

Obtaining Optimal Pareto Front Approximations using Scalarized Preference Information

Marlon Braun
Institute AIFB, KIT, Germany
marlon.braun@kit.edu

Pradyumn Kumar Shukla
Institute AIFB, KIT, Germany
Institut für Mathematik, TU
Clausthal, Germany
pradyumn.shukla@kit.edu

Hartmut Schmeck
Institute AIFB, KIT, Germany
hartmut.schmeck@kit.edu

ABSTRACT

Scalarization techniques are a popular method for articulating preferences in solving multi-objective optimization problems. These techniques, however, have so far proven to be ill-suited in finding a preference-driven approximation that still captures the Pareto front in its entirety. Therefore, we propose a new concept that defines an optimal distribution of points on the front given a specific scalarization function. It is proven that such an approximation exists for every real-valued problem irrespective of the shape of the corresponding front under some very mild conditions. We also show that our approach works well in obtaining an equidistant approximation of the Pareto front if no specific preference is articulated. Our analysis is complemented by the presentation of a new algorithm that implements the aforementioned concept. We provide in-depth simulation results to demonstrate the performance of our algorithm. The analysis also reveals that our algorithm is able to outperform current state-of-the-art algorithms on many popular benchmark problems.

Keywords

multi-objective optimization; scalarization method; preference-based approximation; electromagnetism-like mechanism

1. INTRODUCTION AND MOTIVATION

Multi-objective optimization focuses on finding solutions to problems that have multiple conflicting goals. Opposing aims lead to the problem not possessing a single solution attaining optimal values in each objective. Instead, we obtain a set of Pareto optimal solutions that can only be improved in one goal by deteriorating another objective at the same time. The image of the set of Pareto optimal solutions in the objective space is known as the Pareto optimal front [8].

Obtaining a closed form description of the Pareto front is often too difficult in practice [27]. Therefore, multi-objective

optimization algorithms approximate the Pareto front by a finite set of points in its entirety to enable a decision maker to make an informed choice about the solution he finally implements. Such a decision maker usually already has a certain idea about which type of solutions he prefers, for example avoiding stark tradeoffs or extreme solutions. By incorporating his preferences into a search algorithm, we can generate more options that actually represent equitable candidate implementations [8].

Scalarization techniques articulate preferences by mapping the vector of objectives to real values, thereby imposing a total order on the objective space. Many calculation methods have been suggested in the literature [22]. Usually, these techniques are only capable of identifying a single point on the Pareto front. If only one solution is generated, however, we obtain no additional information about the composition of the Pareto front. There exists no possibility to assess how the option retrieved compares to other alternatives. It is well understood in economics and psychology, however, that human preferences change depending on the alternatives available [20]. One could argue that if the decision maker is able to specify a scalarization function that adheres to his preferences, there would be no need for obtaining other options. In practice however, the scalarization function may produce results that run counter-intuitive to the decision maker's expectations. The weighted sum method, for example, is only able to obtain boundary solutions on concave Pareto fronts [22].

Some scalarization techniques offer the possibility to vary given parameters to generate multiple solutions on the Pareto front. However, manipulating these parameters yields no guarantee for obtaining an approximation that actually adheres to the decision maker's preference. The weighted sum method's inability to find solutions on concave fronts amply demonstrates this circumstance. Other scalarization techniques have been used in conjunction with minimum threshold levels, underneath which all solutions are considered to be equal. Traditional evolutionary selection mechanisms have then been applied to find an approximation of the preferred set [11, 25]. Nevertheless, choosing a threshold level can be difficult. For a given threshold, the approximation of the remaining front does not prioritize solutions whose scalarization value is closer to the global optimum.

The observations in the two previous paragraphs necessitate the development of a mechanism that obtains a preference-driven approximation of the Pareto front. In order to quantify the quality of such an approximation, we require the definition of an optimal distribution of points. Defining

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

GECCO '15, July 11 - 15, 2015, Madrid, Spain

ACM ISBN 978-1-4503-3472-3/15/07...\$15.00

DOI: <http://dx.doi.org/10.1145/2739480.2754674>

global diversity criteria has only recently seen increasing attention. Algorithms implementing traditional solving techniques such as the popular crowding distance metric [13] do not converge to an optimal distribution of points. One of the few, hence very popular global diversity criteria is hypervolume contribution [3]. Calculating hypervolume contributions however remains intractable in higher dimensions. Recent studies [28] have also shown that the optimal distribution of points depends on the chosen reference point. Reference point methods such as [21] or [12] are ill-suited for constrained problems or discontinuous Pareto fronts.

We propose an electrostatic potential energy inspired evolutionary algorithm to solve the aforementioned issues. To the best of our knowledge, [32] and [6] are till the present date the only attempts at designing an electromagnetism-inspired metaheuristic for solving multi-objective optimization problems. Both approaches assign charges to each solution based on their closeness to a randomly selected subset from an archive of non-dominated solutions. The charges are translated to force vectors, which are used to move solutions in the search space. The best solutions found are stored in an archive that uses clustering in [32] and crowding distance in [6] for pruning solutions.

