2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Normalization in R2-based Hypervolume and Hypervolume Contribution Approximation

Guotong Wu, Tianye Shu, Ke Shang*, Hisao Ishibuchi*
Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation
Department of Computer Science and Engineering
Southern University of Science and Technology
Shenzhen, China

***Abstract*—In this paper, we examine the effect of normalization in R2-based hypervolume and hypervolume contribution approximation. The fact is that the region with different scales on objective space brings approximation bias. The basic idea of normalization is to perform a coordinate transformation to make the shape of the approximated region more regular, and then transform it to obtain the final value according to the property of hypervolume and hypervolume contribution. The performance of normalization is evaluated on different datasets by comparing it with the original R2-based method. We use two different metrics to evaluate hypervolume and hypervolume contribution separately, and the results indicate that normalization does exactly improve the approximation accuracy and outperforms the original R2-based method.**

***Index Terms*—Multi-objective Optimization, Hypervolume, Normalization**

I. Introduction

In the field of evolutionary multi-objective optimization (EMO), many indicators such as GD [1], IGD [2], hypervolume [3] and R2 [4] have been proposed to evaluate different multi-objective optimization algorithms. Among these indicators, hypervolume has been widely used due to its ability to evaluate both the convergence and diversity of a solution set. Since the hypervolume is Pareto compliant [5] (i.e., if a solution set A is better than another solution set B, its hypervolume is always larger than the hypervolume of B), some algorithms (i.e., SMS-EMOA [6, 7], FV-MOEA [8], HypE [9]) directly use the hypervolume in their algorithmic step.

For SMS-EMOA, a new solution is generated and added to the population in each generation. To keep the overall population size constant, one solution has to be removed from the entire population, making the hypervolume of the remaining solutions maximized. Therefore we need to calculate the hypervolume contribution of each solution to the whole solution set.

This work was supported by National Natural Science Foundation of China (Grant No. 62002152, 62250710163, 62250710682), Guangdong Provincial Key Laboratory (Grant No. 2020B121201001), the Program for Guangdong Introducing Innovative and Enterpreneurial Teams (Grant No. 2017ZT07X386), The Stable Support Plan Program of Shenzhen Natural Science Fund (Grant No. 20200925174447003), Shenzhen Science and Technology Program (Grant No. KQTD20161125143555 31).
*Corresponding authors: Ke Shang (kshang@foxmail.com), Hisao Ishibuchi (hisao@sustech.edu.cn)

Since computing hypervolume and hypervolume contribution is an NP-hard problem, its computation time grows exponentially as the number of objectives increases. When the number of objectives is small, the exact value of hypervolume can be computed quickly. Some efficient hypervolume calculation methods have been proposed such as WFG [10], QHV [11], and HBDA [12]. When the number of objectives is large, approximation methods are usually used to estimate hypervolume and hypervolume contribution since the calculation of the exact value is very time-consuming.

The approximation methods can be briefly categorized into the point-based method and the line-based method. The point-based method is also known as the Monte Carlo sampling method [13], [14], which uses the proportion of the points in the hypervolume region for approximation. The line-based method is also known as the R2-based method [15]. This method uses the lengths of the line segments in the region for approximation.

In the papers [15, 16], Shang et al. proposed a new R2-based method to approximate the hypervolume and hypervolume contribution. This method further improves the accuracy of the line-based approximation. However, some solution sets are not uniformly distributed, thus the line segments distribute unevenly in the region, which brings a large deviation for calculating the hypervolume. Even if the distribution of the solutions is uniform, due to the mutual positions between each solution, it also brings bias to the approximation of hypervolume contribution in some cases.

Therefore, this paper examines the effect of normalization in hypervolume and hypervolume contribution approximation based on the R2 method. This normalization method has the following two characteristics: 1) It does not change the direction of the vectors and can directly use the R2-based method after normalization. 2) It uses only simple translation and scaling coordinate transformations, which are very easy to operate, while the results can be significantly improved.

This paper is organized as follows. Section II presents the preliminaries of the study. Section III introduces the normalization method of hypervolume and hypervolume contribution respectively. Section IV explains the experimental setting and results. Section V concludes this paper.

978-0-7381-4408-5/23/$31.00 ©2023 IEEE

II. Preliminaries

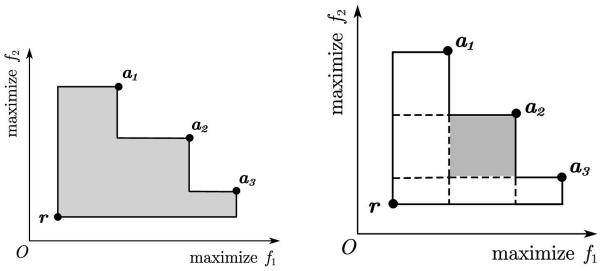
In this section, we first explain the concept of hypervolume and hypervolume contribution. Then the R2-based approximation methods are introduced. It should be noted that all problems in this paper are maximization problems.

A. Hypervolume and Hypervolume Contribution

The hypervolume is an indicator to evaluate the performance of a solution set. Given a two-objective case in Figure 1(a), the hypervolume is the area covered by the non-dominated solution set $A = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$ and the reference point $\boldsymbol{r}$. From a mathematical perspective, hypervolume can be defined as follows:

$$\mathrm{HV}(A, \boldsymbol{r}) = \mathcal{L}\left(\bigcup_{\boldsymbol{a}\in A}\{\boldsymbol{b} \mid \boldsymbol{a} \succ \boldsymbol{b} \succ \boldsymbol{r}\}\right), \tag{1}$$

where $\mathcal{L}$ is the Lebesgue measure of a set, and $\boldsymbol{a} \succ \boldsymbol{b}$ means $\boldsymbol{b}$ is dominated by $\boldsymbol{a}$.



