Examining the Performance of Three Multiobjective Evolutionary Algorithms Based on Benchmarking Problems

Konstantinos Metaxiotis, Konstantinos Liagkouras

Abstract—The objective of this study is to examine the performance of three well-known multiobjective evolutionary algorithms for solving optimization problems. The first algorithm is the Non-dominated Sorting Genetic Algorithm-II (NSGA-II), the second one is the Strength Pareto Evolutionary Algorithm 2 (SPEA-2), and the third one is the Multiobjective Evolutionary Algorithms based on decomposition (MOEA/D). The examined multiobjective algorithms are analyzed and tested on the ZDT set of test functions by three performance metrics. The results indicate that the NSGA-II performs better than the other two algorithms based on three performance metrics.

Keywords—MOEAs, Multiobjective optimization, ZDT test functions, performance metrics.

I. INTRODUCTION

Evolutionary algorithms (EAs) have been widely used in various optimization problems with considerable success. Over the last years many EAs for solving the multiobjective problem have been proposed [1].

A main characteristic of multiobjective EAs (MOEAs) is their ability to find a number of candidate solutions [2] in a single run. Moreover, the MOEAs have been proven to be able to solve complex multiobjective optimization problems (MOPs) when traditional mathematical approaches fail to do so [1].

The majority of MOEAs use the concept of Pareto optimality conditions [26]-[29]. In particular, in MOPs, it is usually impossible to find one optimal solution [3]. Instead, the MOEAs are fabricated to locate a set of near optimum points known as the Pareto optimal set. If a point belongs to the Pareto optimal set, it means that it is not possible to find a feasible candidate solution that improves one criterion without worsening one of the remaining criteria. A MOP can be described as:

\[
\text{Min } f(x) = (f_1(x), \ldots, f_m(x)), \text{ s.t. } x = (x_1, \ldots, x_n) \in D \subseteq \mathbb{R}^n
\]

(1)

Formally, the Pareto efficiency can be defined as follows:

Definition 1. If, given a solution \( \mathbf{y} \), there exists another solution \( \mathbf{x} \) such that \( \forall j = 1, \ldots, m, f_j(x) \leq f_j(y) \) and \( \exists j \in \{1, \ldots, m\} \) such that \( f_j(x) < f_j(y) \) then we will say that solution \( \mathbf{x} \) dominates solution \( \mathbf{y} \) (denoted by \( \mathbf{x} \prec \mathbf{y} \)). If \( f_j(x) \leq f_j(y) \forall j \), we will say that solution \( \mathbf{x} \) weakly dominates solution \( \mathbf{y} \) and will be denoted \( \mathbf{x} \preceq \mathbf{y} \).

Definition 2. A solution \( \mathbf{x} \in D \) is said to be Pareto efficient if and only if \( \nexists \mathbf{y} \in D \) such that \( \mathbf{y} \prec \mathbf{x} \).

Definition 3. The real Pareto optimal set will be denoted with \( P^{true} \). The image of \( P^{true} \) in the objective space is called Pareto front and it will be denoted by \( PF^{true} \).

The EAs can be divided into two distinct categories (i) Pareto dominance based [21]-[25] and (ii) aggregation based. The most popular between the two categories is the one of Pareto dominance based.

There are some other MOEAs that utilize aggregation-based approaches [4]. These techniques create scalar fitness functions. The scalarization of the objective function can be achieved by aggregating the multiple objectives with weighting factors.

Aggregating techniques were very popular in the first era of the multiobjective evolutionary history. The main advantage of aggregating techniques is the simplicity in the implementation process of these methods. MOPs can be converted to single objective problems simply with the assistance of a weighted sum method. The second advantage associated with the aggregating techniques is that there is no need for decision making as all we have to do is to choose the solution that maximizes or minimizes the objective function depending if we are dealing with a minimization or maximization problem, respectively.

