An Improved Version of Volume Dominance for Multi-objective Optimisation

Khoi Le, Dario Landa-Silva, and Hui Li

Automated Scheduling, Optimisation and Planning Research Group
School of Computer Science, The University of Nottingham
{kxl, jds, hzl}@cs.nott.ac.uk

Abstract. This paper proposes an improved version of volume dominance to assign fitness to solutions in Pareto-based multi-objective optimisation. The impact of this revised volume dominance on the performance of multi-objective evolutionary algorithms is investigated by incorporating it into three approaches, namely SEAMO2, SPEA2 and NSGA2 to solve instances of the 2-, 3- and 4- objective knapsack problem. The improved volume dominance is compared to its previous version and also to the conventional Pareto dominance. It is shown that the proposed improved volume dominance helps the three algorithms to obtain better non-dominated fronts than those obtained when the two other forms of dominance are used.

1   Introduction

The application of heuristic and evolutionary techniques to solve difficult real-world multi-objective optimisation problems is a very active research area. In Pareto-based multi-objective optimisation, a set of non-dominated solutions, also known as Pareto front, is sought so that the decision-maker can select the most appropriate one. Evolutionary algorithms and other population-based heuristics seem well suited to deal with Pareto based multi-objective optimisation problems because they can evolve a population of solutions towards the Pareto-optimal front in a single run. A good multi-objective evolutionary algorithm (MOEA) should be able to obtain Pareto fronts that are both well-distributed and well-converged. When designing a MOEA an important issue is how to establish superiority between solutions within the population. That is, how to compare solution fitness in a multi-objective sense. Most modern MOEAs adopt the conventional Pareto dominance relationship. There are few papers that propose different types of dominance relationship such as α-dominance, ϵ-dominance, fuzzy dominance and volume dominance (these approaches are reviewed in Section 2). These alternative forms of dominance aim to help finding solutions in difficult areas (like the extremes of the tradeoff front) or attempt to combine convergence and diversity in order to achieve better Pareto fronts in difficult problems. It has been shown that these alternative forms of dominance can help to obtain better quality Pareto fronts (e.g. [1–3]). In this paper, we present an improved

version of the *volume dominance* proposed by Le and Landa-Silva [4]. This volume dominance compares two solutions with respect to the objective space volume that each of them dominates. The revised version takes into consideration the current non-dominated front as the search progresses and incorporates a crowding technique. We compare the performance of some well-known MOEAs when using the revised volume dominance, the previous volume dominance and the conventional Pareto dominance. Our experiments are conducted using the multi-objective knapsack problem because benchmark results are available for this problem. We also show that the proposed improved volume dominance could be used within other MOEAs.

Section 2 presents a short literature review of Pareto dominance and alternative forms of dominance. Section 3 describes the new volume dominance which is a modification of the one proposed earlier in [4]. Section 4 describes our experiments to assess the impact on the performance of three MOEAs when incorporating the new volume dominance proposed here. We discuss our results in Section 5 while Section 6 gives conclusions and proposes future work.

2 Dominance Relationship

In general, the multi-objective optimisation problem with m-objectives to be maximised can be written as

$$maximise \quad \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})\} \tag{1}$$

subject to the decision vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ belonging to the feasible region S. Then, the objective vector of $\boldsymbol{x}$ is

$$\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})) \tag{2}$$

2.1 Pareto Dominance

Vilfredo Pareto proposed the concept of Pareto dominance (Pareto optimum) in 1896 [5]. Since then, this concept has been extensively used to establish the superiority between solutions in multi-objective optimisation. In Pareto dominance, a solution $\boldsymbol{x}$ is considered to be better than a solution $\boldsymbol{x}^*$ if and only if the objective vector of $\boldsymbol{x}$ dominates the objective vector of $\boldsymbol{x}^*$. More formally:

Pareto Dominance. A solution $\boldsymbol{x} \in S$ dominates a solution $\boldsymbol{x}^* \in S$ ($\boldsymbol{x} \succeq \boldsymbol{x}^*$) if and only if $\boldsymbol{x}$ is not worse than $\boldsymbol{x}^*$ in all objectives ($f_i(\boldsymbol{x}) \geq f_i(\boldsymbol{x}^*) \, \forall i = 1, \ldots, m$) and $\boldsymbol{x}$ is strictly better than $\boldsymbol{x}^*$ in at least one objective ($f_i(\boldsymbol{x}) > f_i(\boldsymbol{x}^*)$ for at least one $i = 1, \ldots, m$).

We can also distinguish between *weak dominance* and *strong dominance* [3] or *loose dominance* and *strict dominance* [6] respectively.

Weak dominance. This is often simply referred to as Pareto dominance. A solution $\boldsymbol{x}$ weakly dominates a solution $\boldsymbol{x}^*$ ($\boldsymbol{x} \succeq \boldsymbol{x}^*$) if $\boldsymbol{x}$ is better than $\boldsymbol{x}^*$ in at least one objective and is as good as $\boldsymbol{x}^*$ in all other objectives.

Strong dominance. A solution $\boldsymbol{x}$ strongly dominates a solution $\boldsymbol{x^*}$ ($\boldsymbol{x} \succ \boldsymbol{x^*}$) if $\boldsymbol{x}$ is strictly better than $\boldsymbol{x^*}$ in all objectives.

Non-dominance. If neither $\boldsymbol{x}$ dominates $\boldsymbol{x^*}$ nor $\boldsymbol{x^*}$ dominates $\boldsymbol{x}$ (weakly or strongly), then both solutions are said to be incomparable or mutually non-dominated. In this case, no solution is clearly preferred over the other.

Pareto-optimal front is the set F consisting of all non-dominated solutions $\boldsymbol{x}$ in the whole search space. Hence, a solution $\boldsymbol{x} \in F$ if there is no solution $\boldsymbol{x^*} \in S$ that dominates $\boldsymbol{x}$, i.e. if x is non-dominated with respect to S. A set of non-dominated solutions that approximates the Pareto optimal front is usually called current Pareto front or known Pareto front.

