Handling Imbalance Between Convergence and Diversity in the Decision Space in Evolutionary Multi-Modal Multi-Objective Optimization

Yiping Liu, *Member, IEEE*, Hisao Ishibuchi, *Fellow, IEEE*, Gary G. Yen, *Fellow, IEEE*, Yusuke Nojima, *Member, IEEE*, and Naoki Masuyama, *Member, IEEE*

***Abstract*—There may exist more than one Pareto optimal solution with the same objective vector to a multi-modal multi-objective optimization problem (MMOP). The difficulties in finding such solutions can be different. Although a number of evolutionary multi-modal multi-objective algorithms (EMMAs) have been proposed, they are unable to solve such an MMOP due to their convergence-first selection criteria. They quickly converge to the Pareto optimal solutions which are easy to find and therefore lose diversity in the decision space. That is, such an MMOP features an imbalance between achieving convergence and preserving diversity in the decision space. In this paper, we first present a set of imbalanced distance minimization benchmark problems. Then we propose an evolutionary algorithm using a convergence-penalized density method (CPDEA). In CPDEA, the distances among solutions in the decision space are transformed based on their local convergence quality. Their density values are estimated based on the transformed distances and used as the selection criterion. We compare CPDEA with five state-of-the-art EMMAs on the proposed benchmarks. Our experimental results show that CPDEA is clearly superior in solving these problems.**

***Index Terms*—Evolutionary multi-modal multi-objective optimization, convergence, decision space diversity, density estimation, test problems.**

I. Introduction

MULTI-objective optimization problems (MOPs) are commonly seen in a variety of disciplines, such as job shop scheduling [1] and financial portfolio management [2]. They involve multiple objectives to be optimized simultaneously. In this study, we consider an MOP with box constraints, which can be formulated as follows:

$$\begin{aligned} &\min \boldsymbol{f}(\boldsymbol{x}) = \min(f_1(\boldsymbol{x}), \dots, f_M(\boldsymbol{x})), \\ &\text{s.t. } \boldsymbol{x} \in S, \end{aligned} \tag{1}$$

where $\boldsymbol{x}$ is an n-dimensional decision vector in the search space S, $f_m(\boldsymbol{x})$ is the m-th objective to be minimized ($m = 1, ..., M$), and M is the number of objectives. $S = \{\boldsymbol{x} \in \boldsymbol{R}^n : x_i^{\text{lower}} \leq x_i \leq x_i^{\text{upper}}, i = 1, \dots, n\}$ where x_i^{upper} and x_i^{lower} are the lower and upper bounds of x_i, respectively. Due to the conflicting nature of objectives, usually an MOP has no single optimal solution which simultaneously optimizes all objectives. It has a set of trade-off solutions, known as the Pareto optimal solution set (PS). The image of the PS in the objective space is called the Pareto optimal front (PF).

For an MOP, if there exist at least two different Pareto optimal solutions with the same objective vector, i.e., corresponding to the same point on the PF, it is also called multi-modal multi-objective optimization problem (MMOP) [3]. Pareto optimal solutions which have the same objective vector are referred to as equivalent Pareto optimal solutions.

Over the past two decades, a number of evolutionary multi-objective algorithms (EMOAs) have been proposed and proven successful in solving MOPs, such as elitist non-dominated sorting genetic algorithm II (NSGA-II) [4] and multi-objective evolutionary algorithm based on decomposition (MOEA/D) [5]. The general goal of EMOAs is to search for a set of solutions with good convergence and diversity in the objective space. The balance between convergence and diversity in the objective space is one of the most important issues in evolutionary multi-objective optimization (EMO) [6]–[9]. In environmental and mating selection, most EMOAs prefer solutions with good convergence and then solutions with good diversity in the objective space (e.g., the non-dominated sorting and the crowding distance in NSGA-II). That is, they use convergence-first selection criteria. It is worth mentioning that there are a few EMOAs which prefer diversity in the objective space first, e.g., MOEA/D-M2M (multi-objective to multi-objective) [10]. The M2M method decomposes the objective space into a number of sub-regions. In each sub-region, at least one solution is preserved regardless of its convergence quality.

However, these EMOAs cannot solve MMOPs due to the lack of diversity maintenance in the decision space, i.e., they are unable to find equivalent Pareto optimal solutions.

This work was supported in part by the National Natural Science Foundation of China under Grant 61876075 and Grant 61803192, in part by the Program for Guangdong Introducing Innovative and Enterpreneurial Teams under Grant 2017ZT07X386, in part by the Shenzhen Peacock Plan under Grant KQTD2016112514355531, in part by the Science and Technology Innovation Committee Foundation of Shenzhen under Grant ZDSYS201703031748284, in part by the Program for University Key Laboratory of Guangdong Province under Grant 2017KSYS008, and in part by Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP19K20358. (*Corresponding Author: Hisao Ishibuchi*)

Y. Liu, Y. Nojima, and N. Masuyama are with Department of Computer Science and Intelligent Systems, Graduate School of Engineering, Osaka Prefecture University, Sakai, Osaka 599-8531, Japan. (yiping0liu@gmail.com, nojima@cs.osakafu-u.ac.jp, masuyama@cs.osakafu-u.ac.jp)

H. Ishibuchi is with Shenzhen Key Laboratory of Computational Intelligence, University Key Laboratory of Evolving Intelligent Systems of Guangdong Province, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China. (hisao@sustech.edu.cn)

G. G. Yen is with the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078, USA. (gyen@okstate.edu)

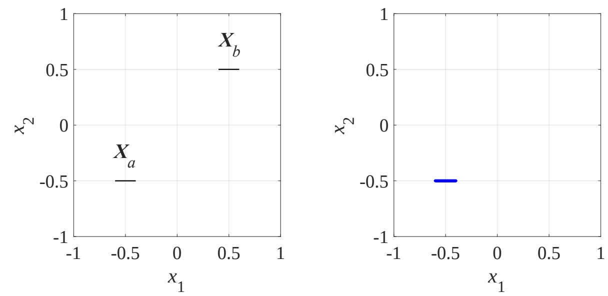
Recently, evolutionary multi-modal multi-objective algorithms (EMMAs) have been proposed to address this issue, such as Omni-optimizer [11], decision space based niching NSGA-II (DN-NSGA-II) [12], double-niched evolutionary algorithm (DNEA) [13], and MOEA/D with addition and deletion operators (MOEA/D-AD) [14]. These algorithms are based on existing EMOAs and niching methods [15] to promote diversity in the decision space.

In evolutionary multi-modal multi-objective optimization (EMMO), in addition to the balance between convergence and diversity in the objective space, we should consider two more balances. One is the balance between diversity in the objective space and that in the decision space. The other is the balance between convergence and diversity in the decision space.

In [16], MOPs whose PFs offer different degrees of search difficulty in different regions are proposed and discussed. The convergence-first EMOAs can obtain only those solutions that are close to easy regions. That is, they fail to maintain diversity in the objective space while converging. This phenomenon is termed as the imbalance between convergence and diversity in the objective space. This issue can be well solved by the diversity-first EMOAs, such as MOEA/D-M2M. Intuitively, such an imbalance may also exist in an MMOP. It could be handled by the same manner proposed in [16]. In [3], [13], MMOPs with the imbalance (or inconsistency) between diversity in the objective space and that in the decision space are proposed and discussed, where a uniformly and widely distributed PS in the objective space is not uniformly or widely distributed in the decision space, or vice versa. Diversity maintenance both in the objective and decision spaces is suggested to handle this imbalance.

To the best of our knowledge, there exists no study on the imbalance between convergence and diversity in the decision space in EMMO. In an MMOP, one equivalent Pareto optimal solution may be more difficult to find than another. Under such a condition, due to the convergence-first selection criteria in existing EMMAs, the population quickly converges toward equivalent Pareto optimal solutions which are easy to find. This results in loss of diversity in the decision space and makes them struggle to search for equivalent Pareto optimal solutions. Due to such an MMOP producing an imbalance between the emphasis on achieving convergence and preserving diversity in the decision space, we refer to it as an MMOP with the imbalance between convergence and diversity in the decision space (MMOP-ICD).

In Fig. 1, we show the solutions obtained by DNEA on the two-objective imbalanced distance minimization problem (IDMP) in Type 1 with $\alpha = 3$ in a single run. This particular run is associated with the result which is the closest to the mean performance metric value among 40 runs. This optimization problem is an MMOP-ICD, which is described in detail in Subsection III-B. As shown in Fig. 1 (a), the PS of this optimization problem are two line segments, i.e., $\boldsymbol{X}_a = \{(y, -0.5) | -0.6 \leq y \leq -0.4\}$ and $\boldsymbol{X}_b = \{(y, 0.5) | 0.4 \leq y \leq 0.6\}$. For any solution in $\boldsymbol{X}_a$, there is a solution in $\boldsymbol{X}_b$ which has the same objective vector. That is, they are equivalent Pareto optimal solutions. However, the solutions in $\boldsymbol{X}_b$ are more difficult to find than those in $\boldsymbol{X}_a$. We can see



(a) PS (b) Obtained Solutions

Fig. 1. The PS of a two-objective imbalanced distance minimization problem and the solutions obtained by DNEA on this problem in a single run.

from Fig. 1 (b) that DNEA cannot find the solutions in $\boldsymbol{X}_b$.

In this study, we first propose a benchmark set called IDMP. We show that existing EMMAs are unable to handle this imbalance by applying them to IDMP. Next, we present a novel evolutionary algorithm using a convergence-penalized density method (CPDEA) to solve these problems. In CPDEA, the distances among solutions in the decision space are transformed based on their local convergence quality. Their density values are estimated based on the transformed distances. The use of these density values as the selection criterion makes CPDEA efficient in searching for equivalent Pareto optimal solutions.

This paper has two main contributions:

1) One is the introduction of the concept of MMOP-ICD. Based on this concept, we propose a new benchmark set called IDMP. To the best of our knowledge, the concept of MMOP-ICD has never been established. Most existing benchmark MMOPs are not MMOP-ICDs. In those existing MMOPs, the difficulty in finding each equivalent Pareto optimal solution is the same, which is not likely occurred in a real-world MMOP. They are easy to solve by existing EMMAs due to this special characteristic. The use of only those existing MMOPs as test problems may mislead the performance evaluation of EMMAs and the research direction in the development of EMMAs. In this study, we introduce the concept of MMOP-ICD, where the difficulties of finding equivalent Pareto optimal solutions are different. MMOP-ICDs should be more general than others in the real world. In the proposed IDMP, the difficulty in finding each equivalent Pareto optimal solution is controllable. The numbers of objectives, decision variables, and equivalent Pareto optimal solutions are scalable. More importantly, IDMP is relevant to real-world applications, which is discussed in Subsection III-C. The use of the proposed IDMP as test problems will greatly facilitate the healthy development of the EMMO research field.
2) The other is the proposal of CPDEA. In CPDEA, we propose (i) a convergence-penalized density method as the selection criterion; (ii) a double k-nearest neighbor method to evaluate the diversity of solutions both in the objective and decision spaces; and (iii) two reproduction operators for exploration and exploitation, respectively. Our CPDEA is the first attempt to solve MMOP-ICDs. The selection criterion in CPDEA is conceptually different

from those in existing EMMAs. The choice of a selection criterion is the key issue in the design of any EMMA. Existing EMMAs are unable to solve MMOP-ICDs due to their convergence-first selection criteria. CPDEA uses a convergence-penalized density method as the selection criterion. This method can not only avoid premature convergence but also efficiently search for equivalent Pareto optimal solutions. Our experimental results demonstrate that CPDEA performs significantly better than the convergence-first EMMAs on IDMP.

The remainder of this paper is organized as follows. In Section II, related studies on EMMO are reviewed for the completeness of the presentation. The proposed IDMP and CPDEA are described in detail in Sections III and IV, respectively. Section V shows the experimental design and results. Section VI concludes the paper and provides future research directions.

