionary algorithms based on coordinated selection strategy. *IEEE Transactions on Evolutionary Computation*, 21(2):220–233.

[20] Hernandez-D'iaz, A. G., Santana-Quintero, L. V., Coello Coello, C. A., and Molina, J. (2007). Pareto-adaptive ε-dominance. *Evolutionary computation*, 15(4):493–517.

[21] Huband, S., Hingston, P., Barone, L., and While, L. (2006). A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation*, 10(5):477–506.

[22] Hughes, E. J. (2003). Multiple single objective pareto sampling. In *The 2003 Congress on Evolutionary Computation, 2003. CEC '03.*, volume 4, pages 2678–2684 Vol.4.

[23] Ikeda, K., Kita, H., and Kobayashi, S. (2001). Failure of pareto-based moeas: does non-dominated really mean near to optimal? In *Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546)*, volume 2, pages 957–962 vol. 2.

[24] Ishibuchi, H., Tsukamoto, N., and Nojima, Y. (2008). Evolutionary many-objective optimization. In *2008 3rd International Workshop on Genetic and Evolving Systems*, pages 47–52.

[25] Ishibuchi, Hisao, H. Y. T. N. and Nojima, Y. (2010). Many-objective test problems to visually examine the behavior of multiobjective evolution in a decision space. In Schaefer, Robert, C. C. K. J. and Rudolph, G.,

editors, *Parallel Problem Solving from Nature, PPSN XI*, pages 91–100, Berlin, Heidelberg. Springer Berlin Heidelberg.

[26] Jiang, S. and Yang, S. (2017). A strength pareto evolutionary algorithm based on reference direction for multiobjective and many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 21(3):329–346.

[27] Knowles, J. and Corne, D. (2003). Properties of an adaptive archiving algorithm for storing nondominated vectors. *IEEE Transactions on Evolutionary Computation*, 7(2):100–116.

[28] Kollat, J. B. and Reed, P. M. (2005). Comparison of multi-objective evolutionary algorithms for long-term monitoring design. In *Impacts of Global Climate Change*, pages 1–11.

[29] Lee, S., S. S. K. K. P. H. and Jeon, M. (2008). Statistical properties analysis of real world tournament selection in genetic algorithms. *Applied Intelligence*, 28(2):195–205.

[30] Lei, H., Wang, R., Zhang, T., Liu, Y., and Zha, Y. (2016). A multi-objective co-evolutionary algorithm for energy-efficient scheduling on a green data center. *Comput. Oper. Res.*, 75(C):103–117.

[31] Li, K., Wang, R., Zhang, T., and Ishibuchi, H. (2018a). Evolutionary many-objective optimization: A comparative study of the state-of-the-art. *IEEE Access*, 6:26194–26214.

[32] Li, M., Grosan, C., Yang, S., Liu, X., and Yao, X. (2018b). Multiline distance minimization: A visualized many-objective test problem suite. *IEEE Transactions on Evolutionary Computation*, 22(1):61–78.

[33] Li, M., Yang, S., and Liu, X. (2014). Shift-based density estimation for pareto-based algorithms in many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 18(3):348–365.

[34] Li, M., Zhen, L., and Yao, X. (2017). How to read many-objective solution sets in parallel coordinates [educational forum]. *IEEE Computational Intelligence Magazine*, 12(4):88–100.

[35] Peng Wang, Bo Liao, W. Z. L. C. S. R. M. C. Z. L. and Li, K. (2018). Adaptive region adjustment to improve the balance of convergence and diversity in moea/d. *Applied Soft Computing*, 70:797–813.

[36] P.M. Reed, D. Hadka, J. H. J. K. and Kollat, J. (2013). Evolutionary multiobjective optimization in water resources: The past, present, and future. *Advances in Water Resources*, 51:438–456. 35th Year Anniversary Issue.

[37] Salazar, J. Z., Reed, P. M., Herman, J. D., Giuliani, M., and Castelletti, A. (2016). A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. *Advances in Water Resources*, 92:172–185.

[38] Sato, Hiroyuki, A. H. E. and Tanaka, K. (2007). Controlling dominance area of solutions and its impact on the performance of moeas. In Obayashi, Shigeru, D. K. P. C. H. T. and Murata, T., editors, *Evolutionary Multi-Criterion Optimization*, pages 5–20, Berlin, Heidelberg. Springer Berlin Heidelberg.

[39] Wang, R., Purshouse, R. C., and Fleming, P. J. (2013). Preference-inspired coevolutionary algorithms for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 17(4):474–494.

[40] Wang, R., Zhou, Z., Ishibuchi, H., Liao, T., and Zhang, T. (2018). Localized weighted sum method for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 22(1):3–18.

[41] While, L., Hingston, P., Barone, L., and Huband, S. (2006). A faster algorithm for calculating hypervolume. *IEEE Transactions on Evolutionary Computation*, 10(1):29–38.

[42] Xiang, Y., Zhou, Y., Li, M., and Chen, Z. (2017). A vector angle-based evolutionary algorithm for unconstrained many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 21(1):131–152.

[43] Xu, Y., Li, K., Hu, J., and Li, K. (2014). A genetic algorithm for task scheduling on heterogeneous computing systems using multiple priority queues. *Information Sciences*, 270:255–287.

[44] Yang, S., Li, M., Liu, X., and Zheng, J. (2013). A grid-based evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 17(5):721–736.

[45] Zhang, Q. and Li, H. (2007). Moea/d: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731.

[46] Zhang, Q., Zhu, W., Liao, B., Chen, X., and Cai, L. (2018). A modified pbi approach for multi-objective optimization with complex pareto fronts. *Swarm and Evolutionary Computation*, 40:216–237.