The contributions of our work may be summarized in the following way:

- we propose a new mathematical notion for defining an optimal distribution of points according to a global diversity criterion. This notion can be combined with scalarized preference information to drive the distribution towards preferred subsets of the Pareto front,
- we develop a new algorithm for solving multi-objective optimization problems that implements aforementioned notion,
- a quantitative and qualitative study demonstrates the performance of our algorithm. We can show that our method outperforms current state-of-the-art algorithms in finding a diverse approximation to the Pareto front on many popular test problems.

The next section outlines some basic multi-objective optimization notions and presents the electrostatic energy based optimization concept. From this concept, we develop a multi-objective evolutionary algorithm in the subsequent section. Afterwards, the performance of our algorithm is evaluated. We first analyze its behavior under no preference articulation by assessing how its final population compares to an energy minimum distribution of points. We then compare its performance to other state-of-the-art algorithms. The evaluation section illustrates the example runs using various scalarization techniques whereas the concluding remarks at the end of the article show possible extensions.

2. AN ELECTROSTATIC ENERGY BASED OPTIMIZATION CONCEPT

Without loss of generality, we only consider minimization problems in this work. We minimize m objective functions $\mathbf{f}(\mathbf{x}) := (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ using n decision variables $\mathbf{x} := (x_1, \dots, x_n)$. The feasible search space is denoted by X and its image by \mathcal{Y} . Domination and Pareto optimality can then be defined in the following way.

DEFINITION 1 (DOMINATION [24]). Let $\mathbf{x}^1, \mathbf{x}^2 \in X$ be given. \mathbf{x}^1 dominates \mathbf{x}^2 , expressed as $\mathbf{x}^1 \succ \mathbf{x}^2$, if $f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2)$ for all $i = 1, \dots, m$ with strict inequality for at least one i .

DEFINITION 2 (PARETO OPTIMALITY [24]). A solution $\mathbf{x}^1 \in X$ is called Pareto optimal if no $\mathbf{x}^2 \in X$ exists so that $f_i(\mathbf{x}^2) \leq f_i(\mathbf{x}^1)$ for all $i = 1, \dots, m$ with strict inequality for at least one i .

We denote the set of Pareto optimal solutions by X_p and its image, the Pareto front, by \mathcal{Y}_p . Next, we formally define scalarization functions.

DEFINITION 3 (SCALARIZATION FUNCTION). A scalarization function is a map $W : \mathbb{R}^m \mapsto \mathbb{R}$.

Our approach for approximating the Pareto front is founded on the physical phenomenon of electrostatic potential energy. In physics, charged particles interact by Coulomb forces with each other. In a closed system, these Coulomb forces induce an electrostatic potential energy. The energy that a given particle exhibits with respect to another particle is equal to the product of both their charges multiplied by Coulomb's constant and divided by their Euclidean distance to each other. The sum of all pairwise energies constitutes the energy of the system. Given a finite set of charges, there exists a distribution of particles that minimizes the energy of the closed system [19].

We translate the physical system to the optimization context. The Pareto front constitutes the closed system and a finite set of points on the front $\mathcal{S} = (\mathbf{x}^1, \dots, \mathbf{x}^N)$ represents the particles, with $|\mathcal{S}| = N$ and $\mathbf{x}^i \in X_p$ for all i . Charges are induced by the scalarization function. Our aim is to find an approximation to the Pareto front that minimizes the function

$$U(\mathcal{S}) = \sum_{i=1}^N \sum_{j=i+1}^N \frac{W(\mathbf{f}(\mathbf{x}^i)) \cdot W(\mathbf{f}(\mathbf{x}^j))}{\|\mathbf{f}(\mathbf{x}^i) - \mathbf{f}(\mathbf{x}^j)\|_2}. \quad (1)$$

THEOREM 1 (EXISTENCE OF MINIMUM). Let W be continuous and $W(\mathbf{f}(\mathbf{x})) > 0$ for all $\mathbf{x} \in X_p$. Further let \mathcal{Y}_p be compact and $|\mathcal{Y}_p| > N > 1$. Then, U attains its minimum on \mathcal{Y}_p .

PROOF. The function U possesses a lower bound, since we require $W(\mathbf{f}(\mathbf{x})) > 0$ and because the denominator of U is positive. Let $\mathcal{P} := (\mathbf{y}^1, \dots, \mathbf{y}^N)$ with $\mathbf{y}^i \in \mathbb{R}^m$ for all i and $M := \{\mathcal{P} \mid U(\mathcal{P}) \leq \max_{\mathbf{x} \in X_p} N \cdot W(\mathbf{f}(\mathbf{x}))/\epsilon\}$ denote a lower level set with ϵ greater but close to zero. We obtain that the set $M \cap \mathcal{Y}_p$ is compact and that U is continuous on $M \cap \mathcal{Y}_p$, because W is continuous and the denominator of U never attains the value zero on $M \cap \mathcal{Y}_p$. Then, the conditions for the Weierstraß extreme value theorem are fulfilled [2]. \square

Remark 1. Practically all non-degenerate real-valued optimization problems fulfill the requirements of Theorem 1. Unbounded or open Pareto fronts are seldom encountered in artificial or real problems. Deterministic and evolutionary methods alike struggle to approximate an unbounded Pareto front. The result of Theorem 1 can also be extended to integer-valued or mixed integer-real-valued objective functions. However, a formal proof goes beyond the scope of this paper.