II. Related Works

A. Multi-Modal Multi-Objective Optimization Problems

MMOPs are commonly seen in real-world applications [17], [18]. In [3], the definition of an MMOP is given as follows:

Definition 1. *For an MOP in (1), if there exists at least one local Pareto optimal solution or at least two equivalent Pareto optimal solutions for any point on the PF, then the MOP is considered as an MMOP.*

A local Pareto optimal solution is a solution that is non-dominated in its neighborhood in the decision space. Note that here we do not regard a Pareto optimal solution as a local one. As reported in [3], it is more complicated to solve an MMOP by finding its local Pareto optimal solution(s). Thus our goal in this study is to search for the equivalent Pareto optimal solutions corresponding to a set of well distributed points on the PF.

In [11], [12], [19], an MMOP is also defined as an MOP which have more than one equivalent Pareto optimal subset corresponding to the entire PF. This definition can be regarded as a special case of Definition 1. In this special case, the number of equivalent Pareto optimal solutions for every point on the PF is the same.

Several benchmarks of MMOPs have been proposed under these definitions. In [20], a two-objective MMOP called TWO-ON-ONE was proposed. It has two equivalent Pareto optimal subsets in the two-dimensional decision space. Later in [19], the same authors proposed three types of SYM-PART whose equivalent Pareto optimal subsets are nine line segments. Rotation and transformation of the line segments make the latter two types of SYM-PART more difficult to solve. In [11], the Omni-test problem was proposed. The numbers of both equivalent Pareto optimal subsets and decision variables are controllable.

In [21], MMF1-8 with two objectives and two decision variables were presented. MMF1 and MMF2 are relatively simple functions which were originally proposed in [12]. MMF3-8 are more complicated. In particular, The equivalent Pareto optimal subsets of MMF3 and MMF6 overlap in every dimension. Very recently, more scalable test problems were added to this benchmark set [22].

MMMOP1-6 were proposed in [3]. The numbers of objectives, decision variables and equivalent Pareto optimal solutions are scalable in all the MMMOPs. MMMOP2 and MMMOP5 address the imbalance between diversity in the objective space and that in the decision space. In MMMOP4, the number of equivalent Pareto optimal solutions for every point on the PF is not the same.

Distance minimization problems (DMPs) [23] or Polygon-based problems [24] are another kind of MMOPs. The objectives are to minimize the distances to the vertexes of several polygons in the two-dimensional decision space. Each polygon with the same shape and size is an equivalent Pareto optimal subset. In [13], two new types of DMPs were proposed. One is with the imbalance between diversity in the objective space and that in the decision space. The other has a number of local Pareto optimal solutions by changing the size of polygons. Very recently, four types of constrained DMPs are discussed in [25].

However, those MMOPs have not considered the imbalance between convergence and diversity in the decision space. In most of them, the difficulty in finding each equivalent Pareto optimal solution (subset) is the same. In Section III, we propose a novel benchmark set called IDMP in which some equivalent Pareto optimal subsets are more difficult to find.

B. Evolutionary Multi-Modal Multi-Objective Algorithms

Early attempts to solve MMOPs include Omni-optimizer [11], Niching-covariance matrix adaptation (Niching-CMA) [26], diversity integrating optimizer (DIOP) [27], etc. The framework of Omni-optimizer is NSGA-II. Different from NSGA-II, Omni-optimizer evaluates the crowding distances of solutions both in the objective and decision spaces. This increases its ability to search for multiple equivalent Pareto optimal subsets. In Niching-CMA, the distances between solutions in the objective and decision spaces are summed up to calculate the niche function. DIOP is a set-based algorithm where a weighted sum function of two indicators is used to evaluate a solution set. The two indicators evaluate diversity in the objective and decision spaces, respectively.

A number of new EMMAs were proposed in recent years. DN-NSGA-II [12] is also developed based on NSGA-II. It adopts a decision space based niching method in the mating selection. The same authors later proposed a multi-objective particle swarm optimization algorithm using ring topology and special crowding distance (MO_Ring_PSO_SCD) [21]. The special crowding distance is a modification of that in Omni-optimizer, which de-emphasizes the negative effect from boundary solutions. Besides, MO_Ring_PSO_SCD adopts a ring topology neighborhood [28] to enhance PSO's search ability.

A multi-modal multi-objective differential evolution optimization algorithm (MMODE) was proposed in [29]. The selection criteria in MMODE are similar to those in Omni-optimizer. A modified DE operator was introduced to promote diversity in the decision space.

In [13], DNEA was proposed, where the niche-based sharing function [30] is employed in both the objective and decision spaces to estimate a solution's density. One advantage of this sharing function is its inherent ability to normalize the objective and decision spaces into the same scale.

A modified MOEA/D with addition and deletion operators, termed MOEA/D-AD, was proposed in [14]. In MOEA/D-AD, a set of reference vectors is used to guarantee diversity in the objective space. Then for each reference vector, a niching method is used to promote diversity in the decision space. One characteristic feature of MOEA/D-AD is that each reference vector can have multiple solutions.

A multi-modal multi-objective evolutionary algorithm using two-archive and recombination strategies (TriMOEA-TA&R) was presented in [3]. Before the evolution, TriMOEA-TA&R analyzes the properties of decision variables and the relationships among them. This is useful in the recombination strategy to generate equivalent Pareto optimal solutions. The cooperative work between the convergence and the diversity archives increases the search efficiency. Similar to MOEA/D-AD, the diversity archive also uses a set of reference vectors in the objective space.

All of the above mentioned EMMAs are developed based on convergence-first EMOAs. Therefore, their selection criteria are similar to these EMOAs. In Sections III and V, we will show that the convergence-first EMMAs are unable to maintain diversity in the decision space when solving MMOP-ICDs. To overcome this issue, we propose a novel evolutionary algorithm using a convergence-penalized density method in Section IV.

III. Optimization Problems with Imbalance Between Convergence and Diversity in the Decision Space

A. *Definition of MMOP-ICD*

In this study, we give the definition of an MMOP-ICD as follows:

Definition 2. *An MMOP is defined to be an MMOP-ICD, if it satisfies one or both of the following conditions:*

1) *For a point on the PF, solutions close to one equivalent Pareto optimal solution are more likely to dominate solutions close to another equivalent Pareto optimal solution.*
2) *For a point on the PF, the complexity of searching for one equivalent Pareto optimal solution is lower than that of another equivalent Pareto optimal solution.*

Here, the complexity of searching for an equivalent Pareto optimal solution can be regarded as the required computation cost to find it. When an evolutionary algorithm is used to search for the equivalent Pareto optimal solution, the higher the complexity is, the more fitness evaluations are required to find it. Equivalent Pareto optimal solutions can be replaced by equivalent Pareto optimal subsets in this definition.

It is difficult for a convergence-first EMMA mentioned in Section II to find equivalent Pareto optimal solutions of an MMOP-ICD. Due to the first condition, the solutions close to some equivalent Pareto optimal solutions have better convergence quality than those close to other equivalent Pareto optimal solutions. Due to the second condition, solutions with good convergence quality are quickly found around some equivalent Pareto optimal solutions, while the convergence quality of the solutions found around other equivalent Pareto optimal solutions is poor. We can infer that the population quickly converges to some equivalent Pareto optimal solutions, since the solutions around them are preferred by the convergence-first selection criterion.

In this study, to check whether an MMOP satisfies the first condition, we calculate the average Pareto rank of solutions close to each equivalent Pareto optimal solution (subset). If the average Pareto rank of solutions close to every equivalent Pareto optimal solution (subset) is the same, the first condition is not met.

To check whether the second condition is met, we use an evolutionary algorithm to solve several new optimization problems. Each of these optimization problems is transformed from the MMOP. It has one optimal solution (or PS), which is the same to an equivalent Pareto optimal solution (or subset) of the MMOP. The computational resource of solving the optimization problem reflects the complexity of searching for the corresponding equivalent Pareto optimal solution (subset).

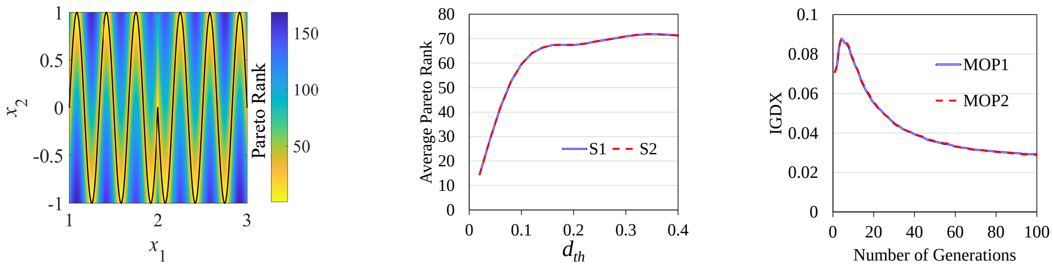
Let us take MMF1 [21] as an example to explain how to check whether an MMOP satisfies these conditions and why it is not an MMOP-ICD. MMF1 is formulated as follows:

$$\begin{aligned} &\min f_1(\boldsymbol{x}) = |x_1 - 2|, \\ &\min f_2(\boldsymbol{x}) = 1 - \sqrt{|x_1 - 2|} + 2(x_2 - \sin(6\pi|x_1 - 2| + \pi))^2, \\ &\text{s.t. } x_1 \in [1, 3], x_2 \in [-1, 1]. \end{aligned} \tag{2}$$

In Fig. 2 (a), we show the fitness landscape based on Pareto rank [31] of MMF1 for an intuitive understanding. In this figure, $10,201$ (101×101) solutions are uniformly sampled in the entire decision space. They are in different colors based on their Pareto ranks. A brighter (warmer) tone corresponds to a lower Pareto rank (i.e., solutions with better convergence quality). The curve segments with $x_1 \in [1, 2]$ and $x_1 \in [2, 3]$ are the first and second equivalent Pareto optimal subsets, respectively. We can see from Fig. 2 (a) that the patterns with $x_1 \in [1, 2]$ and $x_1 \in [2, 3]$ are symmetrical about $x_1 = 2$. This implies that the difficulties in finding the two equivalent Pareto optimal subsets are the same.

To check whether MMF1 meets the first condition, in Fig. 2 (b), we show the average Pareto rank of solutions close to each equivalent Pareto optimal subset based on Fig. 2 (a). We first give a threshold value d_{th}, which is set to $0.02, 0.04, \ldots, 0.4$, respectively. Then, we calculate the distances between each equivalent Pareto optimal subset and the solutions in Fig. 2 (a). The solutions whose distances to the first (second) equivalent Pareto optimal subset are smaller than d_{th} are denoted as a solution set S_1 (S_2). Finally, we use the non-dominated sorting to rank all the solutions in S_1 and S_2, and calculate the average Pareto rank of solutions in each solution set. We can see from Fig. 2 (b) that the average Pareto ranks of S_1 and S_2 are exactly the same. This suggests that MMF1 does not satisfy the first condition.

To check whether MMF1 meets the second condition, we use DNEA [13] to solve two MOPs (denoted as MOP1 and



(a) Fitness landscape (b) Average Pareto rank (c) Results of IGDX

Fig. 2. (a) The fitness landscapes based on Pareto rank of MMF1. The curve segments with $x_1 \in [1, 2]$ and $x_1 \in [2, 3]$ are the first and second equivalent Pareto optimal subsets, respectively. (b) The average Pareto rank of solutions close to each equivalent Pareto optimal subset. S_1 (S_2) is a set of solutions whose distances to the first (second) equivalent Pareto optimal subset are smaller than d_{th}. (c) The IGDX value (averaged over 40 runs) at each generation when DNEA was used to solve MOP1 and MOP2, respectively.