Schaffer [5] proposed an alternative technique to treat multiple, conflicting objectives separately and to search for multiple non-dominated solutions concurrently in a single run. Schaffer [5] introduced the concept of speciation instead of Pareto optimality. According to this concept, the entire population is divided into several sub-populations (called speciation), and the divided sub-population is selected using a selection technique which considers only one objective function for each sub-population. In each generation, the selected speciation makes a new population which is divided into sub-populations again after mutation and crossover operations. Vector Evaluated Genetic Algorithm (VEGA) uses the concept of speciation. The main advantage of this approach is its simplicity. However, this approach presents two major weaknesses: (i) the solutions are biased towards the edge of the Pareto frontier, (ii) the algorithm is severely
affected by the objective values because selection is determined based on the values of the objective vector and not on the domination relationship.

Finally, some algorithms use niche and sharing techniques to spread the searching effort uniformly over the Pareto optimal frontier [6]. The niche and sharing technique help the algorithm to avoid the genetic drift phenomenon by forcing the searching agents not to converge to one point from the beginning of the search. The main advantage of this technique is that it spreads the solutions uniformly across the efficient frontier. On the downside, the sharing technique is affected by the scale difference severely. Also, sharing techniques seem to be opposite to the philosophy of Pareto optimality and domination [7]. By controlling the scales of each parameter, this issue can be partially alleviated.

The rest of this paper is organized as follows. In Section II, a description of the NSGA-II, the SPEA-2, and the MOEAs based on decomposition (MOEA/D) is provided. In Section III, we provide a description of the performance metrics. The experimental results are presented in Section IV, for three different algorithms for the Zitzler-Deb-Thiele's (ZDT) family of test functions with the assistance of three performance metrics. Section V analyses the experimental results and draws conclusions.

II. PRESENTATION OF THE EXAMINED MULTIOBJECTIVE ALGORITHMS

A. NSGA-II

MOEAs are stochastic searching techniques for solving complex optimization problems. One of the most popular MOEA is the NSGA, proposed by Deb et al. [8].

In NSGA [9], a ranking process is executed before the selection operation. Through the ranking process, nondominated solutions in the population are identified at each generation, and the nondominated fronts are formed. After that, the selection, crossover, and mutation operators are performed. NSGA [9] was criticized for the high computational complexity of nondominated sorting, the lack of elitism and the difficulty of achieving diversity of solutions. In 2002, Deb et al. [8] introduced an improved version of NSGA [9] known as NSGA-II [8] which not only addressed all the major issues of the previous version but also included some new features. In particular, NSGA-II uses a faster sorting procedure, an elitism preserving mechanism and a parameter-less niching operator. Fig. 1 provides the pseudocode of the NSGA-II.

B. Strength Pareto EA 2 (SPEA-2)

The Strength Pareto EA (SPEA) was introduced by Zitzler and Thiele [10]. SPEA uses an archive containing nondominated solutions previously found (the so-called external nondominated set). At each generation, nondominated individuals are copied to the external nondominated set. For each individual in this external set, a strength value is computed. The SPEA calculates the fitness of the members of the current population by computing the strengths of all external nondominated solutions that dominate it. The fitness assignment process of SPEA considers (i) the closeness of the derived solutions to the true efficient frontier and (ii) the distribution of the derived solutions. The SPEA is using the principles of Pareto dominance to make sure that the solutions are distributed along the entire length of the Pareto front.

An improved version of Strength Pareto EA, the SPEA-II was introduced by Zitzler et al. [11]. The authors tried to develop a MOEA that eliminates the weaknesses of its predecessor (SPEA) and to incorporate most recent developments. In particular, the SPEA2 incorporates a new fitness assignment scheme which, for each solution, considers how many solutions it dominates and at the same time it is dominated by. Moreover, the SPEA2 incorporates a nearest neighbor estimation technique which provides a more accurate control of the search process. Finally, the SPEA2 incorporates a novel archive truncation technique that assists in preservation of boundary solutions.