Remark 2. The case, where no particular preference information is given, corresponds to choosing $W(\mathbf{f}(\mathbf{x})) = 1$. Such a scenario would result in finding a nearly equidistant approximation to the Pareto front.¹ Theorem 1 guarantees that such an approximation exists irrespective of the shape of the front, ignoring curvature and discontinuities.

The prerequisites for the function W in Theorem 1 are met by most common scalarization techniques [22]. Negative scalarization values can be rescaled to positive values for making them compatible with our approach. For illustrating the capabilities of our concept, we focus on four particular scalarization functions in this work. The sum of objectives and the Chebyshev method [22] are among the most prominent techniques in multi-objective optimization. The Nash bargaining solution [23] is a renowned tool in economics for choosing between Pareto optimal resource allocations. This technique was originally designed for maximization problems. Therefore, we reformulated the notion to make it compatible with minimization and to avoid scalarization values of zero. Finally, proper utility [25] is a concept that has only recently emerged. It is founded on the concept of proper Pareto optimality [18, 29] and describes desirability in terms of tradeoffs. Let $I := \{1, \dots, m\}$ denote the set of objectives. The four methods are defined in the following way:

Sum of objectives:

$$W^s(\mathbf{f}(\mathbf{x})) = \sum_{i \in I} f_i(\mathbf{x}).$$

Chebyshev method:

$$W^c(\mathbf{f}(\mathbf{x})) = \max_{i \in I} (f_i(\mathbf{x}) - u_i^*)$$

with $u_i^* = \min_{\mathbf{y} \in X} f_i(\mathbf{y})$ and $\mathbf{u}^* = (u_1^*, \dots, u_m^*)$ denoting the Utopia point.

Nash bargaining solution:

$$W^n(\mathbf{f}(\mathbf{x})) = \alpha \left(\max_{\mathbf{y} \in X_p} (N(\mathbf{y})) + \min_{\mathbf{y} \in X_p} (N(\mathbf{y})) \right) - N(\mathbf{x})$$

with

$$N(\mathbf{x}) = \prod_{i \in I} \left(u_i^{ndr} - f_i(\mathbf{x}) \right) \quad (2)$$

with $\alpha > 1$, $u_i^{ndr} = \max_{\mathbf{y} \in X_p} f_i(\mathbf{y})$ and $\mathbf{u}^{ndr} = (u_1^{ndr}, \dots, u_m^{ndr})$ denoting the Nadir point.

Proper utility:

$$W^p(\mathbf{f}(\mathbf{x})) = \max_{\mathbf{y} \in X_p \setminus \{\mathbf{x}\}} \frac{\max_{i \in I} (f_i(\mathbf{x}) - f_i(\mathbf{y}))}{\max_{i \in I} (f_i(\mathbf{y}) - f_i(\mathbf{x}))}.$$

COROLLARY 1. Let \mathcal{Y}_p be compact and $|\mathcal{Y}_p| > N > 1$. Then, the function U attains its minimum on \mathcal{Y}_p if W is chosen as either the Chebyshev method, proper utility, the Nash bargaining solution or the weighted sum method for $\mathcal{Y}_p \subset \mathbb{R}_+^m$.

PROOF. Corollary 1 is a consequence of Theorem 1. \square

¹An exact equidistant approximation is not expected as points close to the boundary of the Pareto front naturally exhibit smaller energies compared to inner points, since neighboring solutions are not located in every direction.

3. ALGORITHMIC APPROACH

Finding the minimum of (1) for a given N is a highly difficult task as the problem formulation practically requires complete knowledge of the Pareto front. Additionally, the Jacobian of (1) possesses $N \cdot n$ entries. For this reason, an analytical or deterministic approach seems ill-advised to find the energy minimum in practice. This makes the evolutionary algorithm an ideal candidate for tackling the problem.

Devising an evolutionary strategy for obtaining the energy minimum provides two main challenges. First of all and obviously, the algorithm needs to converge towards the energy minimum state. This implies that the selection mechanism needs to assure the population's convergence to the Pareto front, while retaining the energy minimization focus at the same time. Secondly, once the optimal distribution has been attained, the population should remain at its current position if further search iterations are performed.

We pursue both goals by using a steady-state approach in combination with an archive mechanism. The interaction of both concepts fosters the population gradually converging towards an energy minimum state on the Pareto front, while remaining computationally tractable. We maintain a variable-sized archive A of non-dominated solutions attaining a maximum size of N that contains the currently best known minimum energy approximation. In each iteration, a new solution \mathbf{p} is created by recombining parents from the archive and mutating the resulting offspring solution.² \mathbf{p} is immediately discarded if it is dominated by or possesses the same objective values as any archive member. The newly created solution only enters the archive if either A has not reached its maximum size or if \mathbf{p} can replace an archive member such that \mathbf{p} reduces the overall energy of A .