[47] Zhang, X., Tian, Y., and Jin, Y. (2015). A knee point-driven evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 19(6):761–776.

[48] Zhun Fan, Yi Fang, W. L. X. C. C. W. and Goodman, E. (2019). Moea/d with angle-based constrained dominance principle for constrained multi-objective optimization problems. *Applied Soft Computing*, 74:621–633.

[49] Zitzler, E., Laumanns, M., and Thiele, L. (2001). Spea2: Improving the strength pareto evolutionary algorithm. *in Evolutionary Methods for Design, Optimisation and Control*, pages 95–100.

[50] Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., and da Fonseca, V. G. (2003). Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Transactions on Evolutionary Computation*, 7(2):117–132.

Appendix A. Explain Why ε-Dominance May Eliminate Several Viable Solutions

This appendix provides a detailed description why ε-dominance may eliminate several viable solutions.

In 2002, Laumanns *et al.* proposed a loosening dominance for MOEAs, called ε-dominance. Subsequently, many ε-dominance-based algorithms have been proposed, such as ε-NSGA-II [28] and ε-MOEA [9], etc. This criterion serves as an archiving strategy to ensure both properties of convergence towards the Pareto-optimal front and properties of diversity among the solutions found. Since this criterion guarantees that no two archived solutions are

within an ε_i value from each other in the i-th objective, the ε value is usually provided by the decision maker to control the size of the solution set. Nevertheless, due to the geometrical characteristics of the Pareto-optimal front are commonly unknown by the decision maker, the ε-dominance criterion may lose a large number of viable solutions when the ε value is imperfectly estimated.

Another limitation of ϵ-dominance is the fact that it may lose solutions located on segments of the Pareto front that are almost horizontal or almost vertical, as well as the extreme points of the Pareto front. This has a negative impact on the spread of solutions along the Pareto front.

To illustrate the above phenomenon more directly, we referenced the example in [20].

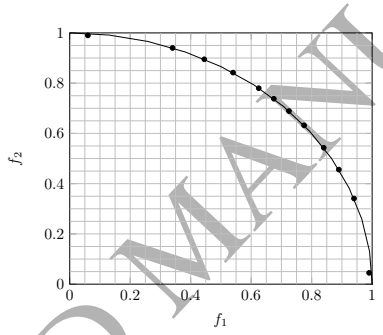
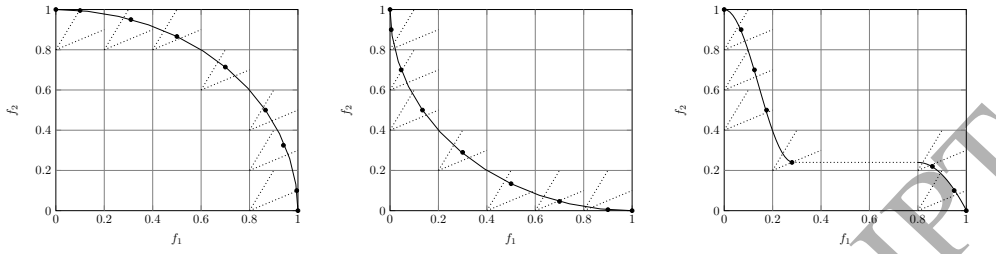


Figure A.11: Illustration of the limitations of ε-dominance, where the objective space is divided into 400 grids (maximum capacity of 20 points) and the objective function is $x^2 + y^2 = 1$ in the figure. This grid allows a maximum of 12 points; the other 8 points are lost because either the extreme points are easily ϵ-dominated.)

As shown in Figure A.11, the objective space are divided into 400 grids, which means that the objective space can accommodate up to 20 points (here the example in [20], However being borrowed). The ϵ-dominance only gets up to 12 points in the space and other 8 points are lost because either the extreme points are easily ϵ-dominated or the precision of the grid is insufficient.

Appendix B. Explain Why Cone ε-Dominance Can Promote Both Convergence And Uniformity

In this appendix, we will analyze in detail why cone ε-dominance can promote both convergence and uniformity.



(a) concave front (b) convex front (c) Disconnected front

Figure B.12: Illustration of the characteristic of cone ε-dominance, where the objective space is divided into 25 grids and the optimal Pareto fronts given are concave, convex and disconnected, respectively. (a) In the concave front, the size of obtained points equals to the maximum non cone ε-dominated points. (b) In the convex front, It can also get the largest non cone ε-dominated points. (c) In the disconnected front, as the same as (a) and (b), each grid touched by the front has one point.

The cone ε-dominance criterion has been proposed by Batista *et al.* [4]. Similar to the ε-dominance, the cone ε-dominance criterion guarantees that no two achieved solutions are within an ε_i value from each other in the i-th objective. Besides, if two points share the same grid, the point is only replaced by a dominating one or by another point closest to the origin of the grid. These characteristics show that the cone ε-dominance can maintain a good distribution and convergence of the population. Besides, the cone ε-dominance introduces a parameter k to control the dominance area so that the loss of extreme points of the Pareto front and points located in segments of the Pareto front that are almost horizontal or vertical can be avoided.

In general, if any connected monotonic front exists between the extreme grids of the hypergrid, then the number of grids that are touched by this front is maximum. Figure B.12(a),(b) illustrate two possible situations in which the number of estimated cone ε-Pareto solutions is maximum, i.e., convex Pareto front and concave Pareto front. For both cases, the number of estimated points is nine (the maximum of non cone ε-dominated points is nine). Figure B.12(c) presents a possible situation in which a disconnected front has been stated. For this case, the maximum size of non cone ε-dominated points cannot be reached. However, it is likely to estimate one solution from each gird touched by the front. The ε-dominance criterion, on the other hand, can only achieve the upper bound for the number of points allowed by a grid when the real Pareto front is linear [20].