

Article

# Search Acceleration of Evolutionary Multi-Objective Optimization Using an Estimated Convergence Point

Yan Pei <sup>1,\*</sup> , Jun Yu <sup>2</sup>  and Hideyuki Takagi <sup>3</sup> <sup>1</sup> Computer Science Division, University of Aizu, Aizuwakamatsu 965-8580, Japan<sup>2</sup> Graduate School of Design, Kyushu University, Fukuoka 815-8540, Japan; yujun@kyudai.jp<sup>3</sup> Faculty of Design, Kyushu University, Fukuoka 815-8540, Japan; takagi@design.kyushu-u.ac.jp

\* Correspondence: peiyan@u-aizu.ac.jp; Tel.: +81-242-37-2765

Received: 21 November 2018; Accepted: 23 January 2019; Published: 28 January 2019



**Abstract:** We propose a method to accelerate evolutionary multi-objective optimization (EMO) search using an estimated convergence point. Pareto improvement from the last generation to the current generation supports information of promising Pareto solution areas in both an objective space and a parameter space. We use this information to construct a set of moving vectors and estimate a non-dominated Pareto point from these moving vectors. In this work, we attempt to use different methods for constructing moving vectors, and use the convergence point estimated by using the moving vectors to accelerate EMO search. From our evaluation results, we found that the landscape of Pareto improvement has a uni-modal distribution characteristic in an objective space, and has a multi-modal distribution characteristic in a parameter space. Our proposed method can enhance EMO search when the landscape of Pareto improvement has a uni-modal distribution characteristic in a parameter space, and by chance also does that when landscape of Pareto improvement has a multi-modal distribution characteristic in a parameter space. The proposed methods can not only obtain more Pareto solutions compared with the conventional non-dominant sorting genetic algorithm (NSGA)-II algorithm, but can also increase the diversity of Pareto solutions. This indicates that our proposed method can enhance the search capability of EMO in both Pareto dominance and solution diversity. We also found that the method of constructing moving vectors is a primary issue for the success of our proposed method. We analyze and discuss this method with several evaluation metrics and statistical tests. The proposed method has potential to enhance EMO embedding deterministic learning methods in stochastic optimization algorithms.

**Keywords:** evolutionary multi-objective optimization; convergence point; acceleration search; evolutionary computation; optimization

## 1. Introduction

In the research area of optimization, there are single objective optimization problems and multi-objective optimization problems. The difference between these two categories of optimization problems lies in the number of fitness functions. The single objective optimization attempts to obtain only one optimal solution in one parameter space, i.e., one fitness landscape. The multi-objective optimization tries to satisfy more than one optimal condition or target, i.e., more than one fitness landscape. Usually, these optimal conditions in multi-objective optimization conflict with each other, and cannot be combined into one optimal condition. Single objective optimization and multi-objective optimization have different search targets because of the requirements of algorithm design. One tries to obtain a better optimum, the other seeks to obtain more non-dominated solutions on Pareto front.

Most multi-objective optimization research pays attention to the diversity and the number of non-dominated Pareto solutions. Evolutionary multi-objective optimization (EMO) algorithms are

efficient and effective when handling multi-objective optimization problems. The EMO keeps the multiple objective functions independently and uses Pareto-based ranking schemes to maintain feasible solutions. However, the deterministic programming methods use a scalarization method that needs to transfer multiple objectives into one objective. State-of-the-art studies on EMO concentrate on Pareto dominance handling in an objective space, which is tied to generating solutions to approximate the Pareto solution frontiers [1]. The primary disadvantages of EMO are its optimization capability and the non-guarantee of Pareto optimality, which cannot be perfectly solved by Pareto dominance studies in EMO.

A fitness approximation method is widely used in the evolutionary computation (EC) community to reduce the computational cost of fitness evaluations and is expected to estimate the global optimum solution area. Reference [2] investigated several approximation methods, such as polynomial models, kriging models, neural networks, and others. Reference [3] proposed a framework to manage approximation models in EC search. Reference [4] made a survey on the advances of approximating a fitness function in EC algorithms and presented some future research challenges. Inspired by the scale-spacing method [5], a uni-modal function was used to approximate a fitness function for estimating the peak point, and the EC algorithm used it as an elite individual to increase convergence speed [6]. The same method was extended into a dimension reduction space for fitness landscape approximation to enhance EC algorithms [7,8]. A fitness approximation mechanism was introduced into genetic algorithms to obtain the optimal solutions satisfactorily and quickly, and can reduce computational cost at the same time [9]. It was a unique approach to filter frequency components for approximating a fitness landscape [10], which uniformly samples in parameter space firstly, and the re-sampled EC individuals are used to obtain frequency information using the discrete Fourier transform. There are many methods to accelerate EC convergence using a fitness approximate model [11–17], and many potential subjects need to be studied further [18,19]. Reference [20] presents a comprehensive survey on the fitness approximation methods in interactive evolutionary computation (IEC). Those methods not only can enhance IEC search, but also can enhance conventional EC search [21].

The estimation method presents a novel perspective using mathematical approaches to calculate a convergence point (Convergence means of modeling the tendency for generic characteristics of populations to stabilize over time, in EC, convergence point means the global optimum in single objective problems ideally.) of a population. This is a great solution that enhances search of a stochastic optimization algorithm using a deterministic method embedded into a stochastic optimization process. Using an estimated point to accelerate EC search is one such method that implements this research philosophy [22]. The individual moving from one generation to the next supports convergence information of EC search condition. We used such information to mathematically estimate a convergence point as an elite individual to enhance single objective optimization search [23]. A clustering method was developed for bipolar tasks (Bipolar tasks mean there are two peaks in a fitness landscape in a problem, e.g., combination of two Gaussian functions  $N(\mu_1, \sigma_1) + N(\mu_2, \sigma_2)$  presents two peaks in the landscape.), which proposed four improvements to increase the accuracy of estimated convergence points and was applied to a multi-modal optimization problem [24]. We attempted to combine EC algorithms with the estimation framework for bipolar tasks and analyzed the effect of proposed four improvements. From our previous studies, we found that this estimation method is effective in single objective optimization [23].

Pareto improvement information from the current generation to the next supports the promising search areas of Pareto frontier solutions in both an objective space and a parameter space. In this paper, we extend the method of estimating a convergence point into the EMO algorithm to find potential non-dominated solution areas in search spaces using the estimation information from the objective spaces. By putting an estimated point into an EMO search and deleting a dominated solution, the EMO algorithm can find more non-dominated solutions in early generations, which is a factor that motivates this study. We use the NSGA-II algorithm to evaluate our study hypothesis and verify the performance

of the framework that combines the estimation method with EMO, and attempt to enhance EMO search. This demonstrates one of original aspects of this work. We undertake a comparative study between the NSGA-II algorithm with and without the estimation method using multi-objective optimization benchmark problems and statistical tests. The advantages and disadvantages of proposed method are presented, analyzed, and discussed using experimental evaluation results.

Following this introductory section, we introduce a variety of EMO techniques and algorithms in the Section 2. We make a brief review on the estimation algorithm for a uni-modal optimization problem in the Section 3. In the Section 4, we explain how to extend such estimation methods to accelerate EMO search. There are three primary steps in this framework. The first is finding the pair information of moving vectors in objective space of EMO, the second is the estimation of a convergence point in a parameter space, and the third is the insertion of an estimated point by deleting a dominated solution to enhance EMO search. In the Section 5, we evaluate our method using the NSGA-II algorithm and some multi-objective benchmark problems, and analyze its optimization performance and characteristics of population distribution in both an objective space and a parameter space. We discuss and analyze the evaluation results using statistical tests in the Section 6. The results demonstrate that the proposed method can obtain more non-dominated Pareto solutions early. Finally, we make a conclusion of the whole work, and present some open research topics and future work in the Section 7.

## 2. Evolutionary Multi-Objective Optimization

Multi-objective optimization problems lie in many real-world applications and they contain multiple optimization objectives that conflict with each other. This makes conventional optimization algorithms (deterministic optimization method), e.g., linear programming method [25] and Newton-Raphson method [26], difficult to apply when solving these problems. One solution of the multi-objective optimization problems is to transform multiple objectives into a single objective by assigning different weights to each objective for a combination. This requires us to have a degree of deep understanding of multi-objective optimization problems.