An outline of the Electrostatic Potential Energy Evolutionary Algorithm (ESPEA) is provided in Algorithm 1. The function $e(\mathbf{a})$ calculates the energy the member \mathbf{a} introduces into the archive:

$$e(\mathbf{a}) = \sum_{\mathbf{b} \in A \setminus \{\mathbf{a}\}} \frac{W(\mathbf{f}(\mathbf{a})) \cdot W(\mathbf{f}(\mathbf{b}))}{\|\mathbf{f}(\mathbf{a}) - \mathbf{f}(\mathbf{b})\|_2}. \quad (3)$$

The energy, \mathbf{p} would introduce into the archive if it was to replace archive member \mathbf{a} , is given by

$$e_{-\mathbf{a}}(\mathbf{p}) = \sum_{\mathbf{b} \in A \setminus \{\mathbf{a}\}} \frac{W(\mathbf{f}(\mathbf{p})) \cdot W(\mathbf{f}(\mathbf{b}))}{\|\mathbf{f}(\mathbf{p}) - \mathbf{f}(\mathbf{b})\|_2}. \quad (4)$$

Step 13 encapsulates the replacement mechanism. As stated before, \mathbf{p} can only replace an archive member if it decreases the overall energy of the archive. It may occur, however, that there exist multiple candidates for replacement. We have identified three different strategies for selecting the archive member to replace in this case. Let $A_{>}$ denote the subset of A consisting only of members that introduce more energy in lieu of \mathbf{p} , i.e. $A_{>} := \{\mathbf{a} \in A | e_{-\mathbf{a}}(\mathbf{p}) < e(\mathbf{a})\}$.

Best feasible position: $\arg \min_{\mathbf{a} \in A_{>}} e_{-\mathbf{a}}(\mathbf{p})$

Of all viable options, \mathbf{p} replaces \mathbf{a} such that it introduces the least energy possible into A .

Worst in archive: $\arg \max_{\mathbf{a} \in A_{>}} e(\mathbf{a})$

The new solution replaces the member that exhibits the highest energy of all elements in $A_{>}$.

²An exact formulation of the offspring generation method applied is provided in Section 4.

Algorithm 1: ESPEA

```
1 begin
2   Generate initial population  $P$ 
3   Copy all non-dominated solutions in  $P$  to archive  $A$ 
4   repeat
5     Generate a single new solution  $\mathbf{p}$ 
6     Remove all solutions from  $A$  dominated by  $\mathbf{p}$ 
7     if  $\mathbf{p} \not\prec \mathbf{a} \wedge \mathbf{f}(\mathbf{p}) \neq \mathbf{f}(\mathbf{a}) \forall \mathbf{a} \in A$  then
8       if  $|A| < N$  then
9          $A := A \cup \{\mathbf{p}\}$ 
10      else
11        Calculate  $e(\mathbf{a})$  for all  $\mathbf{a} \in A$ 
12        Calculate  $\mathbf{e} := (e_{-\mathbf{a}^1}(\mathbf{p}), \dots, e_{-\mathbf{a}^N}(\mathbf{p}))$ 
13        update( $A, \mathbf{p}, \mathbf{e}$ )
14  until stopping criterion
15  return  $A$ 
```

Largest energy decrease: $\arg \max_{\mathbf{a} \in A} (e(\mathbf{a}) - e_{-\mathbf{a}}(\mathbf{p}))$
The energy difference before and after inserting \mathbf{p} is maximized.

The three different selection mechanisms possess several advantages and disadvantages. Best feasible position (BFP) and worst in archive (WIA) put a very strong selection pressure on including only the best solutions into the archive. On the other hand, the total energy loss might only be marginal. The largest energy decrease (LED) builds a bridge between the two aforementioned approaches. However, considering the largest decrease might not remove the least favorable archive members in the long run.

A thorough mathematical proof for the convergence behavior of ESPEA goes beyond the scope of this paper. Whether convergence occurs, also depends on the scalarization function. However, we provide a formal tool, with which the stability of an archive given W can be calculated. We call an archive *evolutionary stable* if no new solution can supersede an existing archive member.

PROPOSITION 1. Let $E^+(\mathbf{a}) := \{\mathbf{p} \in X | e_{-\mathbf{a}}(\mathbf{p}) < e(\mathbf{a})\}$ denote the set of points that would decrease the energy of A if \mathbf{p} was to supersede \mathbf{a} and $D(A) := \{\mathbf{p} \in X | \exists \mathbf{a} \in A : \mathbf{a} \succ \mathbf{p}\}$ the set of points that are dominated by the archive A . A is evolutionary stable if $\bigcup_{\mathbf{a} \in A} E^+(\mathbf{a}) \subseteq D(A)$.

PROOF. A new solution \mathbf{p} may only replace an archive member \mathbf{a} if $e_{-\mathbf{a}}(\mathbf{p}) < e(\mathbf{a})$. According to Proposition 1, this is only the case if any archive member dominates \mathbf{p} . In this case however, \mathbf{p} is automatically discarded before even being considered for insertion. \square

CONJECTURE 1. The ESPEA algorithm is evolutionary stable for the BFP, WIA and LED update mechanism if $W(\mathbf{f}(\mathbf{x})) = 1$.

Remark 3. We base Conjecture 1 on an observation made in two dimensions for an archive of size three and a Pareto front consisting of an arbitrarily-shaped, continuous curve that has fixed boundary points. In this case, the energy minimum consists of the two boundary solutions of the Pareto front \mathbf{x}^l and \mathbf{x}^r and the solution \mathbf{x}^m , whose image lies on the Pareto front and has exactly the same distance from both

boundary points. The complement of $E^+(\mathbf{x}^m)$ in \mathcal{Y} (the image of those points that would introduce more energy into A than \mathbf{x}^m) consists of two circles having their centers at $\mathbf{f}(\mathbf{x}^l)$ and $\mathbf{f}(\mathbf{x}^r)$ intersecting $\mathbf{f}(\mathbf{x}^m)$. Aforementioned space overlaps into $D(A)$ resulting in all solutions that would introduce less energy into the archive being dominated by A . We believe that this observation may be generalized for an arbitrary number of points and objectives.