(a) Hypervolume (b) Hypervolume contribution

Fig. 1. The explanation of the hypervolume and the hypervolume contribution in a two-objective case.

The hypervolume contribution of a single solution $\boldsymbol{s}$ evaluates its contribution to a solution set. Using Figure 1(b) for illustration, the hypervolume contribution of the solution $\boldsymbol{s} = \boldsymbol{a_2}$ is the decreasing area after removing $\boldsymbol{s}$ from the solution set. This process can be written as:

$$\mathrm{HVC}(\boldsymbol{s}, A, \boldsymbol{r}) = \mathrm{HV}(A, \boldsymbol{r}) - \mathrm{HV}(A\backslash\{\boldsymbol{s}\}, \boldsymbol{r}). \tag{2}$$

B. R2-based Hypervolume Approximation

Figure 2 explains the R2-based hypervolume approximation method. A set of vectors $\Lambda = \{\boldsymbol{\lambda}^1, \ldots, \boldsymbol{\lambda}^n\}$ starts from a reference point and intersects with the boundary of the shaded region. The hypervolume (i.e., the area of the shaded region) can be approximated:

$$HV(S, \mathbf{r}) \approx \frac{\pi^{m/2}}{mn2^{m-1}\Gamma(m/2)} \sum_{i=1}^{n} l_i^m, \tag{3}$$

where $\Gamma(x) = \int_0^\infty z^{x-1}e^{-z}dz$ is the Gamma function, l_i is the length of vector $\boldsymbol{\lambda_i}$ between the reference point $\boldsymbol{r}$ and the boundary, and m is the number of objectives.

Each vector in Λ satisfies $\left\|\boldsymbol{\lambda}^i\right\|_2 = 1, \lambda_j^i \geq 0, i = 1, \ldots, n, j = 1, \ldots, m$. From paper [16, 17], the length l_i can be derived as:

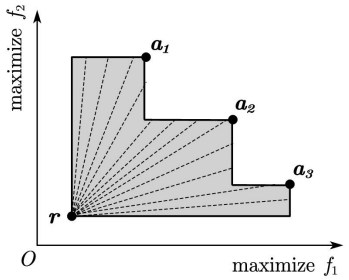


Fig. 2. An example of the hypervolume approximation of a two-objective solution set $A = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$.

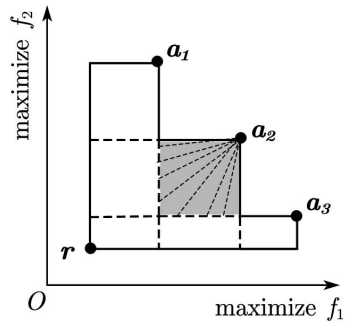


Fig. 3. The hypervolume contribution of $\boldsymbol{a_2}$ to a 2-objective solution set $A = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$.

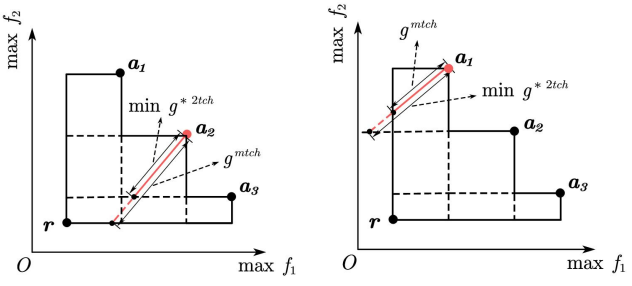
$$l_i = \max_{\mathbf{s}\in A} \min_{j\in\{1,\ldots,m\}} \left\{\frac{|r_j - s_j|}{\lambda_j^i}\right\}, i = 1, \ldots, n. \tag{4}$$

C. R2-based Hypervolume Contribution Approximation

Similar to hypervolume, we also use the average length to approximate the hypervolume contribution in Figure 3. However, these vectors start from the solution $\boldsymbol{s} = \boldsymbol{a_2}$ and intersect with the edge of other points and the reference point. Thus the calculation of l_i is a little different:

$$l_i = \min\{l_i^1(\boldsymbol{s}, A\backslash\{\boldsymbol{s}\}, \boldsymbol{\lambda^i}), l_i^2(\boldsymbol{s}, \boldsymbol{r}, \boldsymbol{\lambda^i})\}, \tag{5}$$

where l_i^1 and l_i^2 represent the length of the vector λ_i in two situations respectively.



(a) Situation 1 (b) Situation 2

Fig. 4. Illustrations of (a) the vector which intersects with the attainment surface of the set $A\backslash\{\boldsymbol{s}\}$, and (b) the vector intersects with the reference point.

To better explain the meaning of l_i^1 and l_i^2, Figure 4 shows the difference between the two situations. In situation 1 (Figure 4(a)), each vector starts from point $\boldsymbol{s}$ and first intersects with the attainment surface of the set $A\backslash\{\boldsymbol{s}\}$.

$$l_1^1(\boldsymbol{s}, A\backslash\{\boldsymbol{s}\}, \boldsymbol{\lambda^i}) = \min_{\boldsymbol{a}\in A\backslash\{\boldsymbol{s}\}}\{g^{*2tch}(\boldsymbol{a}|\boldsymbol{\lambda^i}, \boldsymbol{s})\}, \tag{6}$$

where g^{*2tch} is the 2-Tch function defined as follows:

$$g^{*2tch}(\boldsymbol{a}|\boldsymbol{\lambda^i}, \boldsymbol{s}) = \max_{j\in\{1,\dots,m\}} \frac{s_j - a_j}{\lambda_j^i}, \tag{7}$$