Currently, more popular approaches use evolutionary multi-objective optimization algorithms because of various well-defined features and characteristics, such as strong robustness, ease of use, intelligence, and others. Almost all of these pay attention to finding a set of trade-off optimal solutions (known as Pareto optimal solutions), instead of a single optimal solution. The Pareto dominance and diversity of solutions are two primary subjects in EMO research. One attempts to obtain many non-dominated Pareto solutions, and the other tries to obtain Pareto solutions in a wide area on Pareto solution front. Here, we make a brief review of several techniques, strategies, and algorithms that solve the problems of EMO with regards to these two aspects.

### 2.1. Non-dominated Sorting Method

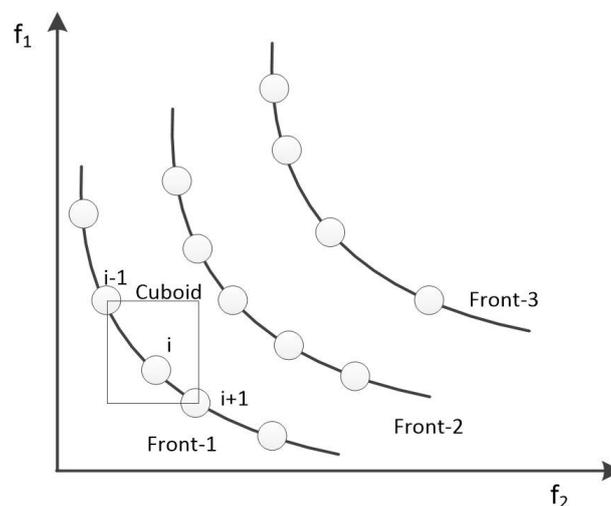
Non-dominated sorting is an elite mechanism for building a new generation of EMO algorithms for handling the non-dominated Pareto solutions. It is one of EMO selection strategies. The main motivation of this method is to find the non-dominated solutions by pairwise comparisons of all individuals. Here, a non-dominated individual means that there are no other individuals, whose all objectives are better than this one. The basic and formal calculation process of non-dominant sorting method can be implemented as the following steps.

1. Getting the first individual as a current individual;
2. Comparing all objectives of the current individual with those of all other individuals;
3. Counting the domination count  $N_p$ , which means the number of individuals that dominant the current individual;
4. Setting the individuals satisfy  $N_p = 0$  as the first front, and remove these individuals from the generation temporarily;
5. Repeating the above process until every individual is processed.

### 2.2. Crowding Distance Techniques

The Pareto solution diversity issue is also an important indicator of measurement in EMO algorithms. If the solution diversity is insufficient, it is easy to lead the Pareto optimal solutions not to be covered. Many methods attempt to maintain the solution diversity as much as possible in EMO algorithms. One of these methods uses a sharing parameter to keep the diversity during EMO search. However, it requires the preset and optimization of the parameter, and EMO performance depends on the setting. Crowding distance technique is another solution for handling this predefined parameter problems, which is calculated using a set of individuals.

The primary motivation of a crowding distance is to measure an individual density by distances. There is an aggregation level of the adjacent individuals in parameter space. Figure 1 presents a two-dimensional example of crowding distance calculation, where  $f_1$  and  $f_2$  are two objectives. The crowding distance can be calculated by averaging the length and width of the cuboid (marked by dash line). Averaging of the length and width of the cuboid (marked by dash line) is used to calculate the crowding distance for each individual.



**Figure 1.** Two-dimensional example of crowding distance [27], the crowding distance is calculated within one Pareto front, its calculation method is  $crowding\_distance = \sum_{o=1}^P \frac{obj_{o(i+1)} - obj_{o(i-1)}}{obj_{o,max} - obj_{o,min}}$ , where  $o$  is the index of the number of objective functions,  $P$  is the number of objective functions,  $i$  is the index of individual, and  $obj$  is the value of an objective function.

Any EMO algorithms sort all individuals according to their first objective. We calculate the difference, i.e.,  $obj_{o(i+1)} - obj_{o(i-1)}$ , between two adjacent individuals for all individuals. The first and last individual are set to infinite. This method calculates all crowding distances using multiple objectives. Finally, the final crowding distance of each individual is set to the mean of its crowding distances in all objectives.

### 2.3. NSGA and Its Variants

The NSGA is the first generation of EMO algorithm that uses non-dominated sorting techniques to find multiple Pareto optimal solutions with a single simulation running [28]. However, it has still suffered from several criticisms, including those relating to its high computational cost, lack of elitism, and the requirement for the setting of sharing parameter. Subsequently, its improved version, NSGA-II, was proposed to overcome all the above limitations at once by introducing fast non-dominated sorting and a tournament selection using a crowding distance to reduce computational complexity [27]. It has become one of the most popular EMO algorithms that are used to solve problems of multi-objective optimization. Recently, a more powerful version, NSGA-III [29], was also proposed, where a clustering operator replaces the crowding distance operator in NSGA-II to solve many-objective optimization

problems. Actually, there are also many other EMO algorithms based on non-dominated sorting and have achieved satisfactory results, such as MOGA [30], NPGA [31], SPEA [32], SPEA2 [33], PESA [34], PESA-II [35], multi-objective chaotic evolution [36], etc.

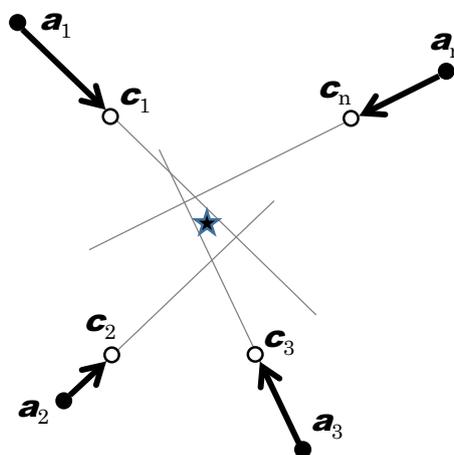
Although various EMO algorithms have been proposed and have achieved outstanding results, most of them only focus on the study in an objective space. We therefore try to use moving vectors as a bridge between a parameter space and an objective space to analyze the landscapes of the two spaces. The motivation of this study promotes to use a mathematical method to estimate a convergence point in a parameter space using these moving vectors' information from its objective space. We expect this research to provoke the EMO researchers' attention towards the parameter space and encourage them to notice the connection between the two spaces for designing better EMO algorithms.

### 3. Estimating a Convergence Point

#### 3.1. Notation Definitions

Before we explain how to estimate a convergence point from moving vectors, we offer some notations for better understanding of this section in advance. When an EC algorithm searches in a  $d$ -dimensional parameter space with  $n$  individuals ( $n, d \in \mathbb{Z}^+$ ), we denote the  $i$ -th individual in the current generation, a corresponding relative individual in the next generation, and their moving vector to be  $\mathbf{a}_i$ ,  $\mathbf{c}_i$ , and  $\mathbf{b}_i = \mathbf{c}_i - \mathbf{a}_i$ , respectively,  $\{\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i \in \mathbb{R}^d; i = 1, 2, \dots, n\}$  (See Figure 2). The unit vector of  $\mathbf{b}_i$  is defined as  $\mathbf{b}'_i = \mathbf{b}_i / \|\mathbf{b}_i\|$  ( $\mathbf{b}'_i{}^T \mathbf{b}_i = 1$ ). There are  $n$  moving vectors, and  $\mathbf{a}_i$  is a starting point of  $\mathbf{b}_i$ .

We denote  $\mathbf{x} \in \mathbb{R}^d$  as the estimated convergence point that has the minimal distance to the lines made by extending the line segments  $\mathbf{b}'_i$ . This point,  $\mathbf{x}$ , has a higher possibility to locate near the optimal solution in EC optimization problems. The  $\mathbf{x}$  is indicated by the  $\star$  mark in the Figure 2. We will explain how to obtain the  $\mathbf{x}$  point by a deterministic mathematical method. In this work, all vectors are presented as column vectors.