We would like to point out that other preference notions such as variable cone orderings [26] or tradeoff thresholds [5], effectively cropping the set of optimal solutions, can be implemented on top of ESPEA by changing the archive update mechanism in deciding if a solution is eligible to enter. Additional preferences information regarding diversity can also be encapsulated in the scalarization function.

Next, we analyze the complexity of ESPEA. Removing dominated solutions from the archive in Step 6 and checking if \mathbf{p} is dominated by or if its objective values are equal to that of any archive member can be conducted at the same time. The worst case effort for this procedure lies in $O(mN)$ for an archive of size N . The computation effort for energy values in Step 11 and 12 hinges on the complexity of the scalarization function. As calculating the Euclidean distance requires an effort of $O(m)$, we gain a complexity of $O(N^2(m + O(W)))$ in Step 11. By exploiting the fact that energies are symmetric, we can halve the number of energy calculations, however this does not change the complexity class. An intelligent evaluation scheme in Step 12 allows us to reduce the effort for calculating the vector \mathbf{e} to $O(N(m + O(W)))$. All three update mechanism in Step 13 exhibit a computational complexity of $O(N)$. Therefore, the overall effort of a single ESPEA iteration lies in $O(N^2(m + O(W)))$. This circumstance makes the electrostatic energy approach computationally tractable in higher dimensions.

4. SIMULATION RESULTS

The computational analysis is split into two parts. In the first part, we assume no specific preferences, in order to benchmark the three archive update mechanisms presented in the previous section. We analyze how well each mechanism is suited to attain the energy minimum on a given set of widely used test problems. Thereafter, we compare ESPEA with state-of-the-art multi-objective evolutionary algorithms. The second part consists of the presentation of various representative example runs using the scalarization techniques presented in Section 2.

We implemented the ESPEA algorithm in the jMetal framework [15]. The code is publicly available and hosted as standalone project³. The following setup was used for performing the quantitative study. ESPEA employed SBX crossover [9] using a distribution index of 20 and a crossover probability of 0.9 in combination with binary tournament selection based on energy values if the archive had not reached its maximum size, yet. This way, we achieve an early selection bias towards the most promising solutions. Such a strategy was already successfully applied in the AMGA2 algorithm [31] in conjunction with the crowding distance metric. When the archive has reached its maximum size, we switch the crossover mechanism to differential evolution and select the

³<http://sourceforge.net/projects/jmetalbymarlonso/>

parents randomly. In recent years, differential evolution has proven to be one of the most competitive selection mechanisms [7]. In order to prevent an early genetic drift, when the archive is still small, we first utilize SBX. Parents are chosen randomly for the differential evolution, so that each solution on the Pareto front is given equal weight. We acknowledge, however, that there might exist better parent selection schemes that will be investigated in future studies. Polynomial mutation [10] using a distribution index of 20 and a mutation probability of one by the number of decision variables was used in all stages of the algorithm. We also used adaptive normalization in calculating Euclidean distances to mitigate the effect of different scales of individual objectives.

Reference points on the Pareto front were computed for each test problem of the study that served as approximate energy minima for a given population size. The final populations obtained by ESPEA using the different update mechanisms were compared against these reference points to assess ESPEA’s performances. The reference points were created in the following way: We first obtained analytical function representations of the Pareto fronts of the test problems. Then, equidistant decision variables were used as starting points in conjunction with the MATLAB optimization toolbox, as they serve as good estimate for the energy minimum. The reference fronts and the code for generating them are contained in the online repository.

Using the approach outlined in the paragraph above limited our selection of test instances to continuous problems, whose analytical description of the Pareto front is known. Still, our final selection of eleven problems constitutes a representative sample of differing challenges. These are composed of convex and concave fronts, shallow and pronounced curvatures and Pareto fronts exhibiting multiple bulges. We employed a maximum archive size of 50 for two-dimensional problems, while three-dimensional problems were solved using an archive size of 100. Each ESPEA configuration was run 100 times on each test problem to provide a representative sample. 10,000 function evaluations were used on the 2D and 20,000 function evaluations on the 3D problems. The following tests suites were used in this study: DEB2DK, DO2DK [4], DTLZ [14], ZDT [34] and an instance of a Lamé hypersphere [16] with a curvature of 0.5.

Table 1: Median and IQR of IGD to minimum energy distribution.