C. MOEA Based on Decomposition (MOEA/D)

The MOEA based on Decomposition (MOEA/D) [12] and the MOEA/D with Dynamical Resource Allocation (MOEA/D-DRA) [13] are using the decomposition principle. The MOEA/D [12] uses a decomposition mechanism for converting the problem of approximation of the Pareto front (PF) into a set of scalar optimization problems. Mathematically the m-objective problem can be represented with the assistance of the following relationship:

\[ \text{Min } f(x) = (f_1(x), f_2(x), \ldots, f_m(x)) \]  

where \( f(x) \) is a m-dimensional objective vector, \( f_i(x) \) is the \( i \)-th objective to be minimized, and \( x \) is the vector of decision variables. The aforementioned minimization problem can be decomposed into a number of scalar optimization problems.
with different weight vectors. For the purposes of the present study, we use the weighted sum and the weighted Tchebycheff as introduced in [12].

The weighted sum is written using the weight vector \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \), as

\[
g_{WS}(x | \lambda) = \lambda_1 f_1(x) + \lambda_2 f_2(x) + \ldots + \lambda_m f_m(x). \tag{3}
\]

The weighted sum is to be minimized in its application to the multiobjective minimization problem.

The weighted Tchebycheff decomposition in [12] is written using the weight vector \( \lambda \) and a reference point \( z^* = (z_1^*, z_2^*, \ldots, z_m^*) \) i.e. \( z_i^* = \min \{ f_i(x) | x \in \Omega \} \) for each \( i = 1, \ldots, m \) as the scalar optimization problems of the form:

\[
g_{TE}(x | \lambda, z^*) = \min_{1,2,\ldots,m} \{ \lambda_i | f_i(x) - z_i^* \}, \text{ subject to } x \in \Omega. \tag{4}
\]

Under certain mild conditions, in each Pareto optimal point \( x^* \), there exists a weight vector, \( \lambda \) such that \( x^* \) is an efficient solution of (4). Clearly, an efficient solution of (4) is a Pareto efficient point of the fitness function \( \min g(x) = (f_1(x), \ldots, f_m(x))^T \). We can obtain different solutions by solving a number of single objective optimization problems with different weight vectors (Tchebycheff).

In MOEA/D [12], all the sub problems are treated equally and receive about the same amount of computational effort. However, a more recent study [13] introduces a different approach where processing power in each sub problem is assigned based on the difficulty of each individual problem. This new approach is called MOEA/D with a dynamical resource allocation (MOEA/D-R). The new approach estimates a utility parameter \( \pi \) for each sub problem \( i \), and thus allows the processing power to be allocated according to the utility value.

III. PERFORMANCE METRICS

A. Hypervolume

Hypervolume [14], [20], also known as S metric, is an indicator of both the convergence and diversity of an approximation set. Thus, given a set \( S \) containing \( m \) points in \( n \) objectives, the hypervolume of \( S \) is the size of the portion of objective space that is dominated by at least one point in \( S \). The hypervolume is estimated relative to a reference point which is no better than every point in \( S \) in every objective [14]. The greater the hypervolume of a solution is, the better considered the solution is.

B. Inverted Generational Distance (IGD)

The inverted generational distance (IGD) [15] can be defined as follows:

\[
IGD(P, S) = \left( \frac{\sum_{i=1}^{n} d_i^2}{|P|} \right)^{1/2}
\]

where \( d_i = \min_{x \in S} \| F(\hat{x}) - F(\hat{x}) \|, \hat{x} \in P, q = 2 \) and \( d_i \) is the smallest distance of \( \hat{x} \in P \) to the closest solutions in \( S \). The smaller the IGD value is, the better is the performance of the approach. The IGD metric is able to provide a measure for both convergence and diversity.

C. Epsilon Indicator (\( \epsilon \))

There are two versions of epsilon indicator; the multiplicative and the additive [16]. In this study, we use the unary additive epsilon indicator. The epsilon indicator of an approximation set \( A \) (\( \epsilon(A) \)) provides the minimum factor \( \epsilon \) by which each point in the real front \( R \) can be added such that the resulting transformed approximation set is dominated by \( A \). The additive epsilon indicator is a good measure of diversity, since it focuses on the worst case distance and reveals whether or not the approximation set has gaps in its trade-off solution set.