**Figure 2.** The convergence point ( $\star$ ) can be estimated by the moving vectors ( $\mathbf{b}_i$ ) between individuals ( $\mathbf{a}_i, i = 1, 2, \dots, n$ ) in the  $k$ -th generation and their offspring ( $\mathbf{c}_i, i = 1, 2, \dots, n$ ) in the  $(k + 1)$ -th generation.

#### 3.2. Estimation Method of a Point from Moving Vectors

This section is primarily adopted from our previous work in reference [22]. From the principle of the law of large numbers, the estimated convergent point is the nearest one to the extension lines of these moving vectors. The lines of moving vectors can be expressed as  $\mathbf{a}_i + t_i \mathbf{b}'_i, t_i \in \mathbb{R}$ . The nearest

point that is close to these extension lines can be obtained by solving the optimization problem shown in Equation (1).

$$\mathbf{x} = \min_{\mathbf{x}, \{t_i\}} J(\mathbf{x}, \{t_i\}) = \min_{\mathbf{x}, \{t_i\}} \sum_{i=1}^n \|\mathbf{a}_i + t_i \mathbf{b}'_i - \mathbf{x}\|^2 \tag{1}$$

The shortest line segment from the estimated convergence point  $\mathbf{x}$  to the extended moving vector  $\mathbf{a}_i + t_i \mathbf{b}'_i$ ,  $t_i \in \mathbb{R}$  is  $\mathbf{a}_i + t_i \mathbf{b}'_i - \mathbf{x}$ ,  $t_i \in \mathbb{R}$ , and the relation of this line segment and the moving vector  $\mathbf{b}$  or its unit vector  $\mathbf{b}'$  is orthogonal. Equation (2) presents this orthogonal condition.

$$\mathbf{b}'_i{}^T (\mathbf{a}_i + t_i \mathbf{b}'_i - \mathbf{x}) = 0 \quad (\text{orthogonal condition}) \tag{2}$$

From this orthogonal condition, we can use  $\mathbf{x}$  to express  $t_i$ , i.e.,  $t_i = \frac{(\mathbf{b}'_i)^T (\mathbf{x} - \mathbf{a}_i)}{\|\mathbf{b}'_i\|^2}$  and introduce it to Equation (1) to reduce the number of optimization parameter. The derivation process is presented in Equation (3), where  $I_d$  is an identity matrix, and  $H_i = \mathbf{b}'_i \mathbf{b}'_i{}^T - I_d$ .

$$\begin{aligned} \mathbf{x} &= \min_{\mathbf{x}, \{t_i\}} \sum_{i=1}^n \|\mathbf{a}_i + t_i \mathbf{b}'_i - \mathbf{x}\|^2 \\ &= \min_{\mathbf{x}} \sum_{i=1}^n \left\| \mathbf{b}'_i \frac{(\mathbf{b}'_i)^T (\mathbf{x} - \mathbf{a}_i)}{\|\mathbf{b}'_i\|^2} - (\mathbf{x} - \mathbf{a}_i) \right\|^2 \\ &= \min_{\mathbf{x}} \sum_{i=1}^n \left\| \left\{ \frac{\mathbf{b}'_i \mathbf{b}'_i{}^T}{\|\mathbf{b}'_i\|^2} - I_d \right\} (\mathbf{x} - \mathbf{a}_i) \right\|^2 \\ &= \min_{\mathbf{x}} \sum_{i=1}^n \|H_i (\mathbf{x} - \mathbf{a}_i)\|^2 \end{aligned} \tag{3}$$

Next, we can obtain the following objective function (Equation (4)) from Equation (1) where we have eliminated the term  $\{t_i\}$ .

$$J(\mathbf{x}) = \sum_{i=1}^n (\mathbf{x} - \mathbf{a}_i)^T H_i^T H_i (\mathbf{x} - \mathbf{a}_i) \tag{4}$$

Our goal is obtained by minimizing  $J(\mathbf{x})$  regarding  $\mathbf{x}$ . Estimation of  $\mathbf{x}$ , i.e.,  $\hat{\mathbf{x}}$ , is obtained by partially differentiating each element of  $\mathbf{x}$  and setting them equal to 0 (shown in Equation (5)).

$$\begin{aligned} \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} &= 2 \sum_{i=1}^n H_i^T H_i (\mathbf{x} - \mathbf{a}_i) \\ &= 2 \left\{ \left( \sum_{i=1}^n H_i^T H_i \right) \mathbf{x} - \left( \sum_{i=1}^n H_i^T H_i \mathbf{a}_i \right) \right\} \\ &= 0 \end{aligned} \tag{5}$$

Thus, the estimation of Equation (6) is obtained.

$$\hat{\mathbf{x}} = \left( \sum_{i=1}^n H_i^T H_i \right)^{-1} \left( \sum_{i=1}^n H_i^T H_i \mathbf{a}_i \right) \tag{6}$$

Since  $H_i$  has the characteristic of  $H_i^T H_i = H_i^2 = H_i$ , i.e., a projection matrix, Equation (6) can be rewritten as in Equation (7), which we can use to estimate a convergence point using moving vectors.

$$\begin{aligned} \hat{x} &= \left( \sum_{i=1}^n H_i \right)^{-1} \left( \sum_{i=1}^n H_i \mathbf{a}_i \right) \\ &= \left\{ \sum_{i=1}^n \left( I_d - \mathbf{b}'_i \mathbf{b}'_i{}^T \right) \right\}^{-1} \left\{ \sum_{i=1}^n \left( I_d - \mathbf{b}'_i \mathbf{b}'_i{}^T \right) \mathbf{a}_i \right\} \end{aligned} \tag{7}$$

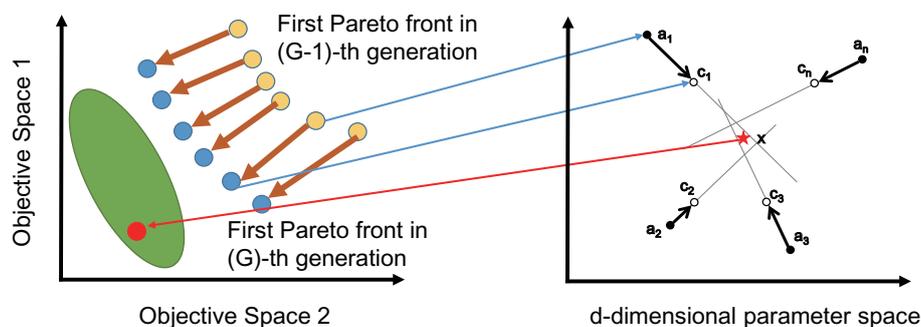
Besides these derivatives exactly estimating a convergence point, two approximated calculation methods are described in [22].

#### 4. Accelerating EMO Search Using an Estimated Convergence Point

##### 4.1. Philosophy of the Proposal

There are two subjects studied in the EMO algorithm research field; one study is the Pareto dominance issue, and the other one is EMO solution diversity issue. Almost all research on these two issues focuses on special handling in an objective space of an EMO algorithm, and frequently ignore the search condition in a parameter space. EMO algorithms try to find more non-dominated solutions with diversity. The solutions on the first Pareto solution frontier from the last generation to the next in an objective space supports information on how moving the variables in a parameter space can find promising Pareto solutions.

In Figure 3, we can find a set of the pairs of moving vectors in a parameter space in accord with the Pareto dominance information obtained in an objective space. We can use the moving vector information to estimate a convergence point that presents a promising area where Pareto solutions would be in a parameter space. We put such an estimated convergence point of a parameter space into EMO search and remove one of dominated solutions. EMO search should be enhanced considering such search information, and hopefully, EMO algorithm can find more non-dominated Pareto solutions quickly. This is a study hypothesis and motivation of our proposal that utilities an estimated convergence point to accelerate EMO search.



**Figure 3.** Estimation of promising Pareto solution area in parameter space using the dominance information from objective space to enhance EMO search.

##### 4.2. Estimation of Pareto Solution Frontier in a Parameter Space from Pareto Improvement Information in an Objective Space

There are three primary steps and/or issues in the proposed method to enhance EMO search. The first step/issue is how to make pairs of moving vectors in a parameter space from Pareto improvement information obtained in an objective space. We make two candidate groups of non-dominated solutions in the current generation and in the last generation, so Pareto solution improvement information can be obtained from these two group individuals. Here, we design two methods to make moving vector pairs ( $\mathbf{b}_i = \mathbf{c}_i - \mathbf{a}_i$  in Figure 2).