	ESPEA BFP	ESPEA WIA	ESPEA LED
DEB2DK k1	2.45e - 3 ₅ .3e-4	1.63e - 3 ₃ .3e-4	1.59e - 3 ₃ .3e-4
DEB2DK k3	5.02e - 3 ₂ .8e-3	1.54e - 3 ₂ .8e-4	1.62e - 3 ₂ .4e-4
DO2DK k2 s1	7.41e - 4 ₂ .7e-3	3.17e - 4 ₁ .1e-5	3.18e - 4 ₁ .2e-5
DO2DK k4 s1	6.57e - 4 ₂ .6e-4	3.35e - 4 ₁ .5e-5	3.34e - 4 ₁ .3e-5
ZDT1	1.12e - 2 ₁ .9e-2	1.62e - 3 ₃ .3e-4	1.63e - 3 ₂ .7e-4
ZDT2	7.62e - 2 ₂ .6e-2	1.22e - 3 ₃ .6e-4	1.27e - 3 ₃ .3e-4
DTLZ1	1.16e - 2 ₅ .8e-3	3.93e - 3 ₅ .2e-4	3.99e - 3 ₅ .0e-4
DTLZ2	9.37e - 3 ₁ .8e-3	5.47e - 3 ₂ .3e-4	5.53e - 3 ₃ .2e-4
DTLZ3	3.91e - 1 ₃ .6e-1	4.22e - 1 ₃ .9e-1	4.34e - 1 ₃ .5e-1
DTLZ4	9.44e - 3 ₅ .9e-2	6.11e - 3 ₆ .2e-2	1.22e - 2 ₆ .2e-2
Lame 0.5	3.58e - 3 ₂ .9e-4	2.56e - 3 ₁ .7e-4	2.57e - 3 ₁ .8e-4

Table 1 lists medians and inter-quartile ranges of the inverted generational distance (IGD) metric [33]. Dark/light gray colored cells hint at the best/second-best result, respectively. The IGD metric reveals how close the final archive is to the reference set that constitutes the energy minimum.

All archive update mechanisms are very capable of obtaining a close approximation to the energy minimum. DTLZ3 is the only problem, for which the inverted generational distance stays above the $10e - 2$ threshold. We observe that WIA and LED exhibit a very similar performance, whereas BFP is outclassed on most problems. Interestingly, BFP obtains very good results on the problems DTLZ3 and DTLZ4, which are known to be difficult to solve.

Table 2: IGD. p-values of a multiple comparison of mean column ranks using Bonferroni adjusted p-values.

	ESPEA WIA	ESPEA LED
ESPEA BFP	0.0019	0.2642
ESPEA WIA		0.2642

We conducted a Friedman test in conjunction with Bonferroni correction to assess, whether the performance differences observed are significant across all tested problems. Table 2 lists the results. Only WIA’s performance differs significantly from BFP. We would like to note, however, that the Bonferroni correction is very conservative and does only utilize ranks instead of absolute differences.

The distance between approximation and the ideal distribution of points is not the only performance indicator that should be considered in this context. Any approximation to the Pareto front exhibiting a small overall energy naturally constitutes a desirable result. Of course, such an approximation is required to have the full number of solution as requested by the user. The solutions retrieved should also be closely located to the Pareto front. We have already resolved the latter issue by showing that the approximations retrieved by ESPEA are close to the ideal distributions of points. Regarding the former issue, all three mechanisms were able to obtain the full number of solutions on all problems besides DTLZ3 in their median runs. However, even on DTLZ3, WIA and LED still obtained a median number of solutions of 97.5 and 92, respectively. BFP was only able to retrieve 69 solutions.

Table 3: Median and IQR of total energy values.

	ESPEA BFP	ESPEA WIA	ESPEA LED
DEB2DK k1	7.17e + 3 ₅ .7e+2	5.68e + 3 ₂ .8e+1	5.67e + 3 ₁ .8e+1
DEB2DK k3	8.27e + 3 ₉ .3e+2	5.82e + 3 ₂ .4e+1	5.82e + 3 ₂ .2e+1
DO2DK k2 s1	7.27e + 3 ₉ .0e+2	5.03e + 3 ₃ .4e+1	5.03e + 3 ₃ .2e+1
DO2DK k4 s1	7.39e + 3 ₆ .7e+2	5.11e + 3 ₃ .8e+1	5.11e + 3 ₃ .5e+1
ZDT1	7.39e + 3 ₅ .7e+2	5.81e + 3 ₃ .1e+1	5.81e + 3 ₃ .1e+1
ZDT2	1.08e + 4 ₄ .1e+3	5.78e + 3 ₂ .7e+1	5.78e + 3 ₂ .4e+1
DTLZ1	1.96e + 4 ₅ .5e+3	1.12e + 4 ₂ .3e+1	1.12e + 4 ₂ .8e+1
DTLZ2	4.82e + 4 ₈ .6e+4	9.07e + 3 ₅ .4e+1	9.06e + 3 ₄ .6e+1
DTLZ3	6.74e + 4 ₅ .3e+5	1.38e + 4 ₂ .3e+4	1.40e + 4 ₁ .3e+4
DTLZ4	4.69e + 4 _N aN	9.22e + 3 ₁ .6e+4	9.21e + 3 ₁ .6e+4
Lame 0.5	1.68e + 4 ₁ .1e+3	1.38e + 4 ₂ .2e+1	1.37e + 4 ₅ .4e+1

Table 3 lists medians and IQRs for the electrostatic potential energy values. BFP is clearly outperformed on all test instances. LED obtains the overall best results across all problems, however the actual performance difference compared to WIA is only marginal. In order to provide a complete picture, we also list the energy values of the optimal distributions calculated using the MATLAB software in Table 4. We observe that the approximations ESPEA produces come very close to the energy minima. On some test

problems the approximations even attain a smaller energy value. This does not necessarily imply, however, that the optimal distributions obtained using MATLAB are approximated too roughly. Since the solutions generated by ESPEA may not reside exactly on the Pareto front, smaller energies may be attained, as energy is mainly a dispersion metric.