MOP2). The objective functions of these MOPs are the same to those of MMF1. However, the bound (constraint) of x_1 is $[1, 2]$ and $[2, 3]$ in MOP1 and MOP2, respectively. Therefore, the PSs of MOP1 and MOP2 are the same to the first and second equivalent Pareto optimal subset of MMF1, respectively. Searching for the PS of MOP1 (MOP2) is equivalent to searching for the first (second) equivalent Pareto optimal subset of MMF1. If it takes less computational resource to find one PS than the other, MMF1 meets the second condition.

In Fig. 2 (c), we show the results of averaged Inverted Generational Distance in the decision space (IGDX) [32] over 40 runs when DNEA was used to solve the MOPs[1]. IGDX measures the distance in the decision space between each reference point on the PS and its nearest point in the obtained solution set. Here, we uniformly sample $1,000$ points on the true PS as the reference points. A faster decrease of the IGDX value over generations indicates that it takes less computational resource for the algorithm to find the PS. That is, the complexity of searching for the PS is lower. Note that the decrease rate of IGDX is also dependent on the search power of the algorithm. Although the results of IGDX cannot tell the exact searching complexity, we can know which problem corresponds to a higher complexity (i.e., which PS is more difficult to find). From Fig. 2 (c), we can see that the curves of IGDX for the two problems are the almost the same (they are very slightly different since evolutionary algorithm is stochastic). Therefore, the complexities of searching for the equivalent Pareto optimal subsets of MMF1 are the same, and the second condition is not met.

According to the above discussions, we can see that MMF1 is not an MMOP-ICD since it satisfies neither of the conditions. In the supplementary material, we also investigate the property of MMMOP1 [3], Omin-test problem [11], and SYM-PART [19] in the same way. The results show that these MMOPs are not MMOP-ICDs. Although a few existing MMOPs, such as the variants of MMF1 [22], satisfy one or both conditions, the concept of MMOP-ICD has not been proposed nor discussed in the literature.

In the next subsection, we introduce four types of two-objective IDMPs which are MMOP-ICDs. The imbalance of the first two and the third types is mainly caused by the first and second conditions, respectively. Both conditions are related to the imbalance of the fourth type.

B. Two-Objective Imbalanced Distance Minimization Problems

The two-objective IDMPs proposed in this study have two equivalent Pareto optimal subsets in the two-dimensional decision space. Their general objective functions are given as follows:

$$\begin{aligned} & f_1(\boldsymbol{x}) = \min\left(|x_1 + 0.6| + g_1(\boldsymbol{x}), |x_1 - 0.4| + g_2(\boldsymbol{x})\right), \\ & f_2(\boldsymbol{x}) = \min\left(|x_1 + 0.4| + g_1(\boldsymbol{x}), |x_1 - 0.6| + g_2(\boldsymbol{x})\right), \\ & \text{ s.t. } x_1 \in [-1, 1], x_2 \in [-1, 1], \end{aligned} \tag{3}$$

where $g_1(\boldsymbol{x}) \geq 0$ and $g_2(\boldsymbol{x}) \geq 0$ are difficulty functions corresponding to the first and second equivalent Pareto optimal subsets, respectively. The first equivalent Pareto optimal subset is $x_1 \in [-0.6, -0.4]$ when $g_1(\boldsymbol{x}) = 0$, and the second equivalent Pareto optimal subset is $x_1 \in [0.4, 0.6]$ when $g_2(\boldsymbol{x}) = 0$. We denote them as EPS1 and EPS2, respectively.

The difficulty functions in each type of two-objective IDMP are given as follows:

Type 1:

$$\begin{aligned} & g_1(\boldsymbol{x}) = |x_2 + 0.5|, \\ & g_2(\boldsymbol{x}) = \alpha|x_2 - 0.5|, \\ & \alpha \geq 1. \end{aligned} \tag{4}$$

Type 2:

$$\begin{aligned} & g_1(\boldsymbol{x}) = 100(x_2 + 0.5)^2, \\ & g_2(\boldsymbol{x}) = 100|x_2 - 0.5|^{2-\alpha}, \\ & \alpha \in [0, 2). \end{aligned} \tag{5}$$

Type 3:

$$\begin{aligned} & g_1(\boldsymbol{x}) = 100(x_2 + 0.5)^2, \\ & g_2(\boldsymbol{x}) = 100(x_2 - 0.5 + \alpha(x_1 - 0.5))^2, \\ & \alpha \in [0, 5]. \end{aligned} \tag{6}$$

[1]Parameter settings of DNEA in Section III: Simulated binary crossover and polynomial mutation are used. The distribution index is specified as 20 in crossover and mutation. The crossover and mutation probabilities are 1.0 and 0.5, respectively. The population size is set to be 60. The termination criterion is that the population has evolved 300 generations. For each problem, 40 independent runs are executed.

Type 4:

$$\begin{aligned} &g_1(\boldsymbol{x}) = 100((x_2+0.5)^2 - \cos(2\pi(x_2+0.5))+1),\\ &g_2(\boldsymbol{x}) = 100((x_2-0.5)^2 - \cos(2\pi\alpha(x_2-0.5))+1),\\ &\alpha \text{ is a positive integer.} \end{aligned} \tag{7}$$

In each type, α is a parameter which determines the difficulty in finding EPS2. The larger value of α, the more difficult to find EPS2. In particular, when α is 1, 0, 0, and 1 in Types 1, 2 , 3, and 4, respectively, the difficulties in finding EPS1 and EPS2 are exactly the same. When α is larger than 1, 0, 0, and 1 in Types 1, 2 , 3, and 4, respectively, EPS2 is more difficult to find than EPS1.

In the following, we use the same method as that in Subsection III-A to explain why these optimization problems are MMOP-ICDs:

(1) For an intuitive understanding, in Fig. 3, we show the fitness landscape based on Pareto rank of these optimization problems, where the line segments are EPS1 ($x_1 \in [-0.6, -0.4]$) and EPS2 ($x_1 \in [0.4, 0.6]$).

(2) To check the first condition, based on Fig. 3, Fig. 4 shows the average Pareto rank of solutions close to each equivalent Pareto optimal subset in each type, where S_1 and S_2 are the solutions whose distances to EPS1 and EPS2 are smaller than d_{th}, respectively. ($d_{th} = 0.02, 0.04, \ldots, 0.4$)

(3) To check the second condition, in each sub-figure of Fig. 5, we show the results of IGDX over 40 runs when DNEA was used to solve the following optimization problem:

$$\begin{aligned} &f_1(\boldsymbol{x}) = |x_1 - 0.4| + g_2(\boldsymbol{x}),\\ &f_2(\boldsymbol{x}) = |x_1 - 0.6| + g_2(\boldsymbol{x}),\\ &\text{s.t. } x_1 \in [-1,1], x_2 \in [-1,1], \end{aligned} \tag{8}$$

where $g_2(\boldsymbol{x})$ is the same as that in the corresponding type. For each type, α in $g_2(\boldsymbol{x})$ is set to four different values as shown in Table I. The PS of the MOP in (8) is the same with EPS2 of the MMOP in (3). When α is 1, 0, 0, and 1 in Types 1, 2 , 3, and 4, respectively, searching for this PS (EPS2) is equivalent to searching for EPS1.

For Type 1, we can observe from Figs. 3 (a) and 4 (a) that solutions close to EPS1 generally have lower Pareto ranks than solutions close to EPS2. This indicates that the DMP in Type 1 satisfies the first condition. However, we can see from Fig. 5 (a) that the curves of IGDX with different settings of α are almost the same. This implies that the complexities of searching for EPS1 and EPS2 with these values of α are almost the same. Therefore, the IDMP in Type 1 is an MMOP-ICD, and the imbalance is mainly caused by the first condition, i.e., solutions close to one equivalent Pareto optimal subset have a higher chance to dominate solutions close to the other.

From Figs. 3 (b), 4 (b), and 5 (b), we can see that the situation in Type 2 is similar to that in Type 1. Thus the imbalance in Type 2 is also mainly caused by the first condition.

We can see from Fig. 3 (c) that the pattern around EPS2 looks a rotated version from the pattern around EPS1. The difference between the average Pareto ranks of S_1 and S_2 in Fig. 4 (c) is quite small. Therefore, the first condition is unlikely to be the primary cause for the imbalance in Type 3.

TABLE I
RESULTS OF DNEA IN SOLVING THE TWO-OBJECTIVE IDMPS UNDER DIFFERENT SETTINGS OF α

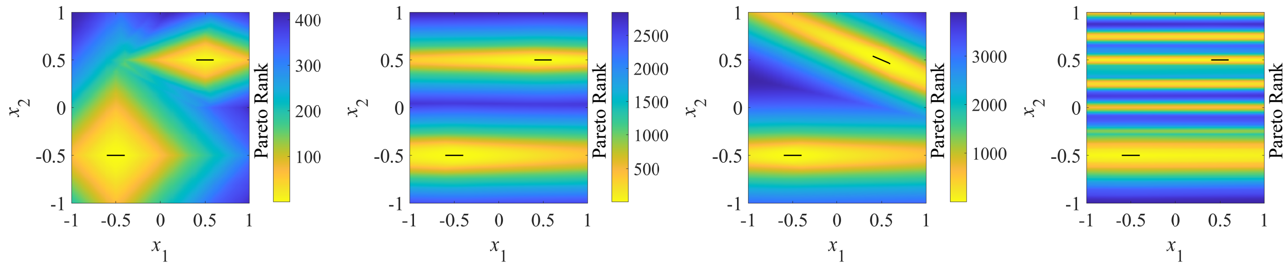
		Percentage of Obtianed Solutions		Number of Runs		
		Close to EPS1	Close to EPS2	Only EPS1	Both	Only EPS2
Type 1	α = 1	49.5%	50.5%	13	14	13
	α = 2	82.9%	17.1%	27	12	1
	α = 3	92.8%	7.2%	34	6	0
	α = 4	95.2%	4.8%	37	3	0
Type 2	α = 0	50.2%	49.8%	9	22	9
	α = 0.2	66.5%	33.5%	21	11	8
	α = 0.4	74.4%	25.6%	23	13	4
	α = 0.6	94.0%	6.0%	37	1	2
Type 3	α = 0	50.2%	49.8%	9	22	9
	α = 0.2	59.0%	41.0%	12	20	8
	α = 0.4	73.6%	26.4%	13	24	3
	α = 0.6	90.9%	9.1%	27	12	1
Type 4	α = 1	48.9%	51.1%	11	17	12
	α = 2	77.7%	22.3%	27	8	5
	α = 4	93.7%	6.3%	37	1	2
	α = 8	97.5%	2.5%	39	0	1

The interaction (or correlation) between x_1 and x_2 in $g_2(\boldsymbol{x})$ increases the complexity of searching for EPS2, which makes the second condition satisfied and results in the imbalance in Type 3. The larger value of α, the stronger interaction between x_1 and x_2 in $g_2(\boldsymbol{x})$ and the higher complexity of searching for EPS2. This can be verified from Fig. 5 (c), where a large value of α slows down the convergence rate according to IGDX. Thus the second condition mainly causes the imbalance in Type 3.

For Type 4, on one hand, we can see from Figs. 3 (d) and 4 (d) that solutions close to EPS1 are more likely to have lower Pareto ranks than solutions close to EPS2. On the other hand, due to the local optima in $g_2(\boldsymbol{x})$, there are some local Pareto optimal regions close to EPS2 (e.g., the regions where x_2 is around 0.25 and 0.75 in Fig. 3 (d)). This increases the complexity of searching for EPS2 since the population may trap into the local Pareto optimal regions during the search process. The number of local Pareto optimal regions increases as α increases. We can see from Fig. 5 (d) that increasing α leads to a slow decrease of the IGDX value over generations. Therefore, the imbalance of Type 4 is induced by both conditions.