- We pick up one of non-dominated solutions in an objective space from one group, and find the nearest non-dominated solution in the other group, and then find their corresponding individuals in a parameter space to make these two solutions form a pair. (Estimation in objective space)
- We pick up one of non-dominated solutions in an objective space from one group, and find its corresponding individual in a parameter space, and then find this individual’s nearest individual in a parameter space to make these two solutions form a pair. (Estimation in parameter space)

After this, we delete these two solutions from the two groups, and we repeat this processing until one of groups becomes empty.

The second step estimates a convergence point in a parameter space using the moving vector pairs obtained from the first step. The estimation method uses Equation (7) to implement. The estimated point has high potential in the non-dominated Pareto solution frontier, and can therefore accelerate EMO search.

- Besides estimating only one estimated point, we can also estimate one point from only one single objective space each by each individually and use them together to accelerate EMO search (Estimation in each single objective space).

In the third step, we put the estimated convergence point as a search elite individual into EMO algorithms, and delete one/more of the dominated solutions in the current generation to enhance EMO search. This is the primary implementation within our proposal.

### 5. Experimental Evaluations

#### 5.1. Experiment Setting

We use five multi-objective benchmark functions from the ZDT test suite [37] to evaluate our proposed methods. We embed our proposed method into conventional NSGA-II [27] with different constructing methods of moving vector, and compare our proposed method with NSGA-II. Table 1 presents the benchmark function’s mathematical expressions. We examine these functions with three dimensional settings, i.e., two dimensions (2-D), 10-D, and 30-D. Table 2 shows the parameter settings of conventional NSGA-II algorithm used in the evaluation experiments.

**Table 1.** Multi-objective benchmark function used in evaluation [27]. All the Pareto frontier are  $g(x) = 1$ .

Functions	Definition
ZDT1	$f_1(x) = x_1$ $f_2(x) = g(x)[1 - \sqrt{\frac{x_1}{g(x)}}]$ $g(x) = 1 + 9 \frac{\sum_{i=2}^n x_i}{n-1}$
ZDT2	$f_1(x) = x_1$ $f_2(x) = g(x)[1 - (\frac{x_1}{g(x)})^2]$ $g(x) = 1 + 9 \frac{\sum_{i=2}^n x_i}{n-1}$
ZDT3	$f_1(x) = x_1$ $f_2(x) = g(x)[1 - \sqrt{\frac{x_1}{g(x)}} - \frac{x_1}{g(x)} \sin(10\pi x_1)]$ $g(x) = 1 + 9 \frac{\sum_{i=2}^n x_i}{n-1}$
ZDT4	$f_1(x) = x_1$ $f_2(x) = g(x)[1 - (\frac{x_1}{g(x)})^2]$ $g(x) = 1 + 10(n - 1) + \sum_{i=2}^n [x_i^2 - 10 \cos(4\pi x_i)]$
ZDT6	$f_1(x) = 1 - \exp(-4\pi x_1) \sin^6(6\pi x_1)$ $f_2(x) = g(x)[1 - (\frac{f_1(x)}{g(x)})^2]$ $g(x) = 1 + 9[\frac{\sum_{i=2}^n x_i}{n-1}]^{0.25}$

Three experiments are designed where different methods of constructing moving vectors, and these combined with the conventional NSGA-II algorithm. The legends displayed in figures and tables have the following meanings.

- **NSGA-II**; conventional NSGA-II algorithm;
- **Estimation in objective space**; we construct moving vectors from two subsequent non-dominated solution set in an objective space;
- **Estimation in parameter space**; we find the nearest offspring individual for each one in a parent generation, and make pairs in a parameter space; and
- **Estimation in each single objective space**; we consider each objective independently and estimation convergence point for each objective, where the estimated points may not be best on all objectives, but they have good potential in some objectives.

**Table 2.** NSGA-II algorithm parameter setting.

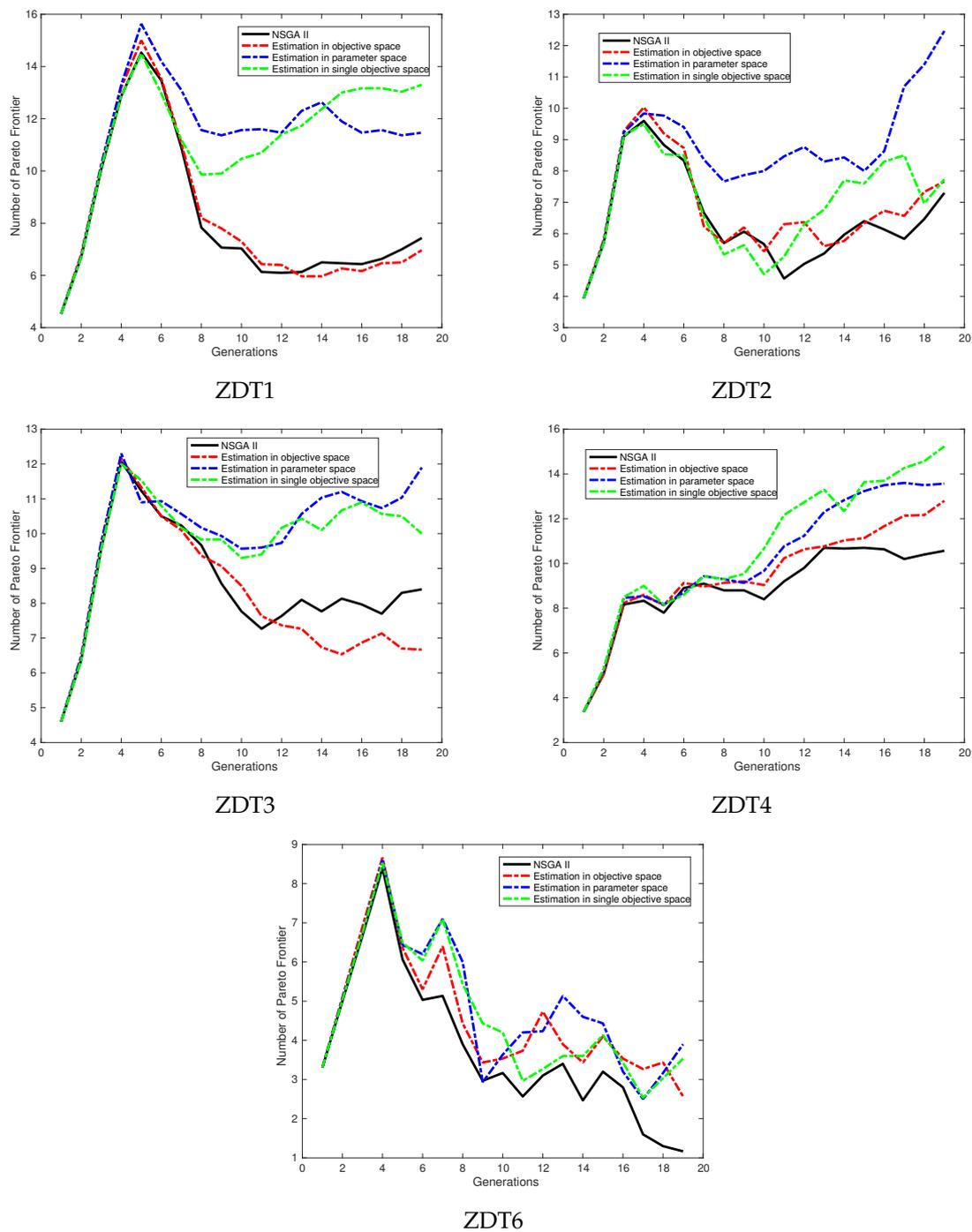
population size for 2-D, 10-D, and 30-D	20, 50, and 100
crossover rate	0.8
mutation rate	0.05
max. # of fitness evaluations, $MAX_{NFC}$ , for 2-D, 10-D, and 30-D search	400, 1000, and 10,000
dimensions of benchmark functions, $D$	2, 10, and 30
# of trial runs	30