IV. PRESENTATION OF THE BENCHMARK PROBLEMS AND EXPERIMENTAL RESULTS

A. The Benchmark Problems

The Zitzler-Deb-Thiele's (ZDT) test suite is widely used for evaluating algorithms solving MOPs [1], [17]-[19]. The following three bi-objective MOPs named ZDT1, ZDT2, ZDT3 were used for comparing the NSGA-II, the Strength Pareto EA 2 (SPEA-2) and the MOEAs based on decomposition (MOEA/D). The Pareto front shapes of Zitzler-Deb-Thiele's (ZDT) test suite are convex, nonconvex, and disconnected. ZDT1, ZDT2 and ZDT3 use 30 decision variables [30].

Zitzler-Deb-Thiele’s function N.1 problem:

\[
\begin{align*}
\text{Min} &= \begin{cases} 
  f_1(x) = x_1 \\
  f_2(x) = g(x) = 1 + \frac{9}{2} \sum_{i=2}^{30} x_i \\
  h(f_1(x), g(x)) = 1 - \frac{f_1(x)}{g(x)} \\
  \text{for } 0 \leq x_i \leq 1 \text{ and } 1 \leq i \leq 30
\end{cases}
\end{align*}
\]

Zitzler-Deb-Thiele’s function N.2 problem:

\[
\begin{align*}
\text{Min} &= \begin{cases} 
  f_1(x) = x_1 \\
  f_2(x) = g(x) = 1 + \frac{9}{2} \sum_{i=2}^{30} x_i \\
  h(f_1(x), g(x)) = 1 - \left( \frac{f_1(x)}{g(x)} \right)^2 \\
  \text{for } 0 \leq x_i \leq 1 \text{ and } 1 \leq i \leq 30
\end{cases}
\end{align*}
\]

Zitzler-Deb-Thiele’s function N.3 problem

\[
\begin{align*}
\text{Min} &= \begin{cases} 
  f_1(x) = x_1 \\
  f_2(x) = g(x) = 1 + \frac{9}{2} \sum_{i=2}^{30} x_i \\
  h(f_1(x), g(x)) = 1 - \frac{f_1(x)}{g(x)} \sin(10\pi f_1(x)) \\
  \text{for } 0 \leq x_i \leq 1 \text{ and } 1 \leq i \leq 30
\end{cases}
\end{align*}
\]

B. Experimental Results

Table I presents the results of ZDT1 test function. Tables II
and III present the results for ZDT2 and ZDT3 problem respectively. Tables I-III present the mean, standard deviation (STD), median and interquartile range (IQR) of all the independent runs carried out for Hypervolume (HV) and IGD and Epsilon indicator respectively.

The higher the value of HV indicator is, the better the computed front is. The second indicator IGD \[30\] examines the convergence and diversity of solutions across the Pareto front. The smaller the value of this indicator is, the better the distribution of the solutions is. This indicator takes a zero value for an ideal distribution of the solutions in the Pareto front.

### Table I: Mean, Std, Median and IQR for HV, IGD and Epsilon metric for ZDT1 problem

<table>
<thead>
<tr>
<th>Problem: ZDT1</th>
<th>NSGA-II</th>
<th>SPEA-2</th>
<th>MOEA/D</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV. Mean and Std</td>
<td>6.39e-01</td>
<td>6.22e-01</td>
<td>1.71e-01</td>
</tr>
<tr>
<td>HV. Median and IQR</td>
<td>6.39e-01</td>
<td>6.20e-01</td>
<td>1.76e-01</td>
</tr>
<tr>
<td>IGD. Mean and Std</td>
<td>5.47e-04</td>
<td>9.95e-04</td>
<td>1.38e-02</td>
</tr>
<tr>
<td>IGD. Median and IQR</td>
<td>5.33e-04</td>
<td>1.02e-03</td>
<td>1.33e-02</td>
</tr>
<tr>
<td>Epsilon. Mean and Std</td>
<td>2.45e-02</td>
<td>5.58e-02</td>
<td>5.28e-01</td>
</tr>
<tr>
<td>Epsilon. Median and IQR</td>
<td>2.46e-02</td>
<td>4.93e-02</td>
<td>5.01e-01</td>
</tr>
</tbody>
</table>