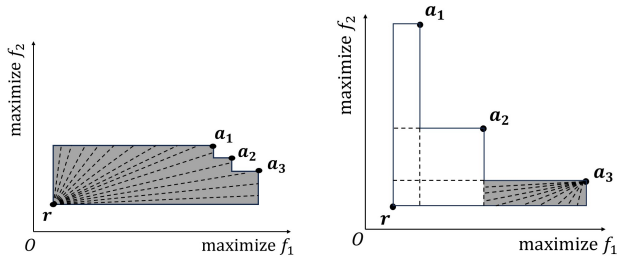
In situation 2 (Figure 4(b)), the intersection point is located on the boundary of the whole attainment surface, which is the edge of the reference point:

$$l_i^2(\boldsymbol{s}, \boldsymbol{r}, \boldsymbol{\lambda^i}) = g^{mtch}(\boldsymbol{r}|\lambda^i, \boldsymbol{s}) = \min_{j\in\{1,\dots,m\}} \frac{|s_j - r_j|}{\lambda_j^i}. \tag{8}$$

Note that Equation (7) is only applicable for maximizing all objectives, which is assumed throughout this paper. In the case of minimization, the positions of s_j and a_j should be swapped.

III. Normalization in R2-based approximation

In this section, we first introduce the intuition about why normalization is needed for the hypervolume and hypervolume contribution approximation. Then the detailed processes of normalization in hypervolume and hypervolume contribution approximation will be presented.



(a) HV approximation (b) HVC approximation

Fig. 5. Example of (a) the hypervolume approximation and (b) the hypervolume contribution approximation.

When using the R2-based method to calculate hypervolume and hypervolume contribution, it aims to approximate the area of the corresponding region. However, in the actual calculation, it can be found that when the shape is irregular (i.e., different scales in different objectives), it may cause a large deviation in the results. Here is a two-dimensional example of hypervolume approximation, as shown in Figure 5(a), the points are almost distributed in the bottom right corner. However, since the vectors are generated uniformly, in this case, the distribution of vectors close to the reference point is very dense, while the distribution of vectors close to $\{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$ is very sparse. The scale of the first objective (f_1) and the second objective (f_2) are significantly different (i.e., $f_1(\boldsymbol{a_3})$ is much larger than $f_2(\boldsymbol{a_3})$), which can produce a large bias.

Another example is shown in Figure 5(b). When approximating the hypervolume contribution of $\boldsymbol{a_3}$, the shaded region has a similar problem (i.e., large difference between the scale of f_1 and f_2). In this case, even though the distribution of the solution set is relatively uniform, the position of each solution can also lead to an irregular region.

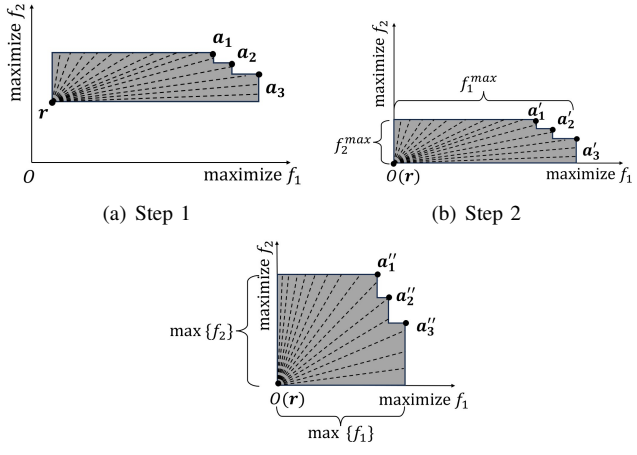
Hence, when confronted with the aforementioned example, it is imperative to apply normalization to eliminate the effect of irregular shapes. The normalization will be applied directly to the solution set, while the vector set remains unchanged. In the following, we will introduce the corresponding transformation properties for hypervolume and hypervolume contribution [18]:

- **Property 1**. For any positive real vector $\boldsymbol{\alpha} \in \mathbb{R}^m_{>0}$, $HV(S, \boldsymbol{r}) = \frac{1}{\prod_{i=1}^m \alpha_i} HV(\boldsymbol{\alpha}\odot S, \boldsymbol{\alpha}\odot\boldsymbol{r}), HVC(\boldsymbol{s}, S, \boldsymbol{r}) = \frac{1}{\prod_{i=1}^m \alpha_i} HVC(\boldsymbol{\alpha}\odot\boldsymbol{s}, \boldsymbol{\alpha}\odot S, \boldsymbol{\alpha}\odot\boldsymbol{r})$, where $\odot$ denotes the element-wise multiplication.
- **Property 2**. For any real vector $\boldsymbol{\beta} \in \mathbb{R}^m$, $HV(S, \boldsymbol{r}) = HV(S + \boldsymbol{\beta}, \boldsymbol{r} + \boldsymbol{\beta})$, $HVC(\boldsymbol{s}, S, \boldsymbol{r}) = HVC(\boldsymbol{s} + \boldsymbol{\beta}, S + \boldsymbol{\beta}, \boldsymbol{r} + \boldsymbol{\beta})$.
- **Property 3**. $HV(S, \boldsymbol{r}) = HV(-S, -\boldsymbol{r})$, $HVC(\boldsymbol{s}, S, \boldsymbol{r}) = HVC(-\boldsymbol{s}, -S, -\boldsymbol{r}))$. In this property, $HV(-S, -\boldsymbol{r})$ and $HVC(-\boldsymbol{s}, -S, -\boldsymbol{r}))$ is calculated for maximization problems whereas $HV(S, \boldsymbol{r})$ and $HVC(\boldsymbol{s}, S, \boldsymbol{r})$ is calculated for minimization problems.