## 5.2. Evaluation Metrics

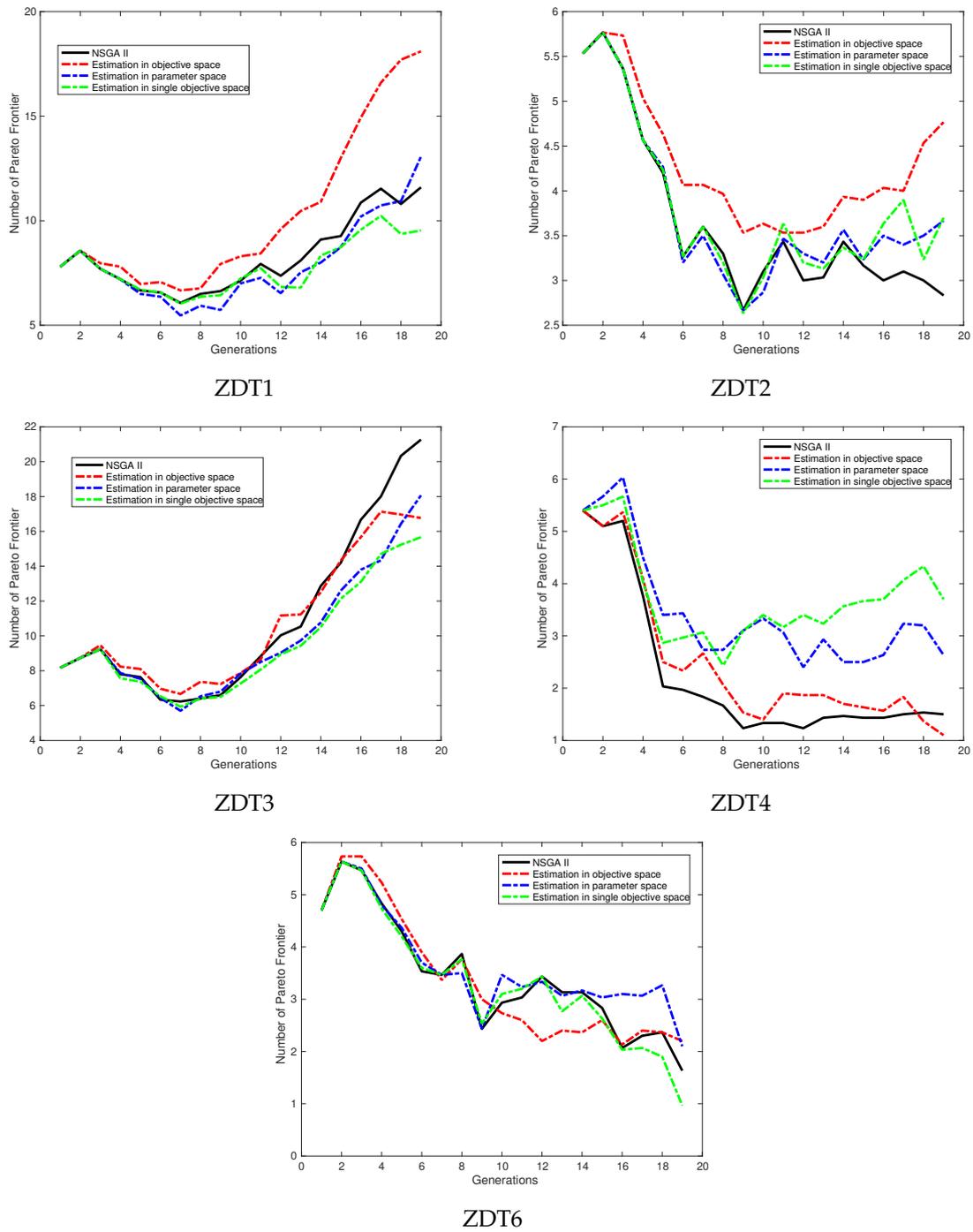
We set the stop conditions of each evaluation using the number of fitness calls instead of generations for fair evaluation, because our proposed methods increase additional fitness cost consumption. We set the stop conditions as 400 times, 1000 times, and 10,000 times of fitness evaluations in 2-D, 10-D, and 30-D problems, respectively. Besides, we test each benchmark function with 30 trial runs in three different dimensional settings.

Conventional NSGA-II is adopted as an example algorithm; other EMO algorithms can be also applied. Although there are many ways to generate estimated points, the greedy replacement strategy, where the estimated points will replace with the worst ranked and low diversity individuals to keep the same population size, is adopted in the proposed acceleration framework. To analyze the effect of the proposed acceleration framework, we calculate the number of non-dominated Pareto solutions in each generation shown in Figures 4–6.

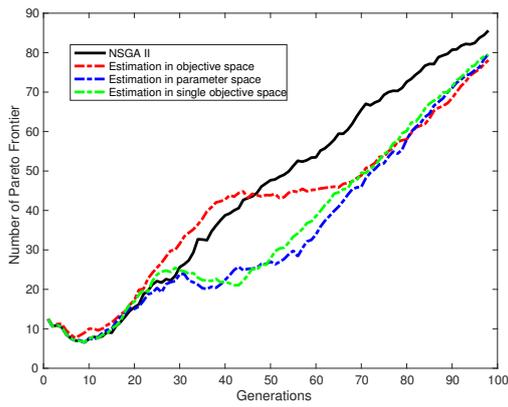
Hyper volume [38] is used to evaluate the diversity and acceleration performance of our proposal. Table 3 presents the hyper volume values of our proposed method and conventional NSGA-II algorithm at the stop condition in three different dimensional settings. We apply Wilcoxon signed-rank test for 30 trial runs data to evaluate the significance of hyper volume obtained by conventional NSGA-II and our proposal. Some functions without hyper volume value is due to reference point  $[-1, 1]$  setting.



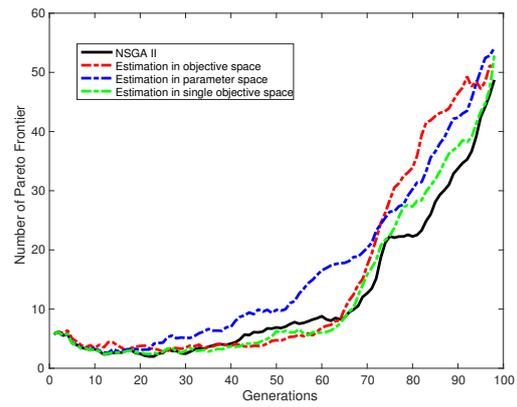
**Figure 4.** The number of Pareto solutions in every generation for 2-D benchmark problems. We can observe that proposed method can obtain more Pareto solutions for the most of cases.



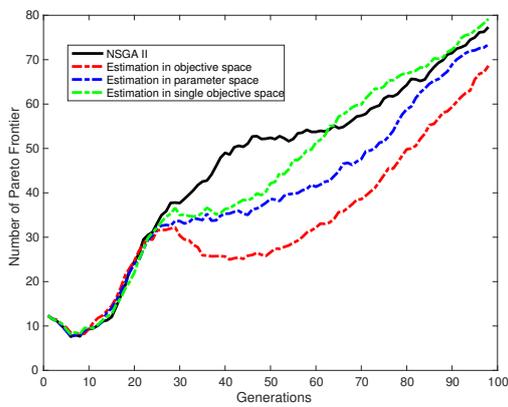
**Figure 5.** The number of Pareto solutions in every generation for 10-D benchmark problems. We can observe that proposed method can obtain more Pareto solutions for the most of cases.