2.2   Alternative Forms of Dominance

Pareto dominance is widely adopted in multi-objective optimisation algorithms. Several alternative forms of dominance have been proposed recently. It has been shown that *relaxing* the conventional Pareto dominance can improve the performance of multi-objective optimisation algorithms. Some of these relaxed forms of Pareto dominance are more effective in finding solutions in the extremes of the feasible region S and in tackling optimisation problems with irregular Pareto-optimal fronts or problems for which it is difficult to generate feasible solutions.

In general, relaxed forms of Pareto dominance allow a solution $\boldsymbol{x}$ to *dominate* another solution $\boldsymbol{x^*}$ for which $\boldsymbol{x}$ does not Pareto-dominate $\boldsymbol{x^*}$. Relaxed forms of Pareto dominance include: structure domination [7], α-dominance [8], ϵ-dominance [1], extended Pareto dominance [9], the fuzzification of Pareto dominance [2,10] and contracting/expanding Pareto dominance [11]. Le and Landa-Silva proposed a relaxed form of Pareto dominance, named *volume dominance* [4]. This form of dominance is based on the volume of the objective space that a solution dominates. This property makes volume dominance distinguisable from conventional Pareto dominance and other relaxed forms of dominance which directly compare the objective vector of solutions in one way or another.

The volume dominance relationship between $\boldsymbol{x}$ and $\boldsymbol{x^*}$ is based on comparing their corresponding dominated volumes, $V(\boldsymbol{x})$ and $V(\boldsymbol{x^*})$ respectively, to a reference volume called *shared dominated volume* [4]. The dominated volume of $\boldsymbol{x}$, $V(\boldsymbol{x})$, and the shared dominated volume of $\boldsymbol{x}$ and $\boldsymbol{x^*}$, $SV(\boldsymbol{x}, \boldsymbol{x^*})$, are calculated with respect to the reference point $\boldsymbol{r} = (r_1, r_2, \ldots, r_m)$ as follows:

$$V(\boldsymbol{x}) = \prod_{i=1}^{m}(f_i(\boldsymbol{x}) - r_i) \tag{3}$$

$$SV(\boldsymbol{x}, \boldsymbol{x^*}) = \prod_{i=1}^{m}(min(f_i(\boldsymbol{x}), f_i(\boldsymbol{x^*})) - r_i) \tag{4}$$

It is said that for a ratio rSV, $\boldsymbol{x}$ volume-dominates $\boldsymbol{x^*}$ ($\boldsymbol{x} \succeq_V \boldsymbol{x^*}$) if either:

- $V(\boldsymbol{x^*}) = SV(\boldsymbol{x}, \boldsymbol{x^*})$ and $V(\boldsymbol{x}) > SV(\boldsymbol{x}, \boldsymbol{x^*})$ or
- $V(\boldsymbol{x}) > V(\boldsymbol{x^*}) > SV(\boldsymbol{x}, \boldsymbol{x^*})$ and $\frac{V(\boldsymbol{x}) - V(\boldsymbol{x^*})}{SV(\boldsymbol{x}, \boldsymbol{x^*})} > rSV$

Le and Landa-Silva clearly identified the difference between the fundamental principles of volume dominance and the $\mathcal{S}$ metric proposed by Zitzler and Thiele [12] which look very similar at first sight. They also proved that volume dominance covers Pareto dominance. This could not be the case with some other relaxed forms of Pareto dominance such as ϵ-dominance and extended Pareto dominance. Another interesting property of volume dominance is that it is normalised (see proof by Le and Landa-Silva [4]). That is, volume dominance is able to prevent bias towards some directions in cases with non-commensurable objective functions. Another crucial difference between volume dominance and other alternative forms of dominance in the literature is that volume dominance combines all objectives into a single unit vector to establish the superiority between solutions instead of directly comparing each objective in turn. This allows volume dominance to evaluate the whole objective vector to compensate improvement and detriment between objectives [4].

3 Volume Dominance

Volume dominance shows promising results when compared to the conventional Pareto dominance. Volume dominance is able to obtain results driven by different criteria such as better coverage, better size of space covered or better distribution of the objective values. This can be done by adjusting the rSV ratio [4]. However, the volume dominance presented earlier also has some drawbacks. It requires a preset and fixed reference point $\boldsymbol{r}$ in order to calculate the dominated volume of a solution. As the search progresses, the population moves away from the reference point which could lead to a significant increase in the dominated volume of a solution. Hence, volume dominance could be highly effective at the start of the search but less and less influent as the search progresses. Therefore, we propose to update the reference point to reflect the evolution of the population. In other words, the reference point is defined based on some characteristics of the current population.

Another issue of volume dominance is that it does not take into account the current Pareto front. This issue is illustrated in Figure 1. For both cases 1(a) and 1(b) in Figure 1, $\boldsymbol{x} \succeq_V \boldsymbol{x^*}$ for some ratio rSV by using the volume dominance in [4]. However, one can easily point out that $\boldsymbol{x} \succeq_V \boldsymbol{x^*}$ should not be true in 1(b) because both $\boldsymbol{x}$ and $\boldsymbol{x^*}$ seem equally good (close to the Pareto front) and should be regarded as non-volume-dominated solutions. In order to overcome this issue, we modify volume dominance to consider the current Pareto front when establishing superiority between two solutions. We propose a clustering technique as an additional feature incorporated into volume dominance.