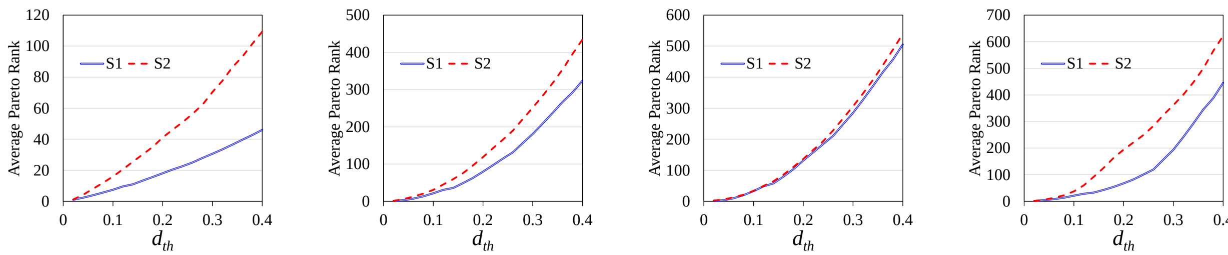
To have a comprehensive understanding of the deficiency of existing EMMAs in solving MMOP-ICDs, we show the results of DNEA in solving the four types of two-objective IDMPs in Table I. Table I lists the average percentage of solutions close to each equivalent Pareto optimal subset over 40 runs. Note that for the results in Table I, if the distance between a solution and EPS1 is smaller than 0.04, it is viewed to be close to EPS1. A solution close to EPS2 is defined in the same way. Table I also lists the number of runs where the obtained solutions are close to only EPS1, only EPS2, and both of them (i.e., some are close to EPS1 and others are close to EPS2), respectively.

We can see from Table I that when α is 1, 0, 0, and 1 in Types 1, 2, 3, and 4, respectively, i.e., the difficulties in finding EPS1 and EPS2 are the same, the percentages of the



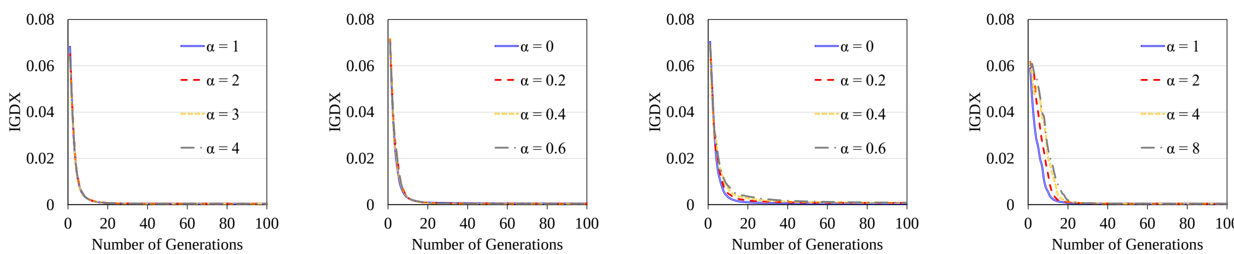
(a) Type 1 ($\alpha = 3$) (b) Type 2 ($\alpha = 0.4$) (c) Type 3 ($\alpha = 0.4$) (d) Type 4 ($\alpha = 4$)

Fig. 3. The fitness landscapes based on Pareto rank of four types of two-objective IDMPs. For each type, $10,201$ (101×101) solutions are uniformly sampled in the entire decision space. They are in different colors based on their Pareto ranks. A brighter (warmer) tone corresponds to a lower Pareto rank. The line segments are the first ($x_1 \in [-0.6, -0.4]$) and second ($x_2 \in [0.4, 0.6]$) equivalent Pareto optimal subsets, i.e., EPS1 and EPS2.



(a) Type 1 ($\alpha = 3$) (b) Type 2 ($\alpha = 0.4$) (c) Type 3 ($\alpha = 0.4$) (d) Type 4 ($\alpha = 4$)

Fig. 4. The average Pareto rank of solutions close to each equivalent Pareto optimal subset based on Fig. 3. S_1 and S_2 are the solutions whose distances to EPS1 and EPS2 are smaller than d_{th}, respectively.



(a) Type 1 (b) Type 2 (c) Type 3 (d) Type 4

Fig. 5. The IGDX value (averaged over 40 runs) with respect to generations when DNEA was to solve the optimization problem in (8) under different settings of α.

obtained solutions close to EPS1 and EPS2 are almost equal to each other (around 50%). As α increases, the percentages of the obtained solutions close to EPS1 and EPS2 increase and decrease, respectively. This suggests that DNEA struggles to search for EPS2. We get similar observations according to the number of runs.

One interesting thing to note is that when α is 1, 0, 0, and 1 in Types 1, 2 , 3, and 4, respectively, the number of runs where the obtained solutions are close to both equivalent Pareto optimal subsets is much less than 40. That is, DNEA may even fail when the difficulties in finding EPS1 and EPS2 are the same. Due to the stochastic search process, there is a chance that more solutions are closer to one equivalent Pareto optimal subset than the other equivalent Pareto optimal subset in some generation, which leads to the convergence to that equivalent Pareto optimal subset through the convergence-first selection criterion. Therefore, even without the imbalance, the convergence-first selection criterion may have a negative effect in searching for all equivalent Pareto optimal subsets.

Note that we have also applied MO_Ring_PSO_SCD, TriMOEA-TA&R, DN-NSGA-II, and Omni-optimizer to these problems under the same settings of α as in Table I. Similar results to those of DNEA are observed. Those results are not shown in this section due to the page limitation. We will show and discuss their performance in Section V.

C. Scalable Imbalanced Distance Minimization Problems

In this subsection, we extend four types of IDMPs with more than two objectives and two decision variables. In the proposed IDMPs, the numbers of objectives, decision variables, and equivalent Pareto optimal subsets are denoted as M, n, and P, respectively. There exist P regular polygons with the same size in the sub-space x_1-x_2. $\boldsymbol{X}_{p,m} = (X_{p,m,1}, X_{p,m,2})$ and $\boldsymbol{C}_p = (C_{p,1}, C_{p,2})$ denote the mth vertex and the center of the pth polygon, respectively ($m = 1, \ldots, M, p = 1, \ldots, P$). The radius r of a polygon is the distance between a vertex and the center. These polygons should be far away from each other. That is, the distances between polygons should be much larger than their radius.

The mth ($m = 1, \ldots, M$) objective function of an IDMP is

$$\begin{aligned} &f_m(\boldsymbol{x}) = \min_{p=1,\ldots,P} \{d(\boldsymbol{X}_{p,m}, x_1, x_2) + g_p(\boldsymbol{x})\}, \\ &\text{s.t. } \boldsymbol{x} = (x_1, \ldots, x_n) \in [-1, 1]^n, \end{aligned} \tag{9}$$

where $d(\boldsymbol{X}_{p,m}, x_1, x_2)$ is the Euclidean distance between $\boldsymbol{x}$ and $\boldsymbol{X}_{p,m}$ in the sub-space x_1-x_2, i.e., $\sqrt{(X_{p,m,1} - x_1)^2 + (X_{p,m,2} - x_2)^2}$, and $g_p(\boldsymbol{x}) \geq 0$ is the difficulty function corresponding to the pth equivalent Pareto optimal subset. The pth equivalent Pareto optimal subset is the inside (including the sides) of the pth polygon with $g_p(\boldsymbol{x}) = 0$, which has been proven in [33].

In each type of IDMP, $g_p(\boldsymbol{x})$ is given as follows:

Type 1:

$$\begin{aligned} &g_p(\boldsymbol{x}) = \sum_{i=3}^{n} \alpha_{p,i} |x_i + 1 - \tfrac{2p}{P+1}|, \\ &\alpha_{p,i} \geq 1, i = 3, \ldots, n. \end{aligned} \tag{10}$$

Type 2:

$$\begin{aligned} &g_p(\boldsymbol{x}) = \sum_{i=3}^{n} 100 |x_i + 1 - \tfrac{2p}{P+1}|^{2-\alpha_{p,i}}, \\ &\alpha_{p,i} \in [0, 2), i = 3, \ldots, n. \end{aligned} \tag{11}$$

Type 3:

$$\begin{aligned} &g_p(\boldsymbol{x}) = \sum_{i=3}^{n} 100(x_i + 1 - \tfrac{2p}{P+1} + \alpha_{p,i} y_p(x_1, x_2))^2, \\ &y_p(x_1, x_2) = x_1 - C_{p,1} + x_2 - C_{p,2}, \\ &\alpha_{p,i} \in [0, \tfrac{p}{(P+1)r}], i = 3, \ldots, n. \end{aligned} \tag{12}$$

Type 4:

$$\begin{aligned} g_p(\boldsymbol{x}) = \sum_{i=3}^{n} 100[&(x_i + 1 - \frac{2p}{P+1})^2 \\ &- \cos(2\pi\alpha_{p,i}(x_i + 1 - \frac{2p}{P+1})) + 1], \\ \alpha_{p,i} \text{ is a} &\text{ non-negative integer}, i = 3, \ldots, n. \end{aligned} \tag{13}$$

For each type, the larger value of $\alpha_{p,i}, p = 1, \ldots, P, i = 3, \ldots, n$, the greater difficulty in finding the pth equivalent Pareto optimal subset.

In this study, we specify a three-objective problem with three decision variables and a four-objective problem with four decision variables for each type, i.e., eight problems in total. Each problem has four polygons with $r = 0.1$. The centers of the first, second, third, and fourth polygons are $(-0.5, -0.5)$, $(0.5, -0.5)$, $(0.5, 0.5)$, and $(-0.5, 0.5)$, respectively. Fig. 6 shows the polygons of the three-objective problems. The settings of $\alpha_{p,i} (p = 1, 2, 3, 4, i = 3, 4)$ in these problems are listed in Table II. By such settings, the difficulty in finding the first, second, third, and fourth equivalent Pareto optimal subsets increases gradually.

It is worthy to note that IDMPs are relevant to real-world applications, such as map-based problems [34]. For a simple example, let us assume Fig. 6 is a two-dimensional map, where $\boldsymbol{X}_{1,1}$ and $\boldsymbol{X}_{2,1}$ are schools, $\boldsymbol{X}_{1,2}$ and $\boldsymbol{X}_{2,2}$ are supermarkets, and $\boldsymbol{X}_{1,3}$ and $\boldsymbol{X}_{2,3}$ are hospitals. The goal is to find an apartment on the map which has the lowest travel cost to

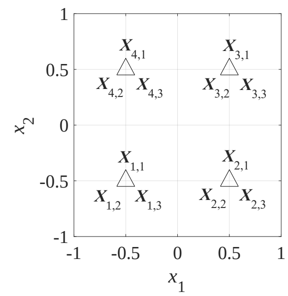


Fig. 6. The polygons of the three-objective IDMPs. $\boldsymbol{X}_{p,m}, p = 1, \ldots, 4, m = 1, \ldots, 3$, is the m-th vertex of the pth polygon. Each Polygon is an equivalent Pareto optimal subset when $g_p(\boldsymbol{x}) = 0$.