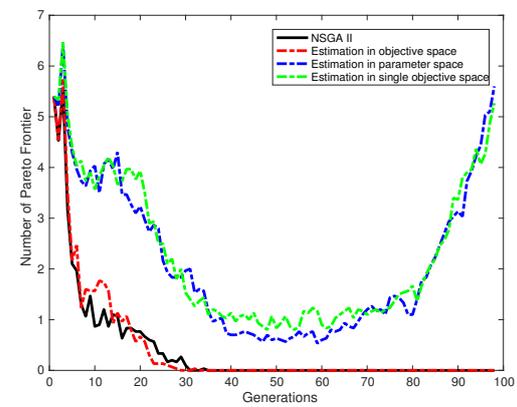
ZDT1



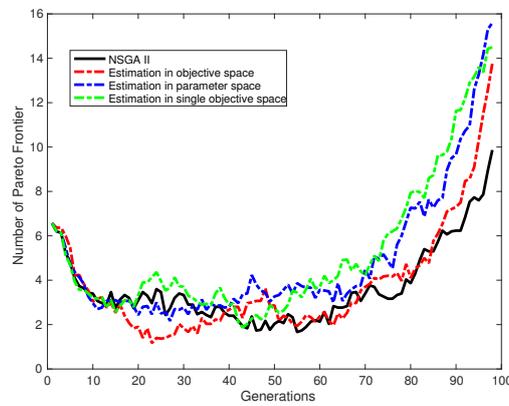
ZDT2



ZDT3



ZDT4



ZDT6

**Figure 6.** The number of Pareto solutions in every generation for 30-D benchmark problems. We can observe that proposed method can obtain more Pareto solutions for the most of cases.

**Table 3.** The average hypervolume values from 30 trials running of 4 methods in 2 dimensions (2-D), 10-D, and 30-D. Symbol † means that there is a significant difference between NSGA-II and Proposed method, i.e., NSGA-II + Estimation Point. The reference point is [1,1]. Obj., Para., and SinglePara. present objective space, parameter space, and each single parameter space, respective.

2-D tasks				
Func.	NSGA-II	Estimation in Obj.	Estimation in Para.	Estimation in SinglePara.
ZDT1	0.414567	0.417300	0.453833 †	0.480533 †
ZDT2	0.109833	0.117367	0.124367	0.113867
ZDT3	0.556733	0.552433	0.622733 †	0.622600 †
ZDT4	0.255733	0.261600	0.284933 †	0.337167 †
ZDT6	0.000033	0.002033	0.001833	0.000967
10-D tasks				
Func.	NSGA-II	Estimation in Obj.	Estimation in Para.	Estimation in SinglePara.
ZDT1	0.328033	0.337767	0.345533	0.339433
ZDT2	0.008633	0.012367	0.010100	0.008933
ZDT3	0.564067	0.545767	0.589400	0.588933
30-D tasks				
Func.	NSGA-II	Estimation in Obj.	Estimation in Para.	Estimation in SinglePara.
ZDT1	0.647167	0.654000	0.651533	0.648633
ZDT2	0.183333	0.157567	0.187700	0.190433
ZDT3	0.796133	0.792600	0.791533	0.794767
ZDT6	0.066233	0.068433	0.069633	0.066733

## 6. Discussions

### 6.1. Pareto Improvement of the Proposal

Pareto dominance and Pareto solution diversity are two metrics to evaluate the performance of EMO algorithms. In this work, we calculated the average number of Pareto solutions in every generation for each benchmark problem; see Figures 4–6. This is one of evaluation metrics for Pareto dominance in EMO. We also calculated hyper volume values at the maximal number of function calls for each dimension setting, and applied Wilcoxon signed-rank test to verify the significant difference among hyper volume values in Table 3. This is a demonstration of Pareto solution diversity for each EMO algorithm. We analyze and discuss our proposed method using these results.

From Figures 4–6, we can observe that methods estimating a convergence point in a parameter space and in a single objective space can obtain more Pareto solutions from all five multi-objective benchmark problems in 2-D setting. Method estimating a convergence point in an objective space fails in two benchmark functions, i.e., ZDT1 and ZDT3 in 2-D tasks. It indicates that moving vectors constructed from information of the nearest points in an objective space cannot exactly estimate the non-dominated Pareto frontier area in a parameter space. The same case can also be found in 10-D benchmark setting for ZDT3 and ZDT4, and 30-D benchmark setting for ZDT1, ZDT3, and ZDT4. We need to further consider improving the estimation the accuracy of estimation method of a convergence point in an objective space.

That in a single objective space works well in most of cases because this method replaces more than one estimated convergence point, and increases the population diversity for EMO algorithms. This indicates that the better individuals in each objective can improve optimization performance of EMO algorithms, although there are conflicts among multi-objective functions when EMO searches for non-dominated Pareto solutions. From this viewpoint, elite strategy-based EC acceleration methods

can be applied not only in single objective problems, but also have a potential to be applied in multi-objective problems.

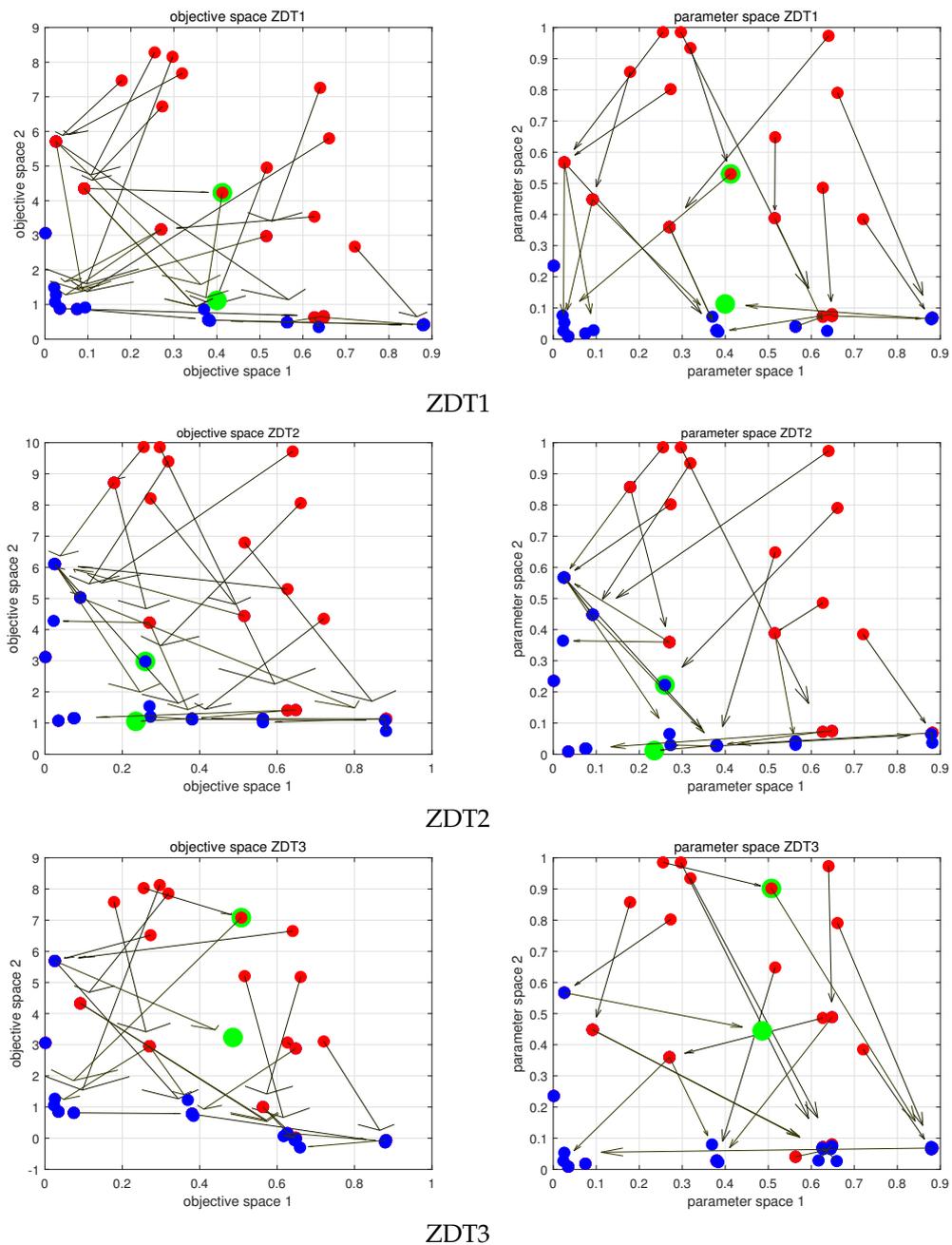
From observation of Table 3, the values of hyper volume from our proposed method are bigger than those from conventional NSGA-II algorithm for the most of tasks in 2-D benchmark problems. The Wilcoxon signed-rank test results showed a significant difference between our proposed method and the conventional NSGA-II algorithm in estimation in a parameter space and estimation in a single objective space. These results demonstrate that our proposed method can obtain non-dominated Pareto solution with more diversities for EMO algorithms. However, it is not significant shown in 10-D and 30-D benchmark problems. It is a limitation for our proposal, and we need to improve it in our future work.