3.1 Dynamic Reference Point

Volume dominance requires a reference point $\boldsymbol{r}$ to calculate the dominated volume of a solution. Le and Landa-Silva [4] proposed a simple strategy to define the reference point $\boldsymbol{r}$ as a fixed point, the origin of coordinates in the objective

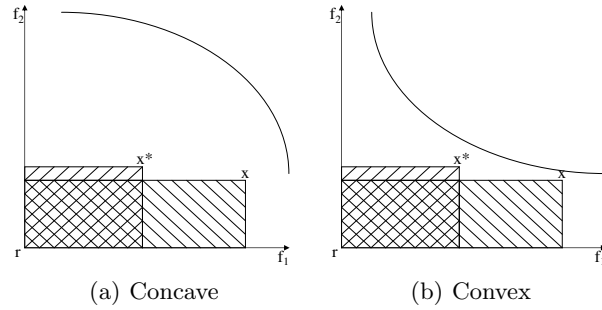(a) Concave   (b) Convex

Fig. 1. Previous Version of Volume Dominance [4]

space. As mentioned above, this simple strategy has the drawback of degrading the effectiveness of volume dominance as the search progresses because the dominated volume of solutions becomes larger and larger. We propose a more elaborate, yet more effective, strategy to estimate the reference point r. The strategy designates a reference point for each solution in the current population P and it also defines two common reference points r^{inf} and r^{sup} for all individuals in P. This strategy also considers the current state of P when determining the reference point. The reference point $\boldsymbol{r^x} = (r_1^x, r_2^x, \ldots, r_m^x)$ for solution $\boldsymbol{x} \in P$ is as follows:

$$r_i^x = f_i(\boldsymbol{x}) - (r_i^{\text{sup}} - r_i^{\text{inf}}) \tag{5}$$

where $\forall i = 1, 2, \ldots, m$

$$r_i^{\text{inf}} = \inf\{f_i(\boldsymbol{x^*}) \mid \boldsymbol{x^*} \in P\} \tag{6}$$

$$r_i^{\text{sup}} = \sup\{f_i(\boldsymbol{x^*}) \mid \boldsymbol{x^*} \in P\} \tag{7}$$

The estimation of $\boldsymbol{r^x} = (r_1^x, r_2^x, \ldots, r_m^x)$ is illustrated in Figure 2.

3.2 Considering the Current Pareto Front

Without considering the current Pareto front during the search, Figure 1 illustrates the drawback of the previous version of volume dominance. While establishing superiority between solutions $\boldsymbol{x}$ and $\boldsymbol{x^*}$, the strength of solution $\boldsymbol{x}$ is defined as the ratio of the dominated volume of $\boldsymbol{x}$ ($V(\boldsymbol{x})$) to the *shared dominated volume* of $\boldsymbol{x}$ and $\boldsymbol{x^*}$ ($SV(\boldsymbol{x}, \boldsymbol{x^*})$) with respect to the reference point $\boldsymbol{r}$. Then, the dominance of $\boldsymbol{x}$ over $\boldsymbol{x^*}$ (or vice versa) is determined based on comparing the difference between their strengths to a ratio rSV.

We propose a different approach in defining the strength of solution $\boldsymbol{x}$. The strength of $\boldsymbol{x}$ is the ratio between the dominated volume of $\boldsymbol{x}$ ($V(\boldsymbol{x})$) and the volume that *fairly* represents the status of the current Pareto set. With respect to $\boldsymbol{x}$, this *fair* representation of the current Pareto set is the subset consisting
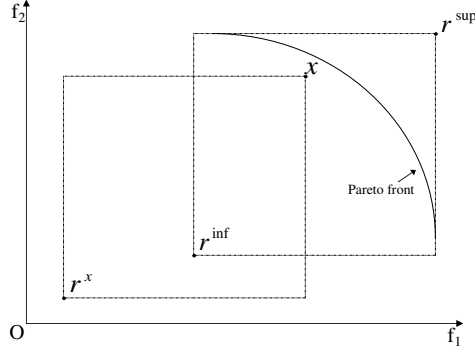
Fig. 2. Volume Dominance - Reference Point

of non-dominated solutions that Pareto-dominate $\boldsymbol{x}$. However, determining the dominated volume of a solution set could be computationally expensive. Therefore, the dominated volume of this solution set is estimated as the dominated volume of the solution that least Pareto-dominates that solution set. Let name this estimated dominated volume (w.r.t $\boldsymbol{x}$) as the reference volume of $\boldsymbol{x}$, $V^{ref}(\boldsymbol{x})$.

$$V^{ref}(\boldsymbol{x}) = \prod_{i=1}^{m}(x_i^{ref} - r_i^x) \tag{8}$$

$$x_i^{ref} = \inf\{\{f_i(\boldsymbol{x})\} \bigcup \{f_i(\boldsymbol{x^*}) \mid \boldsymbol{x^*} \succeq \boldsymbol{x} \ \wedge \ \boldsymbol{x^*} \in ParetoFront\}\} \tag{9}$$

The strength of $\boldsymbol{x}$ is then defined as follows:

$$Str(\boldsymbol{x}) = \frac{V(\boldsymbol{x})}{V^{ref}(\boldsymbol{x})} \tag{10}$$

Therefore, $\boldsymbol{x}$ volume-dominates $\boldsymbol{x^*}$ ($\boldsymbol{x} \succeq_V \boldsymbol{x^*}$) if and only if the following condition holds for a positive ratio r_{Str}:

$$Str(\boldsymbol{x}) - Str(\boldsymbol{x^*}) \geq r_{Str} \tag{11}$$

Additionally, to ensure the improvement of the Pareto set, condition (11) is relaxed whenever $Str(\boldsymbol{x}) = 1$ and $Str(\boldsymbol{x^*}) < 1$. In other words, if none of non-dominated solutions dominates $\boldsymbol{x}$, $x_i^{ref} = f_i(\boldsymbol{x})$ implying $V^{ref}(\boldsymbol{x}) = V(\boldsymbol{x})$, and it is not the case for $\boldsymbol{x^*}$ then $\boldsymbol{x} \succeq_V \boldsymbol{x^*}$.