TABLE II
SETTINGS OF $\alpha_{p,i} (p = 1, 2, 3, 4, i = 3, 4)$ IN THREE- AND FOUR-OBJECTIVE IDMPS

	Three Objective				Four Objective							
	$\alpha_{1,3}$	$\alpha_{2,3}$	$\alpha_{3,3}$	$\alpha_{4,3}$	$\alpha_{1,3}$	$\alpha_{2,3}$	$\alpha_{3,3}$	$\alpha_{4,3}$	$\alpha_{1,4}$	$\alpha_{2,4}$	$\alpha_{3,4}$	$\alpha_{4,4}$
Type 1	1	2	3	4	1	2	3	4	1	2	3	4
Type 2	0	0.2	0.4	0.6	0	0.2	0.4	0.6	0	0.2	0.4	0.6
Type 3	0	0.1	0.2	0.3	0	0.05	0.1	0.15	0	0.05	0.1	0.15
Type 4	1	2	3	4	1	2	3	4	0	0	0	0

either school, supermarket, and hospital. In this optimization problem, the first two decision variables are the location of the apartment, i.e., x_1 and x_2. The travel methods are a variety of public transport services such as railways and buses operated by different companies. The choice of travel methods is the third decision variable, i.e., x_3. The travel cost is the ticket price, which is based on the distance between two locations and government policies. This is a three-objective optimization problem. The mth ($m = 1, 2, 3$) objective, i.e., the lower travel cost from the apartment to $\boldsymbol{X}_{1,m}$ and $\boldsymbol{X}_{2,m}$, can be formulated as

$$f_m(\boldsymbol{x}) = \min_{p=1,2} \{d(\boldsymbol{X}_{p,m}, x_1, x_2) + g_p(x_3)\}, \tag{14}$$

where $d(\boldsymbol{X}_{p,m}, x_1, x_2)$ is the distance between the apartment and $\boldsymbol{X}_{p,m}$, and $g_p(x_3) \geq 0$ is the difficultly function determined by government policies. Triangles $\boldsymbol{X}_{1,1}\boldsymbol{X}_{1,2}\boldsymbol{X}_{1,3}$ and $\boldsymbol{X}_{2,1}\boldsymbol{X}_{2,2}\boldsymbol{X}_{2,3}$ are the first and second equivalent Pareto optimal subsets, respectively.

The region around triangle $\boldsymbol{X}_{2,1}\boldsymbol{X}_{2,2}\boldsymbol{X}_{2,3}$ is downtown. Since this region is too busy and crowded, the government has implemented a policy to raise most ticket prices for rides from the other regions to this region. That is, $g_2(x_3)$ is more likely to be larger than $g_1(x_3)$, which makes the second equivalent Pareto optimal subset more difficult to find than the first one. As a result, it is very difficult to find a good (Pareto optimal) apartment close to $\boldsymbol{X}_{2,1}\boldsymbol{X}_{2,2}\boldsymbol{X}_{2,3}$ for a convergence-first EMMA. People who like to live in the downtown may be disappointed if they use such an EMMA to search for an apartment. Our CPDEA in the next section can help these people by finding both equivalent Pareto optimal subsets.

IV. Proposed Method

A. Motivation

One challenge of handling the imbalance between convergence and diversity in the decision space is to avoid premature convergence. When a convergence-first EMMA is used to solve an MMOP-ICD, the population is very likely to quickly converge to some equivalent Pareto optimal solutions which are easy to find. This phenomenon has been discussed in Subsection III-A and confirmed by the experimental results in Subsection III-B.

One way to deal with this challenge is to treat diversity in the decision space more important than convergence, i.e., to use a decision space diversity measure as the primary selection criterion. To the best of our knowledge, there exist no EMMAs using such a selection criterion. However, we can easily imagine several methods in this way. For example, one intuitive method is to use the density value in the decision space and the convergence quality as the primary and secondary selection criteria, respectively. For another example, we can use the idea of M2M [16] (aforementioned in Section I) in the decision space. Similar to the decomposition in the objective space, we can decompose the decision space into a number of sub-regions. In each sub-region, at least one solution is preserved regardless of its convergence quality.

However, those methods are infeasible to solve MMOP-ICDs due to their low search efficiency, i.e., poor convergence ability. In the first method, the density value is a scalar. When we select (or delete) solutions with the best (or worst) density values one by one, the probability of having two solutions with the same density value is extremely low. Consequently, the secondary selection criterion will never be used. The solutions obtained by this method is most likely to be uniformly distributed in the entire decision space. That is, the population may never converge to the PS. The second method may preserve solutions in a large number of sub-regions with no Pareto optimal solution when the PS locates in a few small sub-regions. In such a case, the search efficiency is very low. Therefore, high search efficiency is another challenge of handling the imbalance between convergence and diversity in the decision space.

We can see from the above discussions that simultaneously dealing with both challenges (i.e., avoiding premature convergence and achieving high search efficiency) is not an easy task. Not only avoiding premature convergence but also achieving high search efficiency is the key to solve MMOP-ICDs.

Our basic idea is simple. We combine the density value in the decision space and the convergence quality into a single selection criterion. More specifically, the density value of a solution is penalized by its local convergence quality. Note that a solution's local convergence quality is mainly measured by its neighborhood, whereas the global one is measured by all the solutions. The reason for using a local convergence quality measure will be explained later in Fig. 7.

Selection based on this convergence-penalized density (CPD) method leads to solutions covering the entire decision space. Thus the population will not prematurely converge to some equivalent Pareto optimal solutions which are easy to find. Meanwhile, solutions with good local convergence quality crowd around some promising regions, since their density values are less penalized than those of solutions with poor local convergence quality. This significantly increases the efficiency of searching for more equivalent Pareto optimal solutions. Therefore, the CPD method can well handle the aforementioned challenges.

Our CPD method is based on the assumption that a solution close to an equivalent Pareto optimal solution has good convergence quality. This is similar to the basic assumption in evolutionary algorithms, where solutions with good fitness values are close to the optimal solution(s).

Fig. 7 illustrates the population in the decision space during the search process when the proposed CPD method is used to solve an MMOP with two equivalent Pareto optimal subsets. In this MMOP, the equivalent Pareto optimal subset on the left is easier to find than the other. We also illustrate that of the convergence-first method in most existing EMMAs (such as DNEA and Omni-optimizer) for comparison. In Fig. 7, for both methods in the initial phase, the population is widely distributed in the entire decision space by random generation.

In the early phase, for the CPD method, the solutions around the left subset are denser than the other solutions due to their global (local as well) good convergence quality. Meanwhile, the other solutions widely spread to explore potential Pareto optimal solutions. Thus, it is promising to generate an offspring close to or in the right subset. By contrast, for the convergence-first method, most solutions close to the right subset are eliminated due to their poor convergence quality. The chance to generate an offspring close to or in the right subset is low.

In the later phase, for the CPD method, the algorithm finds some solutions in the left subset. It also finds some solutions gathering around the right subset due to their good local convergence quality. An offspring closer to or in the right subset is very likely to be generated by these solutions. If we use a global convergence quality measure instead of a local one, these solutions will become sparser since they have worse global convergence quality than those in the left subset. Then the search efficiency will be diminished. This is why we use a local convergence quality measure in the CPD method. On the other hand, for the convergence-first method, the population continues to converge to the left subset in the later phase. There is almost no chance to generate an offspring close to the right subset.

In the final phase (i.e., when the stopping criterion is met), most solutions selected by the CPD method are well distributed in both subsets. However, the convergence-first method can find only solutions in the left subset.

The illustration in Fig. 7 shows the conceptual advantage of our CPD method when compared to the convergence-first method in the existing EMMAs. Note that equivalent Pareto optimal subsets can be replaced with equivalent Pareto optimal solutions in this example. In the following subsections, we will introduce the proposed evolutionary algorithm using the CPD method in detail.

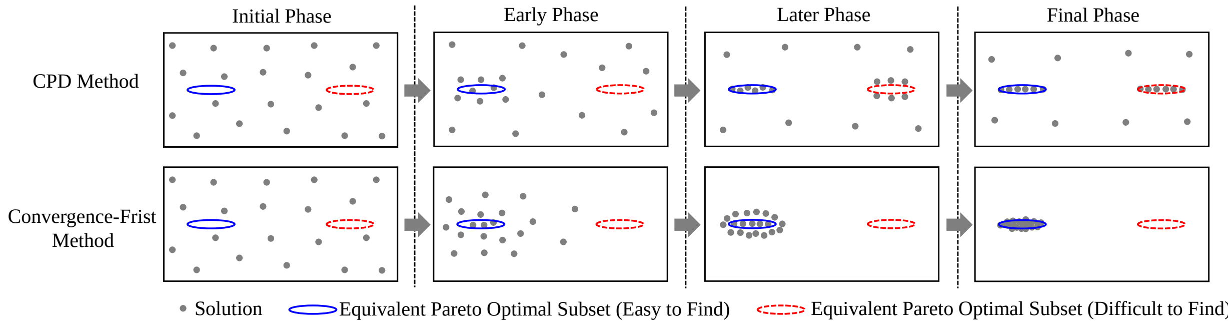


Fig. 7. Illustration of the population in the decision space during the search process when the proposed CPD method and the convergence-first method are used for selection, respectively.

Algorithm 1 General Framework

Require: P (Population), N (Population Size), A (Archive)
1: P = *Initialization* (P);
2: A = *Archive_Update* (P);
3: **while** the stopping criterion is not met **do**
4: Q = *Reproduction* (P, A);
5: P = *CPD-Based_Environmental_Selection* ($P \cup Q$);
6: A = *Archive_Update* ($A \cup Q$);
7: **end while**
8: **return** A

B. General Framework

Algorithm 1 presents the overall framework of the proposed CPDEA. In Algorithm 1, a population P of size N is randomly initialized in the decision space (line 1). An archive A is updated based on P (line 2). We set the maximum size of A equal to N in this study. The stopping criterion is the predefined maximum number N_E of evaluated solutions. While the stopping criterion has not been met (line 3), an offspring Q is generated by a reproduction operator (line 4), and then both P and A are updated by adding Q (lines 5 and 6). Here, we use a steady-state method [35] to update P and A, i.e., only one offspring is produced and added to P and A in each generation. One advantage of the steady-state method is the high selection pressure induced by deleting the worst solution in each generation [36]. P in the next generation is selected by the CPD-based environmental selection. The archive A preserves only non-dominated solutions with good diversity both in the objective and decision spaces. At the end of evolution, A is returned.

We will describe CPD-based environmental selection, archive update, and reproduction operators in Subsections IV-C, IV-D and IV-E, respectively.

C. CPD-Based Environmental Selection

In our CPD-based environmental selection, P is updated by removing the solution with the worst CPD value in $P \cup Q$ since Q contains only one offspring. To calculate CPD values of solutions, we first evaluate their local convergence quality. Then we transform the distance between each pair of solutions based on their local convergence quality. Finally, we estimate the density values based on the transformed distances.

1) Local Convergence Quality: The local convergence quality c_i of solution $\boldsymbol{x}_i \in P \cup Q, i = 1, \ldots, N+1$ is calculated as follows:

$$c_i = \sum_{j=1}^{N+1} w_{i,j} B_{i,j}, \tag{15}$$

where $w_{i,j}$ is a weight parameter, $B_{i,j}$ is 1 if $\boldsymbol{x}_i$ is dominated by $\boldsymbol{x}_j$ and 0 otherwise.

$w_{i,j}$ is defined as follows:

$$w_{i,j} = \tfrac{1}{\sigma\sqrt{2\pi}} \exp\left(-\tfrac{d_{i,j}^2}{2\sigma^2}\right), \tag{16}$$

where $d_{i,j}$ is the Euclidean distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in the decision space. We can see that $w_{i,j}$ is actually the probability density function of normal distribution with mean 0 and standard deviation σ.

σ is given as

$$\sigma = \eta\left(\tfrac{\prod_{l=1}^{n}(x_l^{\text{upper}} - x_l^{\text{lower}})}{N+1}\right)^{\frac{1}{n}}, \tag{17}$$

where η is a parameter given by the user, and x_l^{upper} and x_l^{lower} are the upper and lower bounds of the lth decision variable, respectively. Please refer to Subsection II-A in the supplementary material for the sensitivity analysis of η.