### 6.2. Topological Structure of Moving Vectors and Modality Characteristic of Pareto Improvement

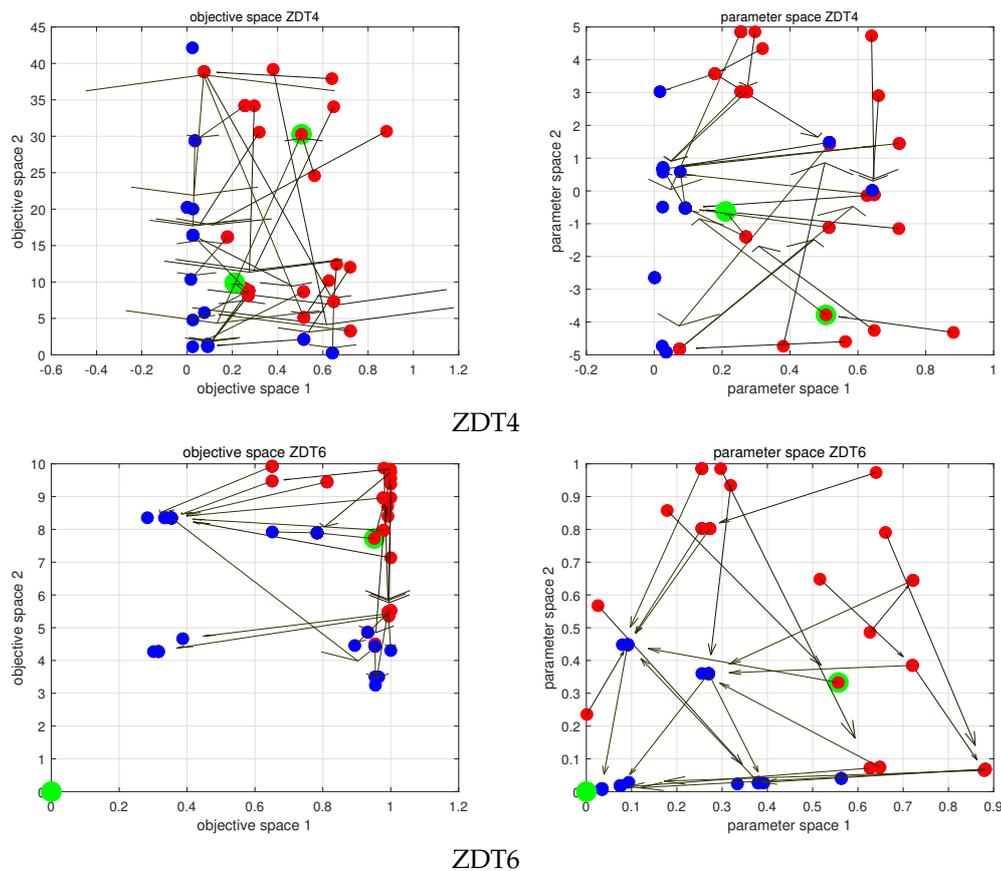
The basic philosophy of our proposed method to accelerate EMO search lies in three hypotheses. First, we can obtain the information to improve non-dominated Pareto solutions through Pareto solution evolutions from the last generation to the current generation in an objective space. Second, after we obtain the information, the moving vectors can be made in an objective space or in a parameter space. Third, the estimated convergence point of these moving vectors has a high possibility that locals in the non-dominated Pareto solution frontier area in a parameter space. From the modality viewpoint, the distribution of Pareto solutions shows a uni-modal characteristic in the objective space. In the case of the Pareto improvement from the last generation to the current generation, do the corresponding individuals also present a uni-modal distribution characteristic in a parameter space? We examine this question here.

We present improved EMO evolutions along three generations' EMO evolution condition both in an objective space and a parameter space for these benchmark functions with 2-D setting (see Figures 7 and 8). The arrows demonstrate the Pareto improvement directions between two generations in both spaces. From Figures 7 and 8, we observe that with regards to the directions of arrows in an objective space, all of them are towards almost the same direction, i.e., their angles are less than 90 degrees. However, in the parameter space, the arrows are not towards the same direction. It displays a multi-modal distribution characteristic, e.g., in the ZDT4 and ZDT6 benchmark problems. From these observations, it indicates that Pareto improvement in the objective space presents a uni-modal characteristic, while it presents a multi-modal characteristic in a parameter space. In Figure 4, the numbers of the first Pareto frontier solution from four methods are almost the same, but their acceleration performances are not significant. This is one our discovery on the modality characteristic of Pareto improvement in both an objective space and a parameter space.

From Figures 7 and 8, there is a multi-modal characteristic in a parameter space when the Pareto improvement occurs from one generation to the next. The third hypothesis of proposed method is not always correct, therefore, the proposed method can work well in the uni-modal condition of Pareto improvement, and by chance well in the multi-modal one. From Table 3, there is not a significant difference between NSGA-II algorithm and our proposed method in ZDT6. This experiment's results verify our analysis and observations. The multi-modal characteristic of Pareto improvement in a parameter space is an issue when applying our proposal to enhance EMO search.



**Figure 7.** Two-dimensional demonstration of Pareto solution improvement in an objective space (left) and their corresponding individuals in parameter space (right) of ZDT1, ZDT2, and ZDT3. The arrows show directions of both Pareto solution improvement and moving vectors. We can observe that there is a uni-modal landscape for Pareto solution improvement in an objective space; however, it is a multi-modal landscape for Pareto improvement in a parameter space. The green point is the estimated convergence point, most of the red points and most of the blue points are in the first generation and in the third generation, respectively.



**Figure 8.** Two-dimensional demonstration of Pareto solution improvement in an objective space (left) and their corresponding individuals in parameter space (right) of ZDT4, and ZDT6. The arrows show directions of both Pareto solution improvement and moving vectors. We can observe that there is a uni-modal landscape for Pareto solution improvement in an objective space; however, it is a multi-modal landscape for Pareto improvement in a parameter space. The green point is the estimated convergence point, most of the red points and most of the blue points are in the first generation and in the third generation, respectively.

### 7. Conclusions and Future Work

In this work, we use an estimated convergence point from dominance information of Pareto solution improvement to enhance EMO search. We use NSGA-II as a test algorithm and five multi-objective functions to qualitatively evaluate our proposal. We found that our proposed method can enhance EMO search in some benchmark problems, especially for the high-dimensional and complex multi-objective problems which can obtain a greater number of Pareto solutions. We also analyzed the modality of the Pareto improvement in both an objective space and a parameter space. We found that the Pareto improvement in an objective space demonstrates a uni-modal characteristic, but a multi-modal one in parameter space. It is one of the discoveries in this work.

In the future, we will further investigate the proposed method in a variety of multi-objective problems, especially for real-world problems. How to find the exact pairs information of moving vectors is one of the potential study subjects in our method. It influences the accuracy of estimated point to make different performances of our proposed by using the point. The multi-modal characteristic of moving vectors in a parameter space is an issue for our estimation method. We will use clustering methods to find the representative moving vectors to find the estimated point in a parameter space. Another study issue is a search condition using multi-objective fitness landscape and an estimated convergence point. These and other study subjects will be involved in our future research work.

**Author Contributions:** Conceptualization, Y.P.; Funding acquisition, Y.P. and H.T.; Investigation, Y.P. and J.Y.; Methodology, Y.P. and H.T.; Project administration, H.T.; Software, Y.P. and J.Y.; Validation, Y.P.; Visualization, Y.P.; Writing—original draft, Y.P.

**Funding:** Japan Society for the Promotion of Science: JP15K00340 and 18K11470.

**Acknowledgments:** The work is supported by the JSPS Grant-in-Aid for Scientific Research C (JP15K00340 and 18K11470).

**Conflicts of Interest:** The author declares that there is no conflict of interests regarding the publication of this paper.