It is noted that the current Pareto front is required in order to apply volume dominance. Therefore, one must use conventional Pareto dominance to obtain the Pareto front as well as $\boldsymbol{x^{ref}}$ to estimate the reference volume of $\boldsymbol{x}$, $V^{ref}(\boldsymbol{x})$.

3.3 Clustering Strategy

We also propose a clustering strategy as part of volume dominance to improve the distribution of non-dominated solutions w.r.t the objective space. This strategy

is only considered when two solutions $\boldsymbol{x}$ and $\boldsymbol{x}^*$ are regarded as non-volume-dominated. It means that both condition (11) and its relaxation do not hold for solutions $\boldsymbol{x}$ and $\boldsymbol{x}^*$. Solution $\boldsymbol{x}$ is said to dominate solution $\boldsymbol{x}^*$ if $\boldsymbol{x}$ is in a less crowded area when comparing to $\boldsymbol{x}^*$. The degree of crowding of $\boldsymbol{x}$ is measured as the number of neighbours of $\boldsymbol{x}$ which is the number of non-dominated solutions in the current Pareto set that ϵ-dominate $\boldsymbol{x}$. The ϵ-dominance deployed here is slightly different from the one proposed by Laumanns et al. [1] and other variants of ϵ-dominance in the literature. To the best of our knowledge, variants of ϵ-dominance either use a dynamic adaptation of the ϵ value or use a different ϵ_i value for each objective i. The variant of ϵ-dominance employed in this paper takes the advantage of both approaches, a different dynamic adaptive ϵ_i value for each objective. The ϵ_i value is estimated based on the current Pareto front as follows:

$$\epsilon_i = (r_i^{\text{sup}} - r_i^{\text{inf}}) \times \mu \tag{12}$$

where r_i^{inf}, r_i^{sup} were given in (6), (7) respectively and μ is a positive constant. Within the context of volume dominance, it is said that $\boldsymbol{x}$ ϵ-dominates $\boldsymbol{x}^*$ ($\boldsymbol{x} \succeq_{\epsilon_v} \boldsymbol{x}^*$) if and only if $f_i(\boldsymbol{x}) \geq f_i(\boldsymbol{x}^*) - \epsilon_i \ \forall i = 1, \ldots, m$ and $f_i(\boldsymbol{x}) > f_i(\boldsymbol{x}^*) - \epsilon_i$ for at least one i. Then, the number of neighbours of $\boldsymbol{x}$ is defined as follows:

$$N(\boldsymbol{x}) = |\{\boldsymbol{x}^* \mid \boldsymbol{x}^* \succeq_{\epsilon_v} \boldsymbol{x} \ \wedge \ \boldsymbol{x}^* \in CurrentParetoFront\}| \tag{13}$$

It is then said that if condition (11) and its relaxation do not hold for either $(\boldsymbol{x}, \boldsymbol{x}^*)$ or $(\boldsymbol{x}^*, \boldsymbol{x})$, then $\boldsymbol{x}$ volume-dominates $\boldsymbol{x}^*$ ($\boldsymbol{x} \succeq_V \boldsymbol{x}^*$) for a positive constant τ if and only if

$$N(\boldsymbol{x}^*) - N(\boldsymbol{x}) \geq \tau \tag{14}$$

4 Experimental Design

Most alternative forms of dominance proposed in the literature attempt to search and maintain extreme points in the objective space and/or points that are difficult to obtain and maintain with Pareto dominance. Few other forms of relaxed dominance attempt to combine convergence and diversity into a single criterion when distinguishing between solutions. These relaxed forms of dominance have been proposed as an integral part of specific multi-objective optimisation algorithms (see [1–3, 7, 8, 10]) with the exception of the contracting/expanding Pareto dominance of Sato et al. [11] which was tested on one existing approach, namely NSGA2. To the best of our knowledge, none of these forms of relaxed dominance has been tested on different multi-objective algorithms.

4.1 Brief Description of the MOEAs Considered

This paper presents experimental results showing that the proposed improved volume dominance performs well on different multi-objective evolutionary algorithms such as SEAMO2 [13], SPEA2 [14] and NSGA2 [15] when solving the multi-objective knapsack problem instances proposed by Zitzler and Thiele [12].

SEAMO2 uses a steady-state population and a simple elitist replacement strategy. Each solution of the population, in turn, acts as the first parent once and a second parent is chosen at random. Offspring is produced by applying cycle crossover on the two parents followed by a single mutation. If the offspring's objective vector improves on any *best-so-far* objective function, it replaces one of the parents and the objective's *best-so-far* is updated. Otherwise, if the offspring dominates one of the parents, it replaces that parent (unless it is a duplicate, then the offspring is deleted). If neither the offspring dominates the parents nor the parents dominate the offspring, the offspring replaces a random solution in the population that the offspring dominates. SPEA2 employs a fixed size archive to store non-dominated solutions in addition to a population. SPEA2 deploys a fine-grained fitness assignment strategy which for each individual p takes into account the number of other individuals that dominate p and that are dominated by p. A nearest neighbour density estimation for environmental selection is used to deal with two situations: when either the archive is too small or too large. The best dominated individuals in the previous archive and the population are copied to the new archive in the first case. In the latter situation, non-dominated individuals in the archive are iteratively removed until the archive's size is not exceeded. The removal of non-dominated individuals from the archive is carefully managed by using an archive truncation method that guarantees the preservation of boundary solutions. NSGA2 uses a fast non-dominated sorting algorithm to classify a population into different non-domination levels. NSGA2 also uses a crowding technique based on the density of solutions surrounding a particular solution to preserve the diversity of the population.