According to (15), the smaller the value of c_i is, the better the local convergence quality is. Particularly, if $\boldsymbol{x}_i$ is a non-dominated solution, c_i is 0. The reason for using (16) is that (16) is a monotonically decreasing function. This function can emphasize the local convergence quality by giving a larger value to $w_{i,j}$ for $\boldsymbol{x}_j$ which is closer to $\boldsymbol{x}_i$. Any other functions which are monotonically decreasing with respect to $d_{i,j}$ could be also used instead of (16).

Fig. 8 illustrates the calculation of c_1 in a solution set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_7\}$. In Fig. 8, the Euclidean distance between $\boldsymbol{x}_1$ and each solution, i.e., $d_{1,j}, j = 1, \ldots, 7$ is $0, 0.8, 0.3, 1.9, 0.6, 1.5, 1.1$, respectively. $\boldsymbol{x}_1$ is dominated by $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, and $\boldsymbol{x}_7$. Then $\{B_{1,j}, j = 1, \ldots, 7\} = \{0, 1, 1, 0, 0, 0, 1\}$. The curve is a probability density function of normal distribution with $\sigma = 0.5$ in (16). $\{w_{1,j}, j = 2, 3, 7\} \approx \{0.22, 0.67, 0.07\}$. $c_1 = w_{1,2} + w_{1,3} + w_{1,7} \approx 0.96$. That is, c_1 is equal to the total length of the solid line segments.

2) Distance Transformation: The transformed distance $d_{i,j}^t$ between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is given as follows:

$$d_{i,j}^t = \tfrac{d_{i,j}}{1 + 0.5(c_i + c_j)}. \tag{18}$$

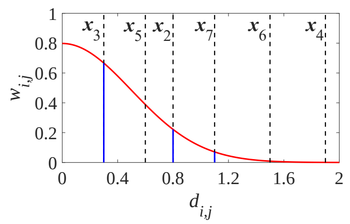


Fig. 8. Illustration of the calculation of local convergence quality in (15).

According to (18), if $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are both non-dominated solutions, $d^t_{i,j} = d_{i,j}$; otherwise, $d^t_{i,j}$ becomes smaller than $d_{i,j}$. By this transformation, distances from a solution with poor local convergence quality to other solutions become shorter. Thus the solution has a worse density value.

3) Density Estimation: Any density estimation method can be used in our framework. Here, we choose the k-nearest neighbor method [37] due to its ability to promote uniformity [38]. By the k-nearest neighbor method, the CPD value, $f_{\text{CPD}}(\boldsymbol{x}_i)$, of $\boldsymbol{x}_i$ is calculated as

$$f_{\text{CPD}}(\boldsymbol{x}_i) = \frac{1}{1+\sum_{k=1}^{K} d^t_{i,k}}, \tag{19}$$

where $d^t_{i,k}, k = 1, \ldots, K$ are the smallest K values of $d^t_{i,j}, j = 1, \ldots, N+1$. The value of f_{CPD} is in the range of $(0, 1]$. The solution with the largest value of f_{CPD} is regarded as the worst one. Such a solution is removed in the environmental selection.

D. Archive Update

We preserve only non-dominated solutions in the archive A. If the number of these solutions exceeds N, we use a double k-nearest neighbor method to evaluate their diversity quality both in the objective and decision spaces. Since only one offspring is added into A in each generation, the solution with worst diversity quality is removed.

The double k-nearest neighbor method evaluates the fitness of solution $\boldsymbol{x}_i, i = 1, \ldots, N+1$ by the following equation:

$$f_{\text{DKN}}(\boldsymbol{x}_i) = \frac{1}{1+\frac{\sum_{k=1}^{K} d^{\text{obj}}_{i,k}}{d^{\text{obj}}_{\text{mean}}}+\frac{\sum_{k=1}^{K} d^{\text{dec}}_{i,k}}{d^{\text{dec}}_{\text{mean}}}}, \tag{20}$$

where $d^{\text{obj}}_{i,k}$ ($d^{\text{dec}}_{i,k}$), $k = 1, \ldots, K$ are the Euclidean distances in the objective (decision) space between $\boldsymbol{x}_i$ and its K nearest neighbors, and $d^{\text{obj}}_{\text{mean}}$ ($d^{\text{dec}}_{\text{mean}}$) is the mean value of these distances over all solutions. That is,

$$d^{\text{obj}}_{\text{mean}} = \frac{\sum_{i=1}^{N+1}\sum_{k=1}^{K} d^{\text{obj}}_{i,k}}{(N+1)K}, \tag{21}$$

and

$$d^{\text{dec}}_{\text{mean}} = \frac{\sum_{i=1}^{N+1}\sum_{k=1}^{K} d^{\text{dec}}_{i,k}}{(N+1)K}. \tag{22}$$

The reason for using $d^{\text{obj}}_{\text{mean}}$ and $d^{\text{dec}}_{\text{mean}}$ in (20) is to normalize the objective and decision spaces into similar scales.

The value of f_{DKN} is in the range of $(0, 1]$. The larger the f_{DKN} value is, the worse the fitness of the solution is. By removing solutions with poor f_{DKN} values during the search process, we can finally obtain a set of non-dominated solutions with good diversity both in the objective and decision spaces in A.

E. Reproduction

In our CPDEA, there are two reproduction operators to choose. The major purpose of the first reproduction operator is exploration in the entire decision space, where both parents come from the population. The second reproduction operator aims at exploitation of solutions in the archive. In each generation, the first reproduction operator is used with a pre-specified probability $p \in [0, 1]$. Otherwise, the second reproduction operator is used.

In the first reproduction operator, we use the tournament selection to choose parents, where two solutions are randomly chosen from the population and compete based on their f_{CPD} values. The solution with smaller f_{CPD} value is chosen as a parent.

In the second reproduction operator, we first choose the solution with the smallest f_{DKN} value in the archive as the first parent. Then, we find the n solutions nearest to the first parent in the decision space in the population and the archive (n is the number of decision variables). We randomly choose one of these solutions as the other parent.

In the early stage of evolution, most solutions in the archive may be close to the equivalent Pareto optimal solutions which are easy to find. If we set a small value of p, i.e., give a high chance to the second reproduction operator, the efficiency of searching for other equivalent Pareto optimal solutions may be reduced. On the contrary, setting a small value of p in the latter stage of evolution may increase the efficiency of approximating all the equivalent Pareto optimal solutions. Therefore, we set $p = 1$ when the number of evaluated solutions is less than $N_E/2$ (N_E is the predefined maximum number of evaluated solutions) and $p = 0.5$ otherwise. Please refer to the supplementary material for the investigation on the effect of p.

V. Experiments

A. Settings of Experiments

In this study, we use IDMPs with 2-4 objectives in each type proposed in Section III as test problems (12 problems in total). For the two-objective IDMPs in Types 1, 2, 3, and 4, we set α to 3, 0.4, 0.4, and 4, respectively. The settings of three- and four-objective IDMPs have been given in Subsection III-C.

We choose five state-of-the-art EMMAs as comparative algorithms in this study. They are MO_Ring_PSO_SCD [21], DNEA [13], TriMOEA-TA&R [3], DN-NSGA-II [12], and Omni-optimizer [11], which have been reviewed in Subsection II-B.

The following parameter settings are adopted by all algorithms. Simulated binary crossover and polynomial mutation are used (except in MO_Ring_PSO_SCD) as the crossover and mutation operators, respectively, where the distribution index in both operators is set to 20. The crossover and mutation probabilities are 1.0 and $1/n$, respectively, where n is the number of decision variables. The termination criterion is the predefined maximum number N_E of evaluated solutions. For the two-, three-, and four-objective problems, N_E is set to $18,000$, $36,000$, and $72,000$, respectively, and the population size N is set to 60, 120, and 240, respectively.

In CPDEA, η and K are set to 2 and 3, respectively. In MO_Ring_PSO_SCD, both C_1 and C_2 are set to 2.05 and W is set to 0.7298 according to the original study [21]. In DN-NSGA-II, the crowding factor is set to a half of the population size as the authors' recommendation [12].

In TriMOEA-TA&R, p_{con}, σ_{niche}, and ϵ_{peak} are set to 0, 0.5, and 0.01, respectively. The reference vectors are generated by the systematic approach [39]. The numbers of reference vectors are 30, 36, and 56, for the two-, three- and four-objective problems, respectively. Note that for the three- and four-objective problems, these reference vectors are inverted due to the PF shape. That is, for each reference vector $\boldsymbol{r}$, we use $\boldsymbol{r}' = 1 - \boldsymbol{r}$ instead of $\boldsymbol{r}$.

In order to compare different algorithms on these test problems, we adopt IGDX [32] and the Inverted Generational Distance-Multi-modal (IGDM) [3] as the performance metrics. IGDX requires a reference solution set in the decsion space, which is typically uniformly distributed on the true PS. By measuring the distance between the reference solution set and an approximate solution set, IGDX gives a comprehensive quantification of both the convergence and the diversity in the decision space of the approximate solution set.

Different from IGDX, IGDM requires a set of reference points which is uniformly distributed on the true PF. The corresponding equivalent Pareto optimal solutions of every reference point are used as the reference solutions in the decision space. Based on these reference points and solutions, IGDM simultaneously measure the quality of an approximate solution set in the following three aspects: convergence, diversity in the objective space, and diversity in the decision space.

The smaller value of IGDX or IGDM, the better performance of the approximate solution set. In this study, for both IGDX and IGDM, $1,000$ reference solutions are sampled in each equivalent Pareto optimal subset for the two-objective IDMPs, and over $2,000$ reference solutions are sampled in each equivalent Pareto optimal subset for the three- and four-objective IDMPs.

All experimental results in this paper are obtained by executing 40 independent runs of each algorithm on each test problem. The Wilcoxon's rank sum test is employed to determine whether one algorithm shows a statistically significant difference from another, and the null hypothesis is rejected at a significant level of 0.05.

TABLE III
RESULTS OF IGDX OBTAINED BY DIFFERENT ALGORITHMS IN SOLVING THE PROPOSED IDMPS

IGDX	CPDEA		MO_Ring_ PSO_SCD		DNEA		TriMOEA-TA&R		DN-NSGA-II		Omni-Optimizer	
IDMP-M2-T1	9.9051E-4	5	5.8238E-2	3	3.1996E-1	1	3.2025E-1	0	1.7046E-1	2	3.0612E-1	2
IDMP-M2-T2	1.0418E-3	5	2.7943E-3	3	2.5280E-1	1	3.0332E-1	0	1.8643E-1	1	2.0597E-1	1
IDMP-M2-T3	1.5349E-3	5	2.7099E-3	2	8.7953E-2	3	2.5573E-1	1	2.2189E-1	0	2.8988E-1	0
IDMP-M2-T4	1.0060E-3	5	1.0455E-1	3	3.1993E-1	3	3.3693E-1	0	3.3686E-1	1	3.2156E-1	2
IDMP-M3-T1	5.6432E-3	5	6.3942E-2	4	4.0848E-1	0	3.9706E-1	0	3.0790E-1	3	3.8916E-1	0
IDMP-M3-T2	5.6551E-3	5	3.5670E-2	4	3.5963E-1	2	3.9149E-1	0	3.5466E-1	0	3.9526E-1	0
IDMP-M3-T3	6.1679E-3	5	7.9801E-3	4	2.6708E-1	2	3.4212E-1	0	2.7391E-1	2	3.3375E-1	0
IDMP-M3-T4	5.6181E-3	5	2.2894E-2	4	4.0470E-1	2	4.5465E-1	0	4.6141E-1	0	4.2578E-1	0
IDMP-M4-T1	4.3122E-3	5	3.0056E-1	4	5.9326E-1	2	5.9636E-1	1	4.9900E-1	2	5.9677E-1	0
IDMP-M4-T2	4.3525E-3	5	6.2201E-2	4	5.2027E-1	0	5.4043E-1	0	5.5508E-1	0	5.5601E-1	0
IDMP-M4-T3	4.7786E-3	5	7.5101E-3	4	4.8355E-1	0	4.9111E-1	0	3.4181E-1	2	4.3368E-1	0
IDMP-M4-T4	4.3816E-3	5	8.0168E-3	4	5.1018E-1	0	5.3027E-1	0	5.2769E-1	0	5.2579E-1	0
Average Performance Score	5.000		3.583		1.333		0.167		1.083		0.417	
+ / - / =	0 / 0 / 12		12 / 0 / 0		12 / 0 / 0		12 / 0 / 0		12 / 0 / 0		12 / 0 / 0	