## References

1. Li, B.; Li, J.; Tang, K.; Yao, X. Many-objective evolutionary algorithms: A survey. *ACM Comput. Surv.* **2015**, *48*, 13. [[CrossRef](#)]
2. Jin, Y. A comprehensive survey of fitness approximation in evolutionary computation. *Soft Comput.* **2005**, *9*, 3–12. [[CrossRef](#)]
3. Jin, Y.; Olhofer, M.; Sendhoff, B. A Framework for evolutionary optimization with approximate fitness functions. *IEEE Trans. Evol. Comput.* **2002**, *6*, 481–494.
4. Jin, Y. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm Evol. Comput.* **2011**, *1*, 61–70. [[CrossRef](#)]
5. Witkin, A.P. Scale-space filtering. In Proceedings of the 8th International Joint Conference Artificial Intelligence, Karlsruhe, Germany, 8–12 August 1983; pp. 1019–1022.
6. Takagi, H.; Ingu, T.; Ohnishi, K. Accelerating a GA convergence by fitting a single-peak function. *J. Soft* **2003**, *15*, 219–229. [[CrossRef](#)]
7. Pei, Y.; Takagi, H. Accelerating IEC and EC searches with elite obtained by dimensionality reduction in regression spaces. *Evol. Intell.* **2013**, *6*, 27–40. [[CrossRef](#)]
8. Pei, Y.; Zheng, S.; Tan, Y.; Takagi, H. Effectiveness of approximation strategy in surrogate-assisted fireworks algorithm. *Int. J. Mach. Learn. Cybern.* **2015**, *6*, 795–810. [[CrossRef](#)]
9. Zhao, N.; Zhao, Y.; Fu, C. Genetic algorithm with fitness approximate mechanism. *J. Natl. Univ. Def. Technol.* **2014**, *36*, 116–121.
10. Pei, Y.; Takagi, H. Fourier analysis of the fitness landscape for evolutionary search acceleration. In Proceedings of the 2012 IEEE Congress on Evolutionary Computation, Brisbane, Australia, 10–15 June 2012; pp. 1–7.
11. Michael, D.S.; Hod, L. Coevolution of fitness predictors. *IEEE Trans. Evol. Comput.* **2008**, *12*, 736–749.
12. Michael, D.S.; Hod, L. Co-evolution of fitness maximizers and fitness predictors. In Proceedings of the Genetic and Evolutionary Computation Conference, Washington, DC, USA, 25–29 June 2005; pp. 1–8.
13. Michael, D.S.; Hod, L. Co-evolving fitness predictors for accelerating evaluations and reducing sampling. *Genet. Programm. Theory Pract. IV* **2006**, *5*, 113–130.
14. Michael, D.S.; Hod, L. Predicting solution rank to improve performance. In Proceedings of the 12th Annual Genetic and Evolutionary Computation Conference, Portland, OR, USA, 7–11 July 2010; pp. 949–955.
15. He, Y.; Yuen, S.Y.; Lou, Y. Exploratory landscape analysis using algorithm based sampling. In Proceedings of the 2018 Genetic and Evolutionary Computation Conference Companion, Kyoto, Japan, 15–19 July 2018; pp. 211–212.
16. Mersmann, O.; Bischl, B.; Trautmann, H.; Preuss, M.; Weihs, C.; Rudolph, G. Exploratory landscape analysis. In Proceedings of the Genetic and Evolutionary Computation Conference, Dublin, Ireland, 12–16 July 2011; pp. 829–836.
17. Wang, G.G.; Tan, Y. Improving metaheuristic algorithms with information feedback models. *IEEE Trans. Cybern.* **2017**. [[CrossRef](#)] [[PubMed](#)]
18. Wang, G.G.; Guo, L.; Gandomi, A.H.; Hao, G.S.; Wang, H. Chaotic krill herd algorithm. *Inf. Sci.* **2014**, *274*, 17–34. [[CrossRef](#)]
19. Pei, Y. Chaotic evolution: fusion of chaotic ergodicity and evolutionary iteration for optimization. *Nat. Comput.* **2014**, *13*, 79–96. [[CrossRef](#)]
20. Pei, Y.; Takagi, H. Research progress survey on interactive evolutionary computation. *J. Ambient Intell. Hum. Comput.* **2018**, 1–14. [[CrossRef](#)]

21. Pei, Y.; Takagi, H. A survey on accelerating evolutionary computation approaches. In Proceedings of the 2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR), Dalian, China, 14–16 October 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 201–206.
22. Murata, N.; Nishii, R.; Takagi, H.; Pei, Y. Analytical estimation of the convergence point of populations. In Proceedings of the 2015 IEEE Congress on Evolutionary Computation, Sendai, Japan, 25–28 May 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 2619–2624.
23. Yu, J.; Pei, Y.; Takagi, H. Accelerating evolutionary computation using estimated convergence points. In Proceedings of the 2016 IEEE Congress on Evolutionary Computation, Vancouver, BC, Canada, 24–29 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1438–1444.
24. Yu, J.; Takagi, H. Clustering of moving vectors for evolutionary computation. In Proceedings of the 2015 7th International Conference of Soft Computing and Pattern Recognition, Fukuoka, Japan, 13–15 November 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 169–174.
25. Dantzig, G.B. *Maximization of a Linear Function of Variables Subject to Linear Inequalities*; John Wiley & Sons: New York, NY, USA, 1951.
26. Wallis, J. A treatise of algebra, both historical and practical. *Philos. Trans.* **1685**, *15*, 1095–1106.
27. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [[CrossRef](#)]
28. Srinivas, N.; Deb, K. Multi-objective function optimization using non-dominated sorting genetic algorithms. *Evol. Comput.* **1994**, *2*, 221–248. [[CrossRef](#)]
29. Deb, K.; Jain, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving Problems With Box Constraints. *IEEE Trans. Evol. Comput.* **2013**, *18*, 577–601. [[CrossRef](#)]
30. Carlos, M.F.; Peter, F. Genetic algorithms for multiobjective optimization: Formulation discussion and generalization. In Proceedings of the 5th International Conference on Genetic Algorithms, Urbana, Champaign, IL, USA, 17–21 July 1993; pp. 416–423.
31. Rey Horn, J.; Nafpliotis, N.; Goldberg, D.E. A niched Pareto genetic algorithm for multiobjective optimization. In Proceedings of the First IEEE Conference on Evolutionary Computation, Orlando, FL, USA, 27–29 June 1994; pp. 82–87.
32. Zitzler, E.; Thiele, L. Multi-Objective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.* **1999**, *3*, 257–271. [[CrossRef](#)]
33. Eckart, Z.; Marco, L.; Lothar, T. SPEA2: Improving the strength Pareto evolutionary algorithm. In *Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems, Proceedings of the EUROGEN2001 Conference, Athens, Greece, 19–21 September 2001*; International Center for Numerical Methods in Engineering: Barcelona, Spain, 2001; pp. 95–100.
34. David, W., C.; Joshua, D., K.; Martin, J., O. The Pareto-envelope based selection algorithm for multi-objective optimization. In Proceedings of the 6th International Conference on Parallel Problem Solving from Nature, Paris, France, 18–20 September 2000; pp. 839–848.
35. Corne, D.W.; Jerram, N.R.; Knowles, J.D.; Oates, M.J.; J, M. PESA-II: Region-based selection in evolutionary multiobjective optimization. In Proceedings of the Genetic and Evolutionary Computation Conference, San Francisco, CA, USA, 7–11 July 2001; pp. 283–290.
36. Pei, Y.; Hao, J. Non-dominated sorting and crowding distance based multi-objective chaotic evolution. In *International Conference in Swarm Intelligence*; Springer: Cham, Switzerland, 2017; pp. 15–22.
37. Zitzler, E.; Deb, K.; Thiele, L. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evol. Comput.* **2000**, *8*, 173–195. [[CrossRef](#)] [[PubMed](#)]
38. Zitzler, E.; Brockhoff, D.; Thiele, L. The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. In *Evolutionary Multi-Criterion Optimization*; Springer: Berlin, Germany, 2007; pp. 862–876.