4.2   Enhancing MOEAs with Volume Dominance

All of these 3 algorithms, SEAMO2 [13], SPEA2 [14] and NSGA2 [15], were implemented according to their original description. Parameter settings for tackling the multi-objective knapsack problem with SPEA2 and NSGA2 were kindly provided by Marco Laumanns by means of email-based discussions. Then, in our experiments we replace the conventional Pareto dominance with the revised volume dominance and analyse the impact on the performance of these three algorithms. We aim to investigate the performance of the improved volume dominance within the three evolutionary approaches with minimum alteration to the original algorithms. The replacement of the conventional Pareto dominance with the volume dominance in each algorithm is described below.

In SEAMO2, we replace Pareto dominance with improved volume dominance to decide on the replacement of offspring by one of its parents or a random solution. This is the only stage where solutions are compared for dominance relationship in SEAMO2. However, this is not the case for SPEA2 and NSGA2. In both SPEA2 and NSGA2, there are three possible stages in which the improved volume dominance could be applied. These are: the fitness assignment to individuals, the environmental selection and the mating selection stages. In fact, Pareto dominance is only applied during the fitness assignment stage. The environmental selection and the mating selection stages use computed individual

fitnesses to compare solutions for superiority or dominance relationship. However, the individual fitness computed during the fitness assignment stage heavily relies on Pareto dominance. Therefore, to some extent, the environmental selection and the mating selection stage in SPEA2 and NSGA2 also involve Pareto dominance. We think that the fitness assignment strategy is an integral part in these two algorithms, SPEA2 and NSGA2. It makes more sense that we do not alter their fitness assignment strategy to preserve their main characteristics. We also suggest that the mating selection stage is less significant than the environmental selection stage. This is because the environmental selection strategy, which uses an archive truncation operator in SPEA2 and front extraction in NSGA2, decides the survival of individuals into the next generation whereas the mating selection strategy choose random parents based on a binary tournament selection operator to produce offspring. Therefore, for the preliminary investigation in this paper, we apply the improved volume dominance to the mating selection stage in SPEA2 and NSGA2. In other words, we replace the comparison of individual fitnesses with the improved volume dominance in order to decide on the superiority between individuals during the mating selection stage.

4.3 Benchmark Problems, Parameters Setting and Metrics

We use the instances with 750 items and 2, 3, and 4 objectives of the knapsack problem proposed in [12]. We executed short and long runs using different values of rSV to investigate the improved volume dominance. The population size used for the 2-, 3-, and 4-objective instances are 250, 300 and 350 individuals respectively. We use the same number of generations for a short run (500 generations) and a long run (1920 generations) as used by Zitzler et al. [14], Deb et al. [15] and Mumford [13]. For the improved volume dominance, we use 5 different values of $rSV = \{0.025, 0.05, 0.075, 0.10, 0.15\}$, $\mu = 0.01$ in equation (12) and $\tau = 5$ in inequality (14). We also replicated the results obtained by applying the previous volume dominance proposed by Le and Landa-Silva [4] with 4 different values of $rSV = \{0.10, 0.15, 0.20, 0.25\}$. We summarise and discuss the results from 30 independent runs. The results in Section 5 are based on $rSV = 0.075$ for the new improved volume dominance proposed in this paper and $rSV = 0.15$ for the previous approach proposed in [4].

We use four metrics to evaluate the non-dominated fronts produced. The first metric is the $\mathcal{S}$ hypervolume proposed by Zitzler and Thiele [12] which measures the overall size of the objective space covered by all the non-dominated solutions. Here, $\mathcal{S}$ is scaled as the percentage of the volume created by the origin and the reference point (39822, 41166), (41968, 41298, 41402), (41841, 40790, 39651, 41630) which is the sum profits of all items in each objective for 2-, 3- and 4-objective instance respectively. The boxplots in Figure 3 represent the distribution of the complement of the $\mathcal{S}$ hypervolume metric $(1-\mathcal{S})$. The vertical axes of the boxplots measure the percentage of the non-dominated objective space. The horizontal axes present Pareto Dominance (pd), the previous volume dominance (vd1) and the improved volume dominance proposed here (vd2) when applied to three different evolutionary algorithms SEAMO2 (se), SPEA2 (sp)

and NSGA2 (ns). The second metric used is the *cluster* CL_μ proposed by Wu and Azarm [16]. The CL_μ cluster metric measures the average number of *indistinct* solutions in each small grid which size is specified by $1/\mu$. The ideal case is when $CL_\mu = 1$ which means that every obtained Pareto solution is distinct. In all other cases, CL_μ is greater than 1. The higher the value of CL_μ, the more clustered the solution set is, and therefore, the less preferred the solution set. We use $\mu = 0.01$ or in other words, $1/\mu = 100$ units in the objective space. The third metric used is the average distance from the obtained non-dominated front to the approximation of the true Pareto front. The lower the value of this metric, the closer the obtained non-dominated front is to the true Pareto front. Finally, the size of the obtained non-dominated fronts is also computed. The higher the value, the better as more non-dominated solutions have been found.