The above three properties can be easily deduced from the properties of the Lebesgue measure. With the help of the above properties, we can easily transform the coordinates of the solutions and the reference point, so that the shape of the region is more regular while ensuring the hypervolume and hypervolume contribution approximation can be more accurate.

A. Normalization in R2-based Hypervolume Approximation

To make the normalization in hypervolume approximation easier to understand, we will explain the processes with a specific example in the following.



(a) Step 1 (b) Step 2

(c) Step 3

Fig. 6. Example of normalization on hypervolume. (a) is the original solution set, (b) is the solution set after the translation transformation, and (c) is the solution set after the scaling transformation.

As shown in Figure 6(a), now we need to approximate the hypervolume of the solution set $A = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$, where the coordinates corresponding to each point are $\boldsymbol{r} = (1, 4), \boldsymbol{a_1} = (20, 10), \boldsymbol{a_2} = (22, 8), \boldsymbol{a_3} = (25, 6)$. The shaded region has very irregular scales on f_1 and f_2.

- Step 1. All points, including the reference point $\boldsymbol{r}$, undergo a global translation to ensure the position of $\boldsymbol{r}$ is transformed to $\boldsymbol{O}$:

$$a'^j = a^j - r^j, j = 1, \ldots, m. \tag{9}$$

  According to **Property 2**, all the coordinates are added to an identical vector at the same time, thus the hypervolume remains unchanged. As shown in Figure 6(b), the coordinates of each point after the transformation are $\boldsymbol{r'} = (0, 0), \boldsymbol{a'_1} = (19, 6), \boldsymbol{a'_2} = (21, 4), \boldsymbol{a'_3} = (24, 2)$.
- Step 2. The coordinates of all points undergo a scaling transformation according to the maximum value above each objective:

$$a''^j = \frac{a'^j}{\max\{f_j\}}, j = 1, \ldots, m, \tag{10}$$

  where $\max\{f_j\}$ is the maximum value of the jth objective for all solutions. According to Property 1, when the coordinates of all points are subject to a scaling transformation, the value of the corresponding hypervolume should be divided by the scaling value. Note that in Figure 6 we normalize the maximum value on each objective to 1 for convenience. As shown in Figure 6(c), the coordinates of each point after the transformation are $\boldsymbol{r'} = (0, 0), \boldsymbol{a''_1} = (\frac{19}{24}, 1), \boldsymbol{a''_2} = (\frac{21}{24}, \frac{4}{6}), \boldsymbol{a''_3} = (1, \frac{2}{6})$.
- Step 3. According to the R2-based method, the hypervolume is approximated and transformed to the final value:

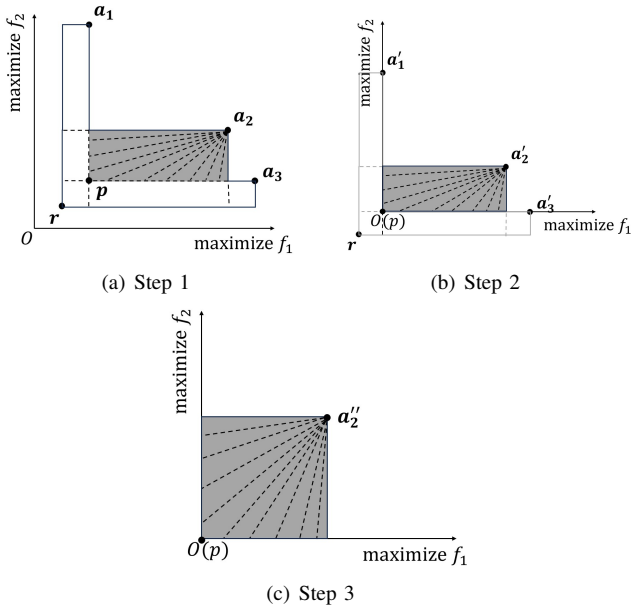
$$\widetilde{\mathrm{HV}} = \prod_{i=1}^{m} \max\{f_j\} \times \widetilde{\mathrm{HV}}', \tag{11}$$

  where $\widetilde{\mathrm{HV}}'$ is the approximated hypervolume obtained by the R2-based method after the coordinate transformation.

B. Normalization in R2-based Hypervolume Contribution approximation

Normalization in the process of approximating hypervolume contribution also requires coordinate transformation. However, the region of the hypervolume is determined by both the solution set and the reference point (as shown in Figure 5(a)). Therefore, we can directly transform the reference point to $\boldsymbol{O}$ and perform the normalization process. In contrast, the reference point may not be related to the hypervolume contribution region (as depicted in Figure 5(b)). Thus, it becomes essential to identify an appropriate base point ($\boldsymbol{p}$) when approximating the hypervolume contribution.

Similarly, the normalization process in approximating hypervolume contribution is explained below with a specific example in Figure 7(a). Suppose we have a solution set $A = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$ and a reference point $\boldsymbol{r}$, where the coordinates



(a) Step 1 (b) Step 2

(c) Step 3

Fig. 7. Example of normalization on hypervolume contribution. (a) is the original solution set, (b) is the solution set after the translation transformation, and (c) is the solution set after the scaling transformation.

corresponding to each point are $\boldsymbol{r} = (1, 1), \boldsymbol{a_1} = (2, 10), \boldsymbol{a_2} = (8, 4), \boldsymbol{a_3} = (10, 2)$. We need to approximate the hypervolume contribution of $\boldsymbol{a_2}$ to A.