TABLE IV
RESULTS OF IGDM OBTAINED BY DIFFERENT ALGORITHMS IN SOLVING THE PROPOSED IDMPS

IGDM	CPDEA		MO_Ring_ PSO_SCD		DNEA		TriMOEA-TA&R		DN-NSGA-II		Omni-Optimizer	
IDMP-M2-T1	1.4124E-2	5	1.8282E-1	3	4.7859E-1	1	5.0449E-1	0	2.8130E-1	1	4.7400E-1	2
IDMP-M2-T2	1.4143E-2	5	3.7313E-2	3	3.8052E-1	1	4.6902E-1	0	2.9222E-1	1	3.4694E-1	1
IDMP-M2-T3	1.8307E-2	5	3.2934E-2	2	1.6449E-1	3	5.0474E-1	0	3.3441E-1	1	4.5614E-1	1
IDMP-M2-T4	1.4381E-2	5	2.0126E-1	3	4.7857E-1	3	5.0462E-1	1	5.0501E-1	0	4.8988E-1	2
IDMP-M3-T1	8.0628E-2	5	2.4865E-1	4	6.4322E-1	2	6.7732E-1	0	5.8313E-1	0	6.5474E-1	0
IDMP-M3-T2	8.0222E-2	5	1.9580E-1	4	5.9635E-1	3	6.3067E-1	1	6.1654E-1	0	6.6028E-1	0
IDMP-M3-T3	8.4936E-2	5	1.1097E-1	4	4.7477E-1	2	5.6472E-1	1	5.1626E-1	1	6.0902E-1	0
IDMP-M3-T4	7.9835E-2	5	1.7693E-1	4	6.4289E-1	3	7.1381E-1	1	7.4751E-1	0	7.0096E-1	2
IDMP-M4-T1	6.2817E-2	5	5.5617E-1	4	7.5730E-1	2	7.6509E-1	0	7.1160E-1	0	7.6067E-1	1
IDMP-M4-T2	6.2369E-2	5	2.5369E-1	4	7.0056E-1	3	7.6503E-1	1	7.7012E-1	0	7.6085E-1	2
IDMP-M4-T3	6.4847E-2	5	1.0586E-1	4	6.9866E-1	1	7.1990E-1	0	6.1194E-1	0	6.9467E-1	0
IDMP-M4-T4	6.2214E-2	5	1.1847E-1	4	7.2776E-1	2	7.4123E-1	1	7.5415E-1	0	7.3205E-1	1
Average Performance Score	5.000		3.583		2.167		0.500		0.333		1.000	
+ / - / =	0 / 0 / 12		12 / 0 / 0		12 / 0 / 0		12 / 0 / 0		12 / 0 / 0		12 / 0 / 0	

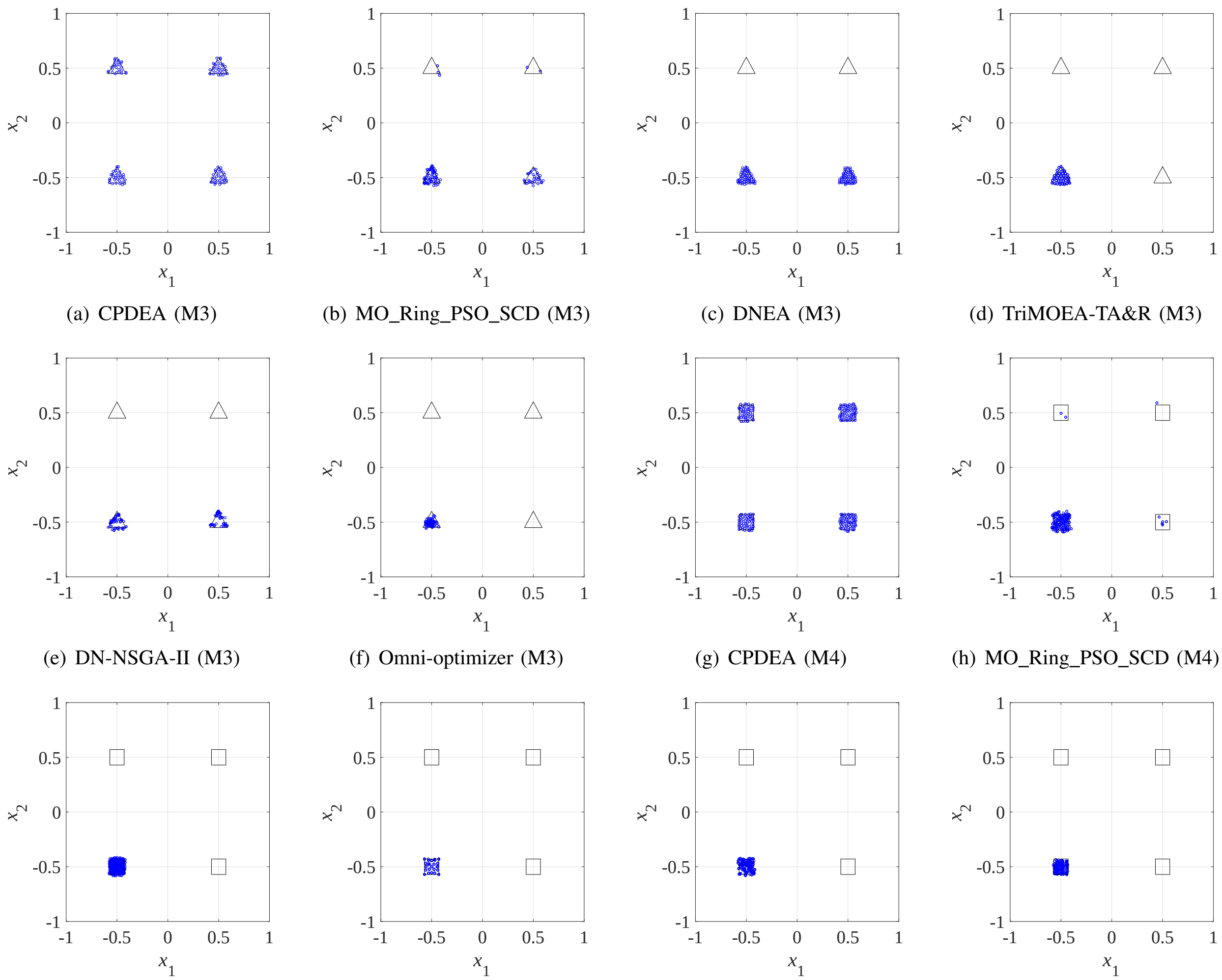
B. Comparison Results

In this subsection, we applied CPDEA, MO_Ring_PSO_ SCD, DNEA, TriMOEA-TA&R, DN-NSGA-II, and Omni-optimizer to the 12 IDMPs. The average IGDX (IGDM) value and the corresponding performance score [40] of each algorithm on each problem are given in Table III (IV). A darker tone corresponds to a larger performance score (i.e., better algorithm). For each test problem, the performance score of an algorithm is the number of the comparative algorithms which perform significantly worse than it according to IGDX (IGDM). We also give the average performance score of each algorithm over all the test problems. At the bottom of Tables III (IV), '+' or '-' mean the number of test problems where CPDEA shows significantly better or worse performance than the algorithm. '=' means that the number of test problems where there exists no significance between the results of CPDEA and the algorithm. Note that the i-th type of IDMP with j objectives is denoted as IDMP-Mj-Ti.

Fig. 9 shows the solutions obtained by these algorithms on IDMP-M3-T1 and IDMP-M4-T1 in a given single run to visually investigate their performance. M3 and M4 in the titles of the sub-figures indicate the results on IDMP-M3-T1 and IDMP-M4-T1, respectively. This particular run is associated with the result which is the closest to the mean IGDX value in Table III. Each polygon (triangle or rectangle) is an equivalent Pareto optimal subset. The difficulty in finding solutions in the polygons on the left-bottom, right-bottom, right-top, and left-top increases gradually.

We can see from Tables III and IV that the proposed CPDEA achieved the highest average performance scores based on both IGDX and IGDM. It significantly outperforms every other algorithm on every test problem. From Fig. 9, we can see that CPDEA can find a set of well distributed solutions in every equivalent Pareto optimal subset (polygon), whereas the other algorithms cannot.

It is interesting to observe from Tables III and IV that MO_Ring_PSO_SCD generally performs better than the other algorithms except for CPDEA. The reason for the relatively good performance of MO_Ring_PSO_SCD is that it inherently evaluates a solution's local convergence quality. In MO_Ring_PSO_SCD, both the personal and neighbor best



(a) CPDEA (M3) (b) MO_Ring_PSO_SCD (M3) (c) DNEA (M3) (d) TriMOEA-TA&R (M3)

(e) DN-NSGA-II (M3) (f) Omni-optimizer (M3) (g) CPDEA (M4) (h) MO_Ring_PSO_SCD (M4)

(i) DNEA (M4) (j) TriMOEA-TA&R (M4) (k) DN-NSGA-II (M4) (l) Omni-optimizer (M4)

Fig. 9. The solutions obtained by different algorithms on IDMP-M3-T1 and IDMP-M4-T1 in a single run. This particular run is associated with the result which is the closest to the mean IGDX value in Table III. Each polygon (triangle or rectangle) is an equivalent Pareto optimal subset. The difficulty in finding solutions in the polygons on the left-bottom, right-bottom, right-top, and left-top increases gradually.

values of each particle (solution) are stored in its own archives (personal best archive and neighbor best archive). The primary selection criterion in the archives is the Pareto rank. The personal and neighbor best archives are updated based on the particle's and its neighbor's historical values, respectively, not the historical values of all particles. That is, the primary selection criteria is actually the local Pareto rank. However, this does not imply that MO_Ring_PSO_SCD is capable to solve MMOP-ICDs. In fact, solutions in MO_Ring_PSO_SCD stuck in their local regions. Thus MO_Ring_PSO_SCD cannot promote diversity around the equivalent Pareto optimal solutions which are difficult to find. This is verified from Fig. 9 (b) and (h). We can see that although MO_Ring_PSO_SCD obtained solutions close to every equivalent Pareto optimal subset, the solutions are much less around the equivalent Pareto optimal subsets which are difficult to find.

DNEA achieved the third best average performance scores based on both IGDX and IGDM from Tables III and IV. The average performance score based on IGDM is higher than that based IGDX. This may be due to the normalization ability of the double-sharing function, which leads to a good diversity in the objective space. TriMOEA-TA&R, DN-NSGA-II, and Omni-optimizer generally perform poorly, especially TriMOEA-TA&R. This is attributed to its very good convergence ability, which makes it quickly converge toward the equivalent Pareto optimal subset which is the easiest to find. We can see from Fig. 9 that DNEA, TriMOEA-TA&R, DN-NSGA-II, and Omni-optimizer can only find solutions in one or two polygons.

Note that we provide further investigations on the behavior of CPDEA in the supplementary material. Readers can refer to them if interested.

VI. Conclusion

In this paper, we have proposed a set of imbalance distance minimization problems, termed IDMPs, and a novel evolutionary algorithm using a convergence-penalized density method, termed CPDEA.