(a) ks2 500 generations   (b) ks2 1920 generations

(c) ks3 500 generations   (d) ks3 1920 generations

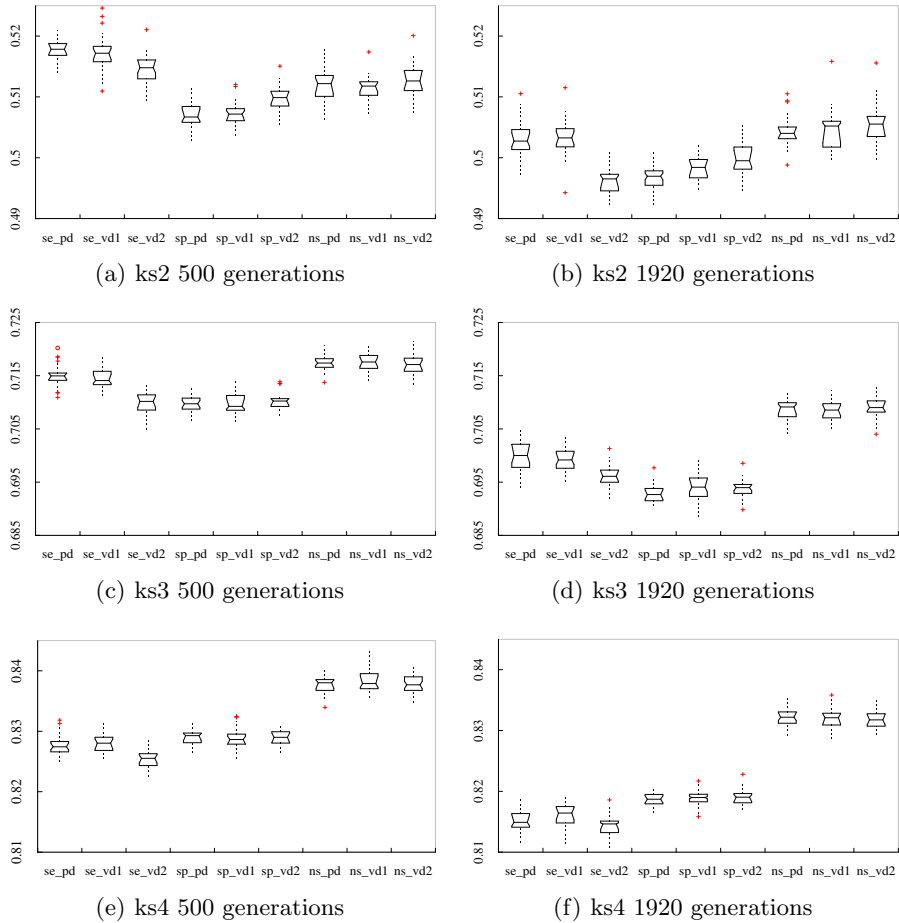(e) ks4 500 generations   (f) ks4 1920 generations

Fig. 3. Distribution of the complement of the hypervolume $\mathcal{S}$ on knapsack problems.

Table 1. Average values (standard deviation) of the size of the non-dominated set.

(a) SEAMO2

knapsack	no. of generations	se_pd	se_vd1	se_vd2
2	500	59.37 (8.1)	59.53 (8.44)	82.9 (9)
2	1920	101.67 (11.37)	104.5 (8.87)	198.9 (16.51)
3	500	199 (19.04)	202.73 (16.29)	299.87 (0.57)
3	1920	244.7 (10.54)	244.87 (11.11)	300 (0)
4	500	284.5 (15.87)	286.77 (15.79)	350 (0)
4	1920	321.27 (8.45)	316.2 (10.24)	350 (0)

(b) SPEA2

knapsack	no. of generations	sp_pd	sp_vd1	sp_vd2
2	500	76.6 (8.81)	77.27 (8.28)	70.9 (8.94)
2	1920	134.63 (10.42)	131.9 (14.22)	126.1 (14.35)
3	500	300 (0)	300 (0)	300 (0)
3	1920	300 (0)	300 (0)	300 (0)
4	500	350 (0)	350 (0)	350 (0)
4	1920	350 (0)	350 (0)	350 (0)

(c) NSGA2

knapsack	no. of generations	ns_pd	ns_vd1	ns_vd2
2	500	66 (8.29)	66.1 (8.19)	62.27 (7.73)
2	1920	85.83 (12.5)	73.2 (21.44)	73.63 (16.96)
3	500	257.37 (9.92)	255.7 (12.55)	256.47 (8.17)
3	1920	272.07 (7.52)	269 (5.38)	266.87 (6.43)
4	500	335.7 (5.64)	335.8 (6.72)	334.37 (5.88)
4	1920	338.7 (3.19)	338.9 (2.45)	338.07 (4.81)

5   Results and Discussion

For the hypervolume $\mathcal{S}$ (Figure 3), the improved volume dominance incorporated into SEAMO2 (se_vd2) outperforms not only the Pareto dominance (se_pd) but also the previous volume dominance (se_vd1) for all knapsack instances both in the short and long runs. The improved volume dominance when incorporated into NSGA2 (ns_vd2) is slightly worst than ns_pd and ns_vd1 in 2-knapsack instance but ns_vd2 is able to compete against ns_pd and ns_vd1 in higher dimension knapsack instances (3 and 4 objectives). We observe a similar result in SPEA2 as in NSAG2 when comparing se_vd2 to se_pd and se_vd1.

We also traced the progress of $\mathcal{S}$ during the evolutionary search. Regarding SPEA2 and NSGA2, vd2, vd1 and pd perform quite similar. However, vd2 incorporated into SEAMO2 is better than vd1 and pd for all 3 knapsack instances from the early stage of the evolutionary search. We omit the graphs related to this results due to space limitations in this paper. The complete results can be obtained on request.

Table 2. Average values (standard deviation) of the cluster metrics CL_μ.