Table 4: Smallest known energy values for population sizes $N = 50$ on 2D and $N = 100$ on 3D problems.

	Minimum
DEB2DK k1	$5.63e + 3$
DEB2DK k3	$5.78e + 3$
DO2DK k2 s1	$5.30e + 3$
DO2DK k4 s1	$5.37e + 3$
ZDT1	$5.75e + 3$
ZDT2	$5.75e + 3$
DTLZ1	$1.23e + 4$
DTLZ2	$9.05e + 3$
DTLZ3	$9.05e + 3$
DTLZ4	$9.05e + 3$
Lame 0.5	$1.38e + 4$

A Friedman test was conducted to assess, whether the observed performance differences are significant. The results are listed in Table 5. We observe that WIA and LED outperform BFP with high confidence. Based on these results, WIA and LED both appear as eligible options, whereas the use of BFP is discouraged. For the remainder of our study, we employed the WIA update strategy.

Table 5: ENERGY. p-values of a multiple comparison of mean column ranks using Bonferroni adjusted p-values.

	ESPEA WIA	ESPEA LED
ESPEA BFP	0.0017	0.0001
ESPEA WIA		1.0000

We opted to compare the performance of ESPEA to three state-of-the-art multi-objective evolutionary algorithms that also aim to obtain a predefined optimal distribution of points in a separate experiment. SMS-EMOA [3], MOEA/D [21] and NSGA-III [12] were chosen for this study. The first two algorithms are readily available in the jMetal framework, whereas NSGA-III was implemented from scratch. Algorithm configurations were taken from the respective original publications. MOEA/D weight vectors were obtained from the University of Essex homepage⁴. We employed 105 reference points for NSGA-III generated by the method suggested in [12]. ESPEA retained its configuration from the previous experiment. The same test problems as for the benchmark study were employed, since NSGA-III is recommended to be used with adaptive reference points on non-continuous Pareto fronts or constrained problems. Adaptive reference points, however, no longer prescribe an optimal distribution of points, which contradicts the approach of this study. We used a population size of 100 and 20,000 function evaluations on all problems.

⁴<http://dces.essex.ac.uk/staff/qzhang/MOEACOMPETITION/CEC09final/code/ZhangMOEAcode/moead030510.rar>

Table 6: Median IGD.

	ESPEA	SMS-EMOA	MOEA/D	NSGA-III
DEB2DK k1	$4.71e - 5$	$4.88e - 5$	$3.43e - 4$	$1.27e - 4$
DEB2DK k3	$4.68e - 5$	$4.91e - 5$	$3.65e - 4$	$1.33e - 4$
DO2DK k2 s1	$5.36e - 5$	$6.04e - 5$	$4.43e - 4$	$4.03e - 4$
DO2DK k4 s1	$5.43e - 5$	$7.55e - 5$	$4.33e - 4$	$2.99e - 4$
ZDT1	$4.56e - 5$	$4.30e - 5$	$3.35e - 4$	$1.22e - 4$
ZDT2	$4.43e - 5$	$5.76e - 5$	$4.24e - 4$	$1.28e - 4$
DTLZ1	$6.97e - 4$	$4.41e - 4$	$6.28e - 4$	$7.00e - 4$
DTLZ2	$6.91e - 4$	$8.33e - 4$	$7.71e - 4$	$5.95e - 4$
DTLZ3	$5.02e - 2$	$4.34e - 2$	$1.69e - 2$	$5.74e - 2$
DTLZ4	$6.97e - 4$	$8.56e - 4$	$1.00e - 3$	$6.05e - 4$
Lame 0.5	$2.64e - 4$	$1.70e - 4$	$2.92e - 4$	$2.50e - 4$

Table 6 lists median results for the IGD metric. IQR values were omitted for space limitations, but are readily available upon request. The reference fronts of DEB2DK and DO2DK were created by sampling 10,000 equidistant points and are contained in the online repository. Other reference fronts were obtained from the jMetal homepage⁵. The results clearly show that ESPEA excels on all test problems. ESPEA is among the top two performers on all but three test problems and scores five first places, the best value among all algorithms. It is noteworthy that the performance gap between ESPEA and the well established MOEA/D and NSGA-III becomes quite large on DEB2DK, DO2DK and the ZDT suite.

In the final part of this section, we present representative example runs of ESPEA on numerous test problems using the four scalarization techniques presented in Section 2. A population size of 50 and 100 were employed for 2D and 3D problems, respectively. The number of function evaluations corresponds to population size times 200.

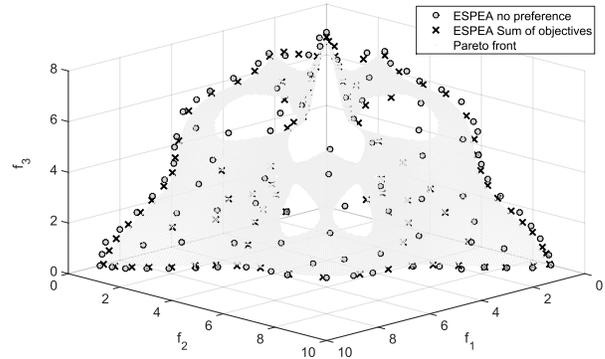


Figure 1: ESPEA example run on DEB3DK with $k = 2$ using the sum of objectives.