There exist two causes for the imbalance between convergence and diversity in the decision space in IDMPs. We have proposed four types of IDMPs according to these two causes. In each type of IDMPs, the difficulty in finding each equivalent Pareto optimal subset is controllable, and the numbers of objectives, decision variables, and equivalent Pareto optimal subsets are scalable. We have formulated 12 IDMPs with different numbers of objectives and decision variables for benchmarking.

In CPDEA, we have proposed a convergence-penalized density method for the environmental selection in the population. The distance between each pair of solutions is transformed based on their local convergence quality. The density value of a solution is penalized due to the transformed distances. Besides the population, an archive is used to store non-dominated solutions during the evolution. We have proposed a double k-nearest neighbor method to select solutions with good diversity in both the objective and decision spaces in the archive. In addition, we have proposed two reproduction operators for exploration and exploitation, respectively.

We have compared CPDEA with MO_Ring_PSO_SCD, DNEA, TriMOEA-TA&R, DN-NSGA-II, and Omni-optimizer by applying them to the proposed IDMPs. The experimental results showed that CPDEA is obviously the best among these algorithms. We have also investigated the effects of the two parameters, η and p on the behavior of CPDEA in the supplementary material. The robust performance of CPDEA was observed in a wide range of η. p is suggest to be set to a large value at the early stage of evolution but smaller later.

One future research direction is to develop other types of imbalanced MMOPs. Another research is to design a strategy to adaptively tune p in CPDEA.

* The codes of IDMP and CPDEA (implemented on PlatEMO [41]) are available on https://github.com/yiping0liu.

REFERENCES

[1] Y. Han, D. Gong, Y. Jin, and Q. Pan, "Evolutionary multiobjective blocking lot-streaming flow shop scheduling with machine breakdowns," *IEEE Transactions on Cybernetics*, no. 99, pp. 1–14, 2017.

[2] R. Qi and G. G. Yen, "Hybrid bi-objective portfolio optimization with pre-selection strategy," *Information Sciences*, vol. 417, pp. 401–419, 2017.

[3] Y. Liu, G. G. Yen, and D. Gong, "A multi-modal multi-objective evolutionary algorithm using two-archive and recombination strategies," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 4, pp. 660–674, 2019.

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

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

[6] K. Li, S. Kwong, J. Cao, M. Li, J. Zheng, and R. Shen, "Achieving balance between proximity and diversity in multi-objective evolutionary algorithm," *Information Sciences*, vol. 182, no. 1, pp. 220–242, 2012.

[7] Y. Yuan, H. Xu, B. Wang, B. Zhang, and X. Yao, "Balancing convergence and diversity in decomposition-based many-objective optimizers," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 2, pp. 180–198, 2016.

[8] Y. Liu, D. Gong, X. Sun, and Y. Zhang, "Many-objective evolutionary optimization based on reference points," *Applied Soft Computing*, vol. 50, no. 1, pp. 344–355, 2017.

[9] Y. Liu, D. Gong, J. Sun, and Y. Jin, "A many-objective evolutionary algorithm using a one-by-one selection strategy," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2689–2702, 2017.

[10] H.-L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 450–455, 2014.

[11] K. Deb and S. Tiwari, "Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1062–1087, 2008.

[12] J. Liang, C. Yue, and B. Qu, "Multimodal multi-objective optimization: A preliminary study," in *IEEE Congress on Evolutionary Computation*. IEEE, 2016, pp. 2454–2461.

[13] Y. Liu, H. Ishibuchi, Y. Nojima, N. Masuyama, and K. Shang, "A double-niched evolutionary algorithm and its behavior on polygon-based problems," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2018, pp. 262–273.

[14] R. Tanabe and H. Ishibuchi, "A decomposition-based evolutionary algorithm for multi-modal multi-objective optimization," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2018, pp. 249–261.

[15] X. Li, M. G. Epitropakis, K. Deb, and A. Engelbrecht, "Seeking multiple solutions: An updated survey on niching methods and their applications," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 4, pp. 518–538, 2017.

[16] H.-L. Liu, L. Chen, K. Deb, and E. D. Goodman, "Investigating the effect of imbalance between convergence and diversity in evolutionary multiobjective algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 3, pp. 408–425, 2017.

[17] F. Kudo, T. Yoshikawa, and T. Furuhashi, "A study on analysis of design variables in Pareto solutions for conceptual design optimization problem of hybrid rocket engine," in *IEEE Congress on Evolutionary Computation*. IEEE, 2011, pp. 2558–2562.

[18] O. Schütze, M. Vasile, and C. A. Coello Coello, "Computing the set of epsilon-efficient solutions in multiobjective space mission design," *Journal of Aerospace Computing, Information, and Communication*, vol. 8, no. 3, pp. 53–70, 2011.

[19] G. Rudolph, B. Naujoks, and M. Preuss, "Capabilities of EMOA to detect and preserve equivalent pareto subsets," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2007, pp. 36–50.

[20] M. Preuss, B. Naujoks, and G. Rudolph, "Pareto set and EMOA behavior for simple multimodal multiobjective functions," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2006, pp. 513–522.

[21] C. Yue, B. Qu, and J. Liang, "A multiobjective particle swarm optimizer using ring topology for solving multimodal multiobjective problems," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 5, pp. 805–817, 2018.

[22] C. Yue, B. Qu, K. Yu, J. Liang, and X. Li, "A novel scalable test problem suite for multimodal multiobjective optimization," *Swarm and Evolutionary Computation*, vol. 48, pp. 62–71, 2019.

[23] M. Köppen and K. Yoshida, "Substitute distance assignments in NSGA-II for handling many-objective optimization problems," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2007, pp. 727–741.

[24] H. Ishibuchi, Y. Hitotsuyanagi, N. Tsukamoto, and Y. Nojima, "Many-objective test problems to visually examine the behavior of multiobjective evolution in a decision space," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2010, pp. 91–100.

[25] Y. Nojima, T. Fukase, Y. Liu, N. Masuyama, and H. Ishibuchi, "Constrained multiobjective distance minimization problems," in *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 2019, pp. 586–594.

[26] O. M. Shir, M. Preuss, B. Naujoks, and M. T. Emmerich, "Enhancing decision space diversity in evolutionary multiobjective algorithms." in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2009, pp. 95–109.

[27] T. Ulrich, J. Bader, and L. Thiele, "Defining and optimizing indicator-based diversity measures in multiobjective search," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2010, pp. 707–717.

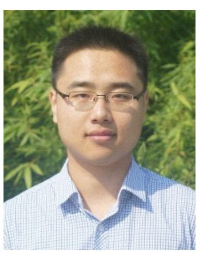
[28] X. Li, "Niching without niching parameters: Particle swarm optimization using a ring topology," *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 1, pp. 150–169, 2010.

[29] J. Liang, W. Xu, C. Yue, K. Yu, H. Song, O. D. Crisalle, and B. Qu, "Multimodal multiobjective optimization with differential evolution," *Swarm and Evolutionary Computation*, vol. 44, pp. 1028–1059, 2019.

[30] J. Horn, N. Nafpliotis, and D. E. Goldberg, "A niched Pareto genetic algorithm for multiobjective optimization," in *Proceedings of IEEE World Congress on Computational Intelligence*. IEEE, 1994, pp. 82–87.

[31] R. Cheng, M. Li, K. Li, and X. Yao, "Evolutionary multiobjective optimization-based multimodal optimization: Fitness landscape approx-

imation and peak detection," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 5, pp. 692–706, 2018.
[32] A. Zhou, Q. Zhang, and Y. Jin, "Approximating the set of Pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 5, pp. 1167–1189, 2009.
[33] O. Shoval, H. Sheftel, G. Shinar, Y. Hart, O. Ramote, A. Mayo, E. Dekel, K. Kavanagh, and U. Alon, "Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space," *Science*, vol. 336, no. 6085, pp. 1157–1160, 2012.
[34] H. Ishibuchi, N. Akedo, and Y. Nojima, "A many-objective test problem for visually examining diversity maintenance behavior in a decision space," in *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 2011, pp. 649–656.
[35] G. Syswerda, "Uniform crossover in genetic algorithms," in *Proceedings of the third international conference on Genetic algorithms*. Morgan Kaufmann Publishers, 1989, pp. 2–9.
[36] M. Lozano, F. Herrera, and J. R. Cano, "Replacement strategies to preserve useful diversity in steady-state genetic algorithms," *Information Sciences*, vol. 178, no. 23, pp. 4421–4433, 2008.
[37] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: improving the strength Pareto evolutionary algorithm," Eidgenössische Technische Hochschule Zürich (ETH), Institut für Technische Informatik und Kommunikationsnetze (TIK), Tech. Rep. 103, 2001.
[38] M. Li, S. Yang, and X. Liu, "Shift-based density estimation for Pareto-based algorithms in many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 348–365, 2014.
[39] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.
[40] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evolutionary Computation*, vol. 19, no. 1, pp. 45–76, 2011.
[41] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization," *IEEE Computational Intelligence Magazine*, vol. 12, no. 4, pp. 73–87, 2017.



Yiping Liu (M'18) received the B.Eng. degree in electrical engineering and automation and the Ph.D. degree in control theory and control engineering from China University of Mining and Technology, Xuzhou, China in 2012 and 2017, respectively.

During 2016-2017, he was a visiting scholar in the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK, USA. He is currently a Research Assistant Professor in the Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Sakai, Japan. His research interests include evolutionary computation, multi-objective optimization, and machine learning.

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

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

Gary G. Yen (S'87-M'88-SM'97-F'09) received the Ph.D. degree in electrical and computer engineering from the University of Notre Dame, Notre Dame, IN, USA, in 1992.

He was with the Structure Control Division, U.S. Air Force Research Laboratory, Albuquerque, NM, USA. He joined Oklahoma State University (OSU), Stillwater, OK, USA, in 1997, where he is currently a Regents Professor with the School of Electrical and Computer Engineering. His current research interests include intelligent control, computational intelligence, conditional health monitoring, signal processing, and their industrial/defense applications.

Dr. Yen received the Andrew P Sage Best Transactions Paper Award from the IEEE Systems, Man and Cybernetics Society in 2011, and the Meritorious Service Award from the IEEE Computational Intelligence Society in 2014. He served as the General Chair for the 2003 IEEE International Symposium on Intelligent Control held in Houston, TX, USA, and the 2006 IEEE World Congress on Computational Intelligence held in Vancouver, BC, Canada. He served as the Vice President for the Technical Activities from 2005 to 2006 and then President from 2010 to 2011 of the IEEE Computational Intelligence Society. He was an Associate Editor of the *IEEE Control Systems Magazine*, the *IEEE Transactions on Control Systems Technology*, *Automatica*, *Mechatronics*, the *IEEE Transactions on Systems, Man and Cybernetics, Parts A and B*, and the *IEEE Transactions on Neural Networks*. He was the Founding Editor-in-Chief of the *IEEE Computational Intelligence Magazine* from 2006 to 2009. He is currently serving as an Associate Editor for the *IEEE Transactions on Evolutionary Computation* and the *IEEE Transactions on Cybernetics*.

Yusuke Nojima (S'99–M'04) received the B.S. and M.S. Degrees in mechanical engineering from Osaka Institute of Technology, Osaka, Japan, in 1999 and 2001, respectively, and the Ph.D. degree in system function science from Kobe University, Hyogo, Japan, in 2004.

Since 2004, he has been with Osaka Prefecture University, Sakai, Japan, where he was a Research Associate and is currently an Associate Professor in the Department of Computer Science and Intelligent Systems. His research interests include evolutionary machine learning and evolutionary multi-objective optimization.

Naoki Masuyama (S'12–M'16) received the degree from Nihon University, Funabashi, Japan, in 2010, the M.E. degree from Tokyo Metropolitan University, Hino, Japan in 2012, and the Ph.D. degree from Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia in 2016.

He is currently an Assistant Professor with the Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Sakai, Japan. His research interests include data mining and machine learning.