### Table II: Mean, Std, Median and IQR for HV, IGD and Epsilon metric for ZDT2 problem

<table>
<thead>
<tr>
<th>Problem: ZDT2</th>
<th>NSGA-II</th>
<th>SPEA-2</th>
<th>MOEA/D</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV. Mean and Std</td>
<td>2.97e-01</td>
<td>8.28e-02</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>HV. Median and IQR</td>
<td>2.98e-01</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>IGD. Mean and Std</td>
<td>8.02e-04</td>
<td>1.51e-02</td>
<td>2.89e-02</td>
</tr>
<tr>
<td>IGD. Median and IQR</td>
<td>8.18e-04</td>
<td>2.10e-02</td>
<td>3.27e-02</td>
</tr>
<tr>
<td>Epsilon. Mean and Std</td>
<td>7.22e-02</td>
<td>7.43e-01</td>
<td>1.40e+00</td>
</tr>
<tr>
<td>Epsilon. Median and IQR</td>
<td>6.30e-02</td>
<td>1.03e+00</td>
<td>1.47e+00</td>
</tr>
</tbody>
</table>

### Table III: Mean, Std, Median and IQR for HV, IGD and Epsilon metric for ZDT3 problem

<table>
<thead>
<tr>
<th>Problem: ZDT3</th>
<th>NSGA-II</th>
<th>SPEA-2</th>
<th>MOEA/D</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV. Mean and Std</td>
<td>4.95e-01</td>
<td>4.83e-01</td>
<td>1.39e-01</td>
</tr>
<tr>
<td>HV. Median and IQR</td>
<td>4.95e-01</td>
<td>4.81e-01</td>
<td>1.52e-01</td>
</tr>
<tr>
<td>IGD. Mean and Std</td>
<td>7.72e-04</td>
<td>1.22e-02</td>
<td>2.01e-02</td>
</tr>
<tr>
<td>IGD. Median and IQR</td>
<td>8.04e-04</td>
<td>1.25e-02</td>
<td>1.97e-02</td>
</tr>
<tr>
<td>Epsilon. Mean and Std</td>
<td>4.42e-02</td>
<td>7.18e-02</td>
<td>7.32e-01</td>
</tr>
<tr>
<td>Epsilon. Median and IQR</td>
<td>4.26e-02</td>
<td>7.03e-02</td>
<td>7.40e-01</td>
</tr>
</tbody>
</table>

Tables IV-VI use boxplots to present graphically, the performance of the NSGA-II, the Strength Pareto EA 2 (SPEA-2) and the MOEAs based on decomposition (MOEA/D), for the three performance indicators; namely, HV, IGD, and Epsilon. Boxplot is a convenient way of depicting graphically groups of numerical data.
V. ANALYSIS OF THE RESULTS - CONCLUSIONS

We applied the NSGA-II, the Strength Pareto EA 2 (SPEA-2) and the MOEAs based on decomposition (MOEA/D) to the ZDT1-3 test functions using the following parameter specifications:

- Coding: real encoding
- Population size: \( N = 100 \)
- Termination condition: The algorithm stops after 100,000 function evaluations for each test instance.
- Number of runs for each test problem: 100 runs.
- Crossover operator for NSGA-II and SPEA-2: Simulated Binary Crossover (SBX).
- Crossover operator for MOEA/D: Differential Evolution Crossover (CR: 1.0 and F: 0.5)
- Mutation operator: Polynomial Mutation (PLM), with a mutation probability \( p_m = 1/n \).

Examining the results (Tables I-III) of the first indicator, the Hypervolume, we observe that the NSGA-II performs better than the SPEA-2 and MOEA/D for all test instances. The NSGA-II performs better than the SPEA-2 and MOEA/D with regard to the IGD metric. Finally, when examining the results regarding the Epsilon indicator, the NSGA-II outperforms the SPEA-2 and MOEA/D for all test instances.

The boxplots (Tables IV-VI) confirm the aforementioned findings. To conclude, from the analysis of the experimental results of this study, we reach the conclusion that the NSGA-II generates better results in terms of three different performance metrics compared to the SPEA-2 and MOEA/D for all examined test instances.

ACKNOWLEDGEMENT

The publication of this paper has been partly supported by the University of Piraeus Research Center.

REFERENCES