- Step 1. Identify the normalization region of hypervolume contribution to the solution. The rectangle region chosen for normalization should be as tight as possible to the real hypervolume contribution range (i.e., in Figure 7(a), the normalized region needs to be as close as possible to the shaded region). Based on this consideration, for the maximization problem, the upper bound of this region is the solution itself ($\boldsymbol{a_2}$), and the lower bound (e.g., the base point) can be determined according to the following equation [19]:

$$p^j = \min\left\{s'^j \mid \mathbf{s}' \in S \backslash \{\mathbf{s}\} \text{ and } \mathbf{s}' \prec_j \mathbf{s}\right\}, \tag{12}$$

  where $j = 1, \ldots, m$, $\mathbf{s}' \prec_j \mathbf{s}$ denotes that $\boldsymbol{s'}$ dominates $\boldsymbol{s}$ in all but the jth objective. According to Equation (12), in Figure 7(b), the base point can be identified as $\boldsymbol{p} = (2, 2)$, and the normalization region is exactly the shaded part. However, if there is no point satisfying Equation Equation (12), we will directly use the reference point $\boldsymbol{r}$ as the base point $\boldsymbol{p}$.
- Step 2. Transform the base point $\boldsymbol{p}$ to $\boldsymbol{O}$ while all other points are translated accordingly:

$$a'^j = a^j - p^j, j = 1, \ldots, m. \tag{13}$$

  After this step, the coordinates in Figure 7(c) are $\boldsymbol{p} = (0, 0), \boldsymbol{a'_1} = (0, 8), \boldsymbol{a'_2} = (6, 2), \boldsymbol{a'_3} = (8, 0)$.
- Step 3. Every point undergoes a scaling transformation according to Equation (10). Different from the normalization in hypervolume approximation, the scale is decided by $\boldsymbol{a'_2}$:

$$a''^j = \frac{a'^j}{\max\{f_j(\boldsymbol{s}')\}}, j = 1, \ldots, m, \tag{14}$$

where $\boldsymbol{s}' = \boldsymbol{a}'_{\boldsymbol{2}}$, and $\max\{f_j(\boldsymbol{s}')\}$ mean the maximal value of the jth objective for $\boldsymbol{a}'_{\boldsymbol{2}}$. After this step, the coordinates in Figure 7(c) are $\boldsymbol{p} = (0, 0), \boldsymbol{a}''_{\boldsymbol{2}} = (1, 1)$.

- Step 4. Approximate the hypervolume contribution of $\boldsymbol{a}''_{\boldsymbol{2}}$ using the R2-based method, and transform the value to the final hypervolume contribution:

$$\widetilde{\mathrm{HVC}} = \prod_{i=1}^{m} \max\{f_j(\boldsymbol{s}')\} \times \widetilde{\mathrm{HVC}}', \tag{15}$$

where $\widetilde{\mathrm{HVC}}'$ is the approximated hypervolume contribution obtained after the coordinate transformation.

IV. Experiments

In this section, we evaluate the effects of the normalization in R2-based hypervolume and hypervolume contribution approximation. The experiment is conducted on the CPU of Intel(R) Core(TM) i7-8700k CPU @ 3.70GHZ. All codes are implemented in MATLAB R2022b.

A. Performance Metric

When evaluating the accuracy of hypervolume, we use the approximation error ε as a measured metric [19]:

$$\varepsilon = \left| \frac{\widetilde{\mathrm{HV}}(S, \mathbf{r}) - \mathrm{HV}(S, \mathbf{r})}{\mathrm{HV}(S, \mathbf{r})} \right| \tag{16}$$

where $\widetilde{\mathrm{HV}}$ is the hypervolume approximation value, while the HV is the exact hypervolume value. A smaller value of the approximation error means a better approximation quality.

When approximating hypervolume contribution, we use the correct identification rate (CIR) to evaluate the performance of different methods. The CIR is a metric that measures the ability to correctly identify the worst solution in a solution set. The method of approximation used here evaluates l sets of solutions, where each set has n solutions (e.g., $S_1 = \boldsymbol{a}_{\boldsymbol{1}}^{\boldsymbol{1}}, \boldsymbol{a}_{\boldsymbol{1}}^{\boldsymbol{2}}, ..., \boldsymbol{a}_{\boldsymbol{1}}^{\boldsymbol{n}}$). From each set S_i, we identify the k worst solutions (i.e., with small hypervolume contribution approximation) and call this new collection S_i^{R2}. Also, a_i^{Worst} is the worst solution based on the exact hypervolume contribution. We say we have correct identification if a_i^{Worst} is in S_i^{R2}. The CIR is the ratio of the total number of correct identifications to the total number of solution sets. In this paper, we solely consider the situation when $k = 1$.

B. Data Generation

In the experiment, we generate two different types of data. In the first type of solution set, we randomly generate different numbers of non-dominated solutions in each set. The second type is to generate the same number of solutions in each set on a fixed shape of the Pareto front. Both two types cover the 5-objective, 8-objective, and 10-objective solutions.

When generating the first type of solution set, we use the following procedure to obtain 10 different sets. Each solution set contains a different number of solutions, represented by the symbol num.

- Step 1. Randomly sample 2000 solutions within $[0, 1]$ as candidate solutions.
- Step 2. Apply non-dominated sorting to the candidate solutions. For the sorted solutions, check whether the number of non-dominated solutions is larger than num. if not then go back to Step 1.
- Step 3. Construct a solution set by randomly selecting num solutions from the non-dominated solutions obtained in Step 2.

In the second solution set, we generate solutions from m-objective maximization problems with various Pareto front shapes, including linear, convex, concave, inverted linear, inverted convex, and inverted concave. For each Pareto front, we generate 100 solution sets. Each set contains 100 solutions.