(a) SEAMO2

knapsack	no. of generations	se_pd	se_vd1	se_vd2
2	500	5.78 (0.79)	5.69 (0.7)	6.1 (0.8)
2	1920	6.16 (1.05)	6.49 (0.94)	8.74 (0.74)
3	500	7.23 (1.08)	7.25 (0.82)	3.95 (0.38)
3	1920	6.6 (0.71)	6.29 (0.6)	3.31 (0.23)
4	500	5.8 (0.66)	6.1 (0.63)	3.06 (0.31)
4	1920	5.2 (0.35)	5.29 (0.43)	2.77 (0.16)

(b) SPEA2

knapsack	no. of generations	sp_pd	sp_vd1	sp_vd2
2	500	4.15 (0.45)	4.34 (0.52)	4.32 (0.54)
2	1920	6.17 (0.56)	6.16 (0.61)	6.34 (0.72)
3	500	2.09 (0.16)	2.14 (0.25)	2.16 (0.21)
3	1920	1.66 (0.09)	1.69 (0.08)	1.72 (0.09)
4	500	1.45 (0.06)	1.43 (0.06)	1.46 (0.06)
4	1920	1.36 (0.05)	1.38 (0.06)	1.35 (0.05)

(c) NSGA2

knapsack	no. of generations	ns_pd	ns_vd1	ns_vd2
2	500	4.52 (0.57)	4.52 (0.55)	4.48 (0.46)
2	1920	5.04 (0.69)	4.26 (1.02)	4.42 (0.89)
3	500	2.99 (0.25)	3 (0.3)	3.22 (0.36)
3	1920	1.6 (0.07)	1.58 (0.07)	1.6 (0.07)
4	500	1.97 (0.16)	1.9 (0.13)	1.99 (0.17)
4	1920	1.78 (0.14)	1.77 (0.13)	1.82 (0.12)

The performance of vd2 when incorporated in SPEA2 and NSGA2 is quite similar to vd1 and pd with respect to the size of the non-dominated set (Table 1(b), 1(c)) and with respect to the cluster CL_μ of the non-dominated set (Table 2(b), 2(c)). However, for SEAMO2, vd2 is noticably better than vd1 and pd regarding both the size of the non-dominated set (Table 1(a)) and the cluster CL_μ of the non-dominated set (Table 2(a)) in almost all knapsack instances for both short and long runs, except for the 2-knapsack instance when vd2 is worse than vd1 and pd in term of the cluster CL_μ of the non-dominated set.

We should point out here that se_vd2, comparing to se_vd1 and se_pd, is able not only to find more non-dominated solutions but also to reduce the clustering in the non-dominated set. In other words, se_vd2 is able to obtain more diverse solution sets and better extreme solutions. We believe that this promising result is due to the clustering strategy deployed in the improved volume dominance but further experimentation is required.

The average distance of the non-dominated set found by vd2 when incorporated into SEAMO2 is higher than when using vd1 and pd. We argue that this is because se_vd2 is able to find more extreme solutions than se_vd1 and se_pd.

Finding more extreme solutions could be intepreted as obtaining more solutions that are slightly away from the available approximated Pareto fronts. This is the reason for se_vd2 not being competitive with se_vd1 and se_pd with respect to the average distance from the non-dominated set to the approximated Pareto front (Table 3(a)). However, Table 3(b), 3(c) show that vd2 clearly outperforms vd1 and pd, when incorporated into SPEA2 and NSGA2 in all knapsack instances for both short and long runs.

Table 3. Average values (standard deviation) of the distance from the non-dominated set to the approximation of the true Pareto Front.

(a) SEAMO2

knapsack	no. of generations	se_pd	se_vd1	se_vd2
2	500	498.88 (51.62)	491.97 (40.5)	602.42 (70.93)
2	1920	389.18 (35.34)	387.93 (34.69)	432.57 (43.93)
3	500	1381.9 (58.15)	1392.34 (68.96)	1506.43 (72.99)
3	1920	1250.8 (41.71)	1269.06 (35.85)	1318.78 (50.35)
4	500	744.97 (51.94)	733.9 (50)	808.94 (48.64)
4	1920	677.02 (55.06)	668.83 (65.52)	695.78 (33.22)

(b) SPEA2

knapsack	no. of generations	sp_pd	sp_vd1	sp_vd2
2	500	701.4 (53.14)	676.4 (54.15)	638.26 (50.76)
2	1920	466.5 (43.39)	468.89 (53.54)	450.24 (43.8)
3	500	1985.72 (86.65)	1984.3 (68.79)	1960.08 (80.6)
3	1920	1682.92 (47.05)	1687.89 (34.42)	1692.77 (42.44)
4	500	1825.8 (100.38)	1761.69 (98.71)	1769.54 (115)
4	1920	1605.61 (55.04)	1567.62 (67.82)	1571.16 (70.83)

(c) NSGA2

knapsack	no. of generations	ns_pd	ns_vd1	ns_vd2
2	500	613.85 (39.39)	623.58 (43.67)	592.18 (54.07)
2	1920	429 (42.49)	454.39 (39.32)	429.03 (50.83)
3	500	2029.53 (80.73)	2026.49 (90.53)	1933.72 (99.73)
3	1920	1739.44 (74.74)	1729.7 (77.4)	1687.68 (69.35)
4	500	1681.5 (135.55)	1711.23 (143.16)	1640.75 (110.29)
4	1920	1316.23 (95.44)	1290.11 (89.45)	1285.49 (91.13)