The formulations of the linear, concave, and convex Pareto fronts are as follows: $f_1+f_2+...+f_m = 1$, $f_1^2+f_2^2+...+f_m^2 = 1$, $\sqrt{f_1} + \sqrt{f_2} + ... + \sqrt{f_m} = 1$, where $f_i \in [0, 1]$ represents the objective value. For inverted cases, we invert the solution and remapped it to $[0, 1]$ using the formula $f_i = 1 - f_i$, where $i = 1, 2, ..., m$. We use the Unit Norm Vector (UNV) method to generate vectors in the R2-based method, as recommended in [20]. The reference point is fixed at $(0, ..., 0)$.

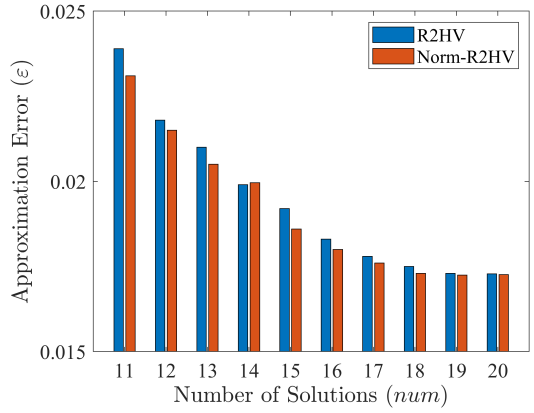
C. Experiments of Hypervolume Normalization

In the experiments of hypervolume approximation, we use both the two solution sets. We specify the number of vectors as $|\Lambda| = 1000$ for each experiment and repeat 21 times to obtain the average value.

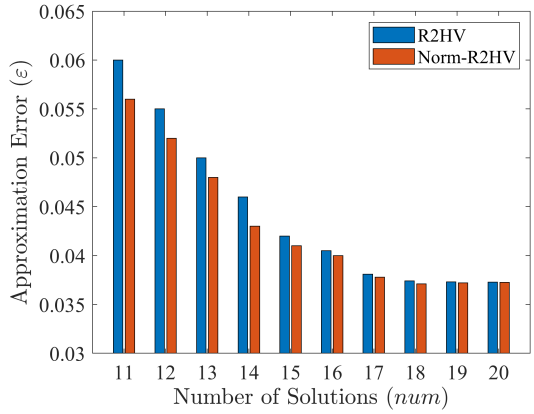
For the first type of solution set, ten different solution sets are generated in 5-objective, 8-objective, and 10-objective separately. The number of solutions (num) in each solution set varies from 11 to 20. The results are shown in Figure 8, where the x-axis represents the different solution sets with different solution numbers. The y-axis represents the approximation error (ε), where a smaller value indicates a more accurate approximation. The blue bar represents the result of the original R2-based (R2HV) and the orange bar represents the result after the normalization (Norm-R2HV). It can be observed that all the orange bar is lower than the blue bar. In other words, the normalization method does improve the accuracy of the hypervolume approximation. Another phenomenon in Figure 8 is that the accuracy of both R2HV and Norm-R2HVC decreases as num increases. And the impact of normalization yields more significant improvements when num is smaller. Comparing Figures 8(a), 8(b), 8(c), It can be observed that a smaller number of objectives corresponds to higher accuracy.

For the second type of solution set, we compare the difference between normalization (Norm-R2HV) and the original R2-based method (R2HV) in detail. Since we can easily compare the exact hypervolume, R2HV, and Norm-R2HV within one solution set, for each type of solution set (i.e., the solution set with the same Pareto Front and objective number),

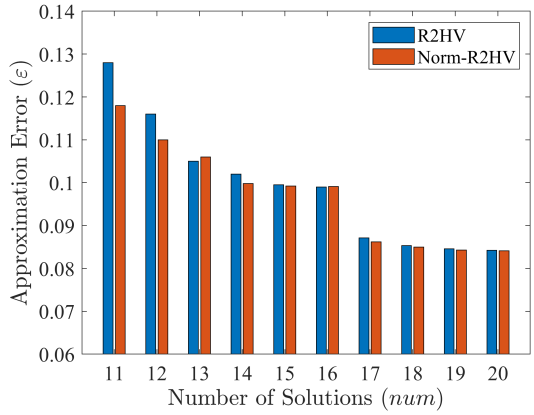
we randomly select one single set for the experiment. Results of 18 different solution sets are demonstrated in Table I (e.g., I-Concave indicates the solution set with an inverted concave shape).



(a) 5-objective problem



(b) 8-objective problem



(c) 10-objective problem

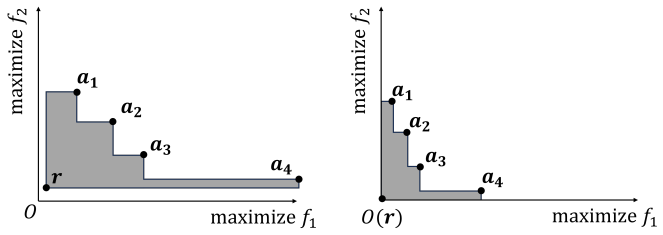
Fig. 8. Results of the first solution set in hypervolume approximation. Comparisons with the original R2-based method (R2HV) and normalization method (Norm-R2HV) in (a) 5-objective, (b) 8-objective, (c) 10-objective problem.

In Table I, we compare the performance between the normalization method and the original R2-based method on these different solution sets. The best results in each solution set are shown in bold. A direct observation is that the normalized results outperform the original R2-based method on most solution sets, which means the normalization indeed has a positive effect. Meanwhile, there are no significant differences between these two methods in some solution sets (i.e., the 10-objective inverted linear solution set). One possible reason is that these solutions have very similar scales on each objective, thus the normalization has almost no effect. However, the normalization may obtain worse results on some solution sets (i.e., 10-objective convex solution set) in Table I.