Figure 4 shows the offline results for the 2-knapsack instance. They are the combined non-dominated solutions from 30 runs. For better visualisation, we show the non-dominated fronts in a lower density. See that vd2, vd1 and pd are quite similar when incorporated into SPEA2 and NSGA2 (Figure 4(b), 4(c)) but Figure 4(a) shows a better performance of vd2 over vd1 and pd.
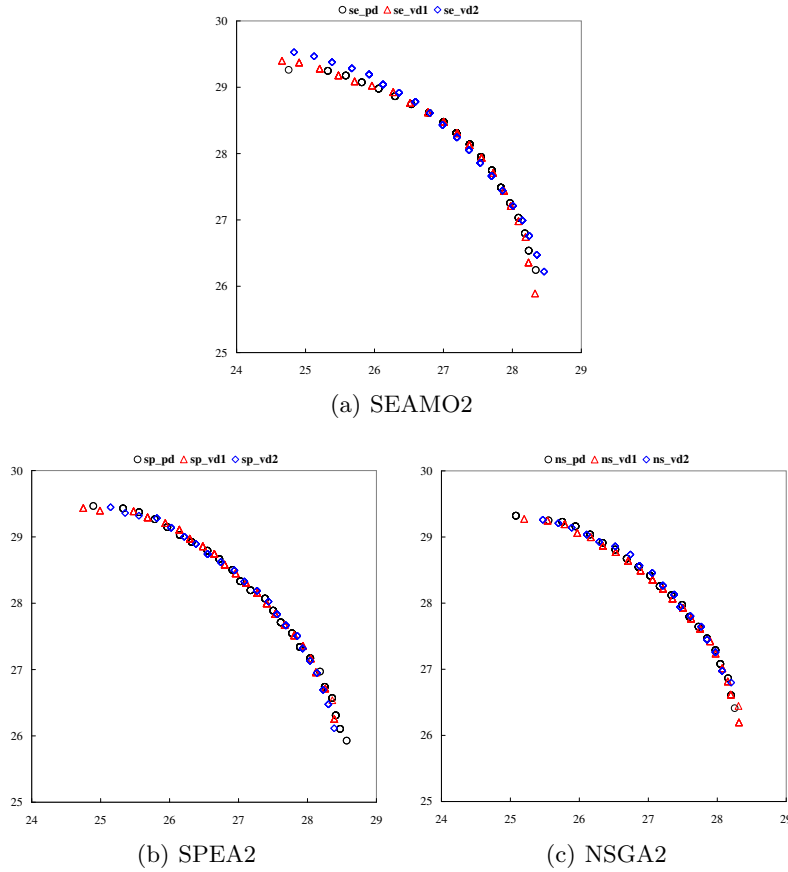
(a) SEAMO2

(b) SPEA2

(c) NSGA2

Fig. 4. Offline Result

6   Final Remarks

This paper proposed an improved volume dominance to the one originally proposed by Le and Landa-Silva [4]. We presented extensive experiments to compare the preformances of this improved volume dominance approach, the previous one and the conventional Pareto dominance using three MOEAs: SEAMO2, SPEA2 and NSGA2. The results show that in most of the cases, using 3 different knapsack instances using short and long runs, the new volume dominance approach performs better than the Pareto dominance and the previous proposed volume dominance. The improved volume dominance is more effective when incorporated into SEAMO2 than when incorporated into SPEA2 and NSGA2. This could be due to the fact that SEAMO2 is a very simple strategy whereas SPEA2 and NSGA2 already deploy more elaborate mechanisms. We believe that this revised volume dominance could be used as a new strategy to assign fitness to solutions in multi-objective evolutionary algorithms.

References

1. Laumanns, M., Thiele, L., Deb, K., Zitzler, E.: Combining Convergence and Diversity in Evolutionary Multi-Objective Optimisation. Evolution Computation **10**(3) (2002) 263–282
2. Koppen, M., Vicente-Garcia, R., Nickolay, B.: Fuzzy-Pareto-Dominance and its Application in Evolutionary Multi-Objective Optimization. In: Evolutionary Multi-Criterion Optimization 3rd Intl. Conference (EMO 2005). Volume 3410 of LNCS., Springer (2005) 399–412
3. Burke, E.K., Landa-Silva, D.: The Influence of the Fitness Evaluation Method on the Performance of Multiobjective Search Algorithms. European Journal of Operational Research **169**(3) (2006) 875–897
4. Le, K., Landa-Silva, D.: Obtaining Better Non-Dominated Sets Using Volume Dominance. In: Proceedings of the 2007 Congress on Evolutionary Computation (CEC 2007), IEEE Press (2007) 3119–3126
5. Pareto, V.: Cours D'Economie Polotique. Lausanne: F. Rouge (1896)
6. Dasgusta, P., Chakrabarti, P.P., DeSarkar, S.C.: Multiobjective Heuristic Search: An Introduction to Intelligent Search Methods for Multicriteria Optimization. Computational Intelligence (1999)
7. Yu, P.L.: Cone Convexity, Cone Extreme Points, and Nondominated Solutions in Decision Problems with Multiobjectives. Journal of Optimization Theory and Applications **14**(3) (1974) 319–377
8. Kokolo, I., Hajime, K., Shigenobu, K.: Failure of Pareto-based MOEAS: Does Non-dominated Really Mean Near to Optimal? In: Proceedings of the 2001 Congress on Evolutionary Computation (CEC 2001), IEEE Press (2001) 957–962
9. Jin, H., Wong, M.L.: Adaptive Diversity Maintenance and Convergence Guarantee in Multiobjective Evolutionary Algorithms. In: Proceedings of the 2003 Congress on Evolutionary Computation (CEC 2003), IEEE Press (2003) 2498–2505
10. Peng, J., Mok, H.M., Tse, W.: Fuzzy Dominance Based on Credibility Distributions. In: Fuzzy Systems and Knowledge Discovery. Volume 3613 of LNCS., Springer (2005) 295–303
11. Sato, H., Aguirre, H., Tanaka, K.: Controlling Dominance Area of Solutions and Its Impact of the Performance of MOEAs. In: Evolutionary Multi-Criterion Optimization 4th Intl. Conference (EMO 2007). Volume 4403 of LNCS., Springer (2007) 5–20
12. Zitzler, E., Thiele, L.: Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. IEEE Transactions on Evolutionary Computation **3**(4) (1999) 257–271
13. Mumford, C.: Simple Population Replacement Strategies for a Steady-State Multiobjective Evolutionary Algorithm. In: Proceedings of the 2004 Genetic and Evolutionary Computation Conference (GECCO 2004). Volume 3102 of LNCS., Springer (2004) 1389–1400
14. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization. In: Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001). (2002) 95–100
15. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation **6**(2) (2002) 182–197
16. Wu, J., Azarm, S.: Metrics for Quality Assessment of a Multiobjective Design Optimization Solution Set. Journal of Mechanical Design **123**(1) (2001) 18–25