TABLE I
HYPERVOLUME APPROXIMATION OF THE NORMALIZATION METHOD (NORM-R2HV) AND THE ORIGINAL R2-BASED METHOD (R2HV) ON THE DIFFERENT SOLUTION SETS. THE WILCOXON RANK SUM TEST IS USED TO COMPARE THE PERFORMANCE, WHERE THE SYMBOL "+", "-" AND "≈" MEANS THE NORM-R2HV IS "SIGNIFICANTLY BETTER THAN", "IS SIGNIFICANTLY WORSE THAN" AND "HAS NO SIGNIFICANT DIFFERENCE WITH" THE R2HV.

Solution Set		R2HV	Norm-R2HV
Linear	5	1.627E-02	**9.124E-03** (+)
	8	4.332E-02	**2.923E-02** (+)
	10	2.742E-02	**4.181E-03** (+)
I-Linear	5	2.557E-02	**5.938E-03** (+)
	8	2.531E-02	**2.461E-02** (+)
	10	3.981E-02	**3.980E-02** (≈)
Convex	5	**4.944E-02**	4.946E-02 (≈)
	8	3.137E-02	**3.007E-02** (+)
	10	**2.134E-02**	5.704E-02 (-)
I-Convex	5	1.521E-03	**1.521E-02** (≈)
	8	4.510E-02	**4.506E-02** (+)
	10	**7.371E-03**	2.710E-02 (-)
Concave	5	2.524E-02	**2.361E-02** (+)
	8	5.812E-02	**4.424E-02** (+)
	10	4.189E-02	**2.442E-02** (+)
I-Concave	5	1.093E-01	**5.355E-02** (+)
	8	5.848E-02	**5.814E-02** (+)
	10	4.371E-02	**4.213E-02** (+)
+/-/≈			13/3/2

To explain the reason why the normalization method of some solution sets in Table I is worse, we carefully examine the distribution of the solutions in these sets. Here we take a two-objective problem as an example for illustration.



(a) The original solutions (b) The normalized solutions

Fig. 9. Example of (a) the original solutions $A = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}, \boldsymbol{a_4}\}$ and (b) these solutions after the normalization process.

In Figure 9(a), we have the four solutions $A = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}, \boldsymbol{a_4}\}$. Among those solutions, $\boldsymbol{a_4}$ has an extreme position coordinate (i.e., $f_1(\boldsymbol{a_4}) \gg f_2(\boldsymbol{a_4})$). Yet the normalization process requires transforming all the solutions according to $\max\{f_1\}$ and $\max\{f_2\}$. The distribution of

$\{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$ is very uniform originally (i.e., the values of $\max\{f_1\}$ and $\max\{f_2\}$ are very close after removing $\boldsymbol{a_4}$), while the normalization process leads to an irregular scale (shown in Figure 9(b)). However, such instances are exceedingly rare. For example, there are only two solution sets with worse normalization results, which does pose some challenges for approximating hypervolume.

TABLE II
RESULTS OF THE SECOND SOLUTION SET FOR THE HYPERVOLUME CONTRIBUTION APPROXIMATION. A COMPARISON WITH THE ORIGINAL R2-BASED METHOD (R2HVC) AND THE NORMALIZATION METHOD (NORM-R2HVC) IN DIFFERENT SOLUTION SETS. THE WILCOXON RANK SUM TEST IS USED TO COMPARE THE PERFORMANCE, WHERE THE SYMBOL "+", "-" AND "≈" MEANS THE NORM-R2HV IS "SIGNIFICANTLY BETTER THAN", "IS SIGNIFICANTLY WORSE THAN" AND "HAS NO SIGNIFICANT DIFFERENCE WITH" THE R2HV.

Solution Set		R2HVC	Norm-R2HVC
Linear	5	**0.8124**	0.7729 (-)
	8	0.6819	**0.6980** (+)
	10	0.5881	**0.6048** (+)
I-Linear	5	0.6100	**0.6681** (+)
	8	**0.4938**	0.4795 (-)
	10	**0.4543**	0.4542 (≈)
Convex	5	**0.8562**	0.8124 (-)
	8	0.7824	**0.8895** (+)
	10	**0.9210**	0.7481 (-)
I-Convex	5	0.8200	**0.8210** (+)
	8	0.8181	**0.8224** (+)
	10	0.7724	**0.8048** (-)
Concave	5	0.6148	**0.6695** (+)
	8	0.5081	**0.5114** (+)
	10	0.4857	**0.4967** (+)
I-Concave	5	0.3238	**0.3333** (+)
	8	**0.3214**	0.3090 (-)
	10	0.3038	**0.3223** (+)
+/-/≈			12/1/5

D. Experiments of Hypervolume Contribution Normalization

In the first type of solution set, each solution set has a small number of solutions ($num \in [11, 20]$), which may bring large randomness when approximating hypervolume contribution. Therefore, in this subsection, we only use the second type of solution set for the experiment. Similarly, we set the number of vectors $|\Lambda| = 1000$. Each experiment is repeated 21 times to obtain the average value.

The CIR values obtained are shown in Table II, and the normalization results (Norm-R2HVC) are compared with the original R2-based method (R2HVC), where the better result in each pair is bold. It can be noticed that a solution set with a smaller objective number can obtain better performance. The results of different shapes are quite different (i.e., the worst result for linear is above 0.6, while the inverted concave is all around 0.3).

As shown in Table II, better results are achieved with normalization on most solution sets. However, same as the hypervolume approximation in Figure 8, some results in the original R2-based method (R2HVC) outperform the normalization result (Norm-HVC) in Table II. The direct reason is shown in Figure 9 as we explained in the hypervolume approximation. Another reason is that the CIR is different from the hypervolume, which is more easily affected by a tiny change. For example, in a 5-objective linear solution set, we conduct the experiments using the original R2-based method (R2HVC), and the normalization-based method (Norm-R2HVC). The approximated values of the hypervolume contribution of these two methods and the exact values of hypervolume contribution (HVC) are shown in Table III separately.

The 10 solutions are in ascending order according to their corresponding values, where the first line is the index of these 10 solutions and the second line is the hypervolume contribution (approximated or exact value). It can be found that the hypervolume contribution obtained by the normalization method is closer to the exact value compared with the original R2-based method (i.e., the only difference is that in Norm-R2HVC, solution No. 2 is in the 5th position, while in HVC, it is in the 9th position). However, this will not affect the final result of CIR (since CIR needs to select the worst solution in the solution set, while all three methods obtained the same worst solution No. 42). In contrast, a small change in a certain value (i.e., we obtain a larger result of the solution No. 42 with the Norm-HVC) may cause the normalization method failing to select the worst solution. Thus the result of the normalization method seems worse on the hypervolume contribution approximation compared with the hypervolume approximation.

V. CONCLUSIONS

In this paper, we evaluate the effects of normalization in approximating hypervolume and hypervolume contribution with the R2-based method. We compared the normalization method with the original R2-based method. The experimental results showed that the normalization method outperformed the original version in terms of the two metrics, which indicates

TABLE III
THE WORST 10 SOLUTIONS IN ONE EXPERIMENT AMONG THE ORIGINAL R2-BASED METHOD (R2HVC), NORMALIZATION METHOD (NORM-R2HVC), AND THE EXACT HYPERVOLUME CONTRIBUTION (HVC).

Solution Index	42	81	16	43	66	98	29	68	94	9
R2HVC	3.66E-05	5.15E-05	5.43E-05	6.04E-05	6.80E-05	6.97E-05	8.01E-05	8.06E-05	8.10E-05	8.34E-05
Solution Index	42	81	16	43	98	66	29	68	2	33
Norm-R2HVC	3.28E-05	5.27E-05	5.84E-05	6.35E-05	8.06E-05	8.10E-05	8.13E-05	8.34E-05	8.50E-05	8.77E-05
Solution Index	42	81	16	43	2	98	66	29	68	33
HVC	5.86E-06	8.53E-06	9.57E-06	1.04E-05	1.33E-05	1.34E-05	1.38E-05	1.43E-05	1.44E-05	1.48E-05

the potential of using the normalization technique to decrease the approximation bias.

Our future work is to add the normalization method into the multi-objective optimization algorithms (e.g., R2HCA-EMOA [21]) and intends to further improve its performance to solve some bias-distributed problems.

REFERENCES

[1] D. A. Van Veldhuizen, *Multiobjective evolutionary algorithms: classifications, analyses, and new innovations*. Air Force Institute of Technology, 1999.

[2] C. A. Coello Coello and M. Reyes Sierra, "A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm," in *MICAI 2004: Advances in Artificial Intelligence: Third Mexican International Conference on Artificial Intelligence, Mexico City, Mexico, April 26-30, 2004. Proceedings 3*. Springer, 2004, pp. 688–697.

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

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

[5] E. Zitzler, D. Brockhoff, and L. Thiele, "The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration," in *Evolutionary Multi-Criterion Optimization: 4th International Conference, EMO 2007, Matsushima, Japan, March 5-8, 2007. Proceedings 4*. Springer, 2007, pp. 862–876.

[6] M. Emmerich, N. Beume, and B. Naujoks, "An EMO algorithm using the hypervolume measure as selection criterion," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2005, pp. 62–76.

[7] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007.

[8] S. Jiang, J. Zhang, Y.-S. Ong, A. N. Zhang, and P. S. Tan, "A simple and fast hypervolume indicator-based multiobjective evolutionary algorithm," *IEEE Transactions on Cybernetics*, vol. 45, no. 10, pp. 2202–2213, 2015.

[9] 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.

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

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

[12] R. Lacour, K. Klamroth, and C. M. Fonseca, "A box decomposition algorithm to compute the hypervolume indicator," *Computers & Operations Research*, vol. 79, pp. 347–360, 2017.

[13] J. Bader, K. Deb, and E. Zitzler, "Faster hypervolume-based search using Monte Carlo sampling," in *Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems: Proceedings of the 19th International Conference on Multiple Criteria Decision Making, Auckland, New Zealand, 7th-12th January 2008*. Springer, 2010, pp. 313–326.

[14] J. Deng and Q. Zhang, "Combining simple and adaptive Monte Carlo methods for approximating hypervolume," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 5, pp. 896–907, 2020.

[15] K. Shang, H. Ishibuchi, and X. Ni, "R2-based hypervolume contribution approximation," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 1, pp. 185–192, 2020.

[16] K. Shang, H. Ishibuchi, M.-L. Zhang, and Y. Liu, "A new R2 indicator for better hypervolume approximation," in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2018, pp. 745–752.

[17] J. Deng and Q. Zhang, "Approximating hypervolume and hypervolume contributions using polar coordinate," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 5, pp. 913–918, 2019.

[18] K. Shang, W. Chen, W. Liao, and H. Ishibuchi, "HV-Net: hypervolume approximation based on deepsets," *IEEE Transactions on Evolutionary Computation*, vol. 27, no. 4, pp. 1154–1160, 2023.

[19] K. Shang, W. Liao, and H. Ishibuchi, "HVC-Net: Deep learning based hypervolume contribution approximation," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2022, pp. 414–426.

[20] Y. Nan, K. Shang, and H. Ishibuchi, "What is a good direction vector set for the R2-based hypervolume contribution approximation," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, 2020, pp. 524–532.

[21] K. Shang and H. Ishibuchi, "A new hypervolume-based evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 5, pp. 839–852, 2020.