Figure 1 shows two example runs on the problem DEB3DK for $k = 2$ [4]. DEB3DK was omitted from our quantitative study, because its Pareto front features two distinct holes that prevent NSGA-III and MOEA/D from finding meaningful approximations without resorting to adaptive weight vectors/reference points. In order to demonstrate the effect of the sum of objectives charges, we also provide an ESPEA run using no preference information. We can observe that if no scalarization function is specified, solutions are distributed equally on the entire front. The sum of objectives introduces a strong focus on convex valleys. The

⁵<http://jmetal.sourceforge.net/>

Pareto front of DEB3DK features many convex and concave bulges. It is interesting to note that ESPEA using sum of objectives charges also captures points in concave regions, where the weighted sum method would fail.

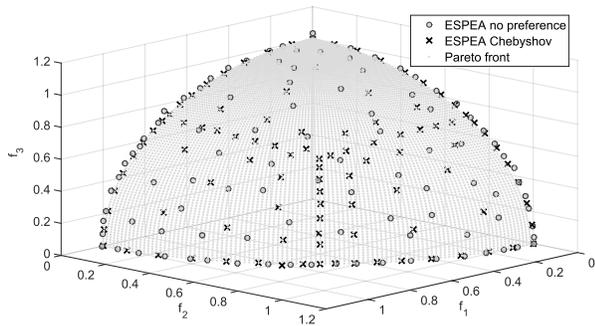


Figure 2: ESPEA example run on DTLZ4 using Chebyshev’s method.

Chebyshev’s method is illustrated in Figure 2 on DTLZ4. If no preferences are specified, solutions are equally distributed on the Pareto front. The Chebyshev technique retains the scope of the entire front, however we can spot a distinct intersection of three lines in the center of the figure. These three lines constitute exactly those points that minimize Chebyshev’s scalarization function.

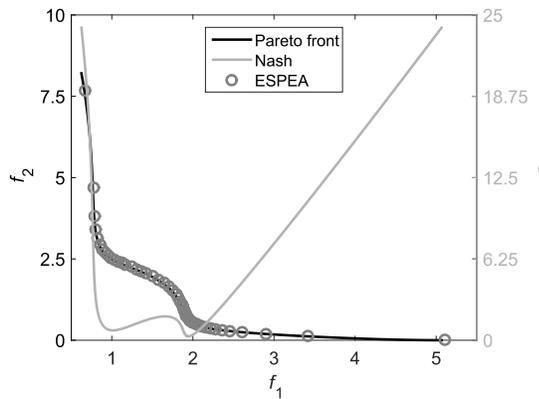


Figure 3: ESPEA example run on DO2DK with $k = 2$, $s = 1$ and $\alpha = 1.1$ using the Nash bargaining solution.

We choose DO2DK to illustrate the performance of ESPEA using the Nash bargaining solution. The current population was used to attain an estimate of the Nadir point and the minimum and maximum of (2). Figure 3 shows an ESPEA run on DO2DK and also the scalarization values for every point of the Pareto front. There is a stark focus of solutions on the barycenter of the curve as this region possesses the smallest scalarization values. Very few points are located on the near horizontal/vertical sections close to the boundary solutions. The Nash bargaining solution naturally avoids extreme points making these regions highly undesirable as indicated by the scalarization values.

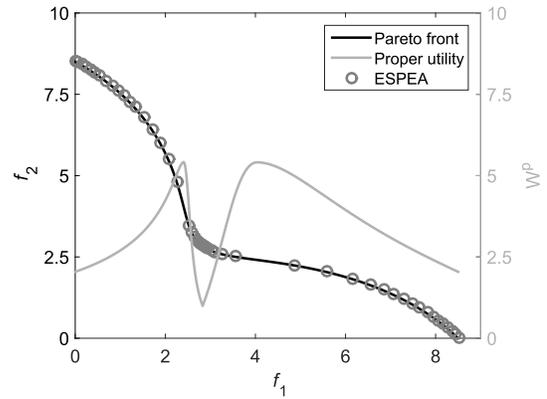


Figure 4: ESPEA example run on DEB2DK with $k = 1$ using proper utility.

Figure 4 illustrates the effect if proper utility values are set as charges. ESPEA places a heavy focus on the most favorable regions that are located at the convex valley in the center and at the boundary of the Pareto front. Utility values exhibit a sharp peak close to the global optimum, which is why ESPEA practically neglects these regions.

The search results presented in Figure 1 to 4 all present the decision maker with multiple solutions that are highly desirable with respect to his preferences. At the same time, a focus on the entire Pareto front is retained enabling him to make a comparison of his preferred solutions to other available options.

5. SUMMARY AND OUTLOOK

This paper has presented a new concept for defining an optimal distribution of points on the Pareto front. The optimal distribution can be biased towards preferred regions using scalarized preference information. We have presented the ESPEA algorithm, which implements aforementioned concept. A thorough computational study has revealed that ESPEA succeeds in obtaining these optimal distributions and compares well against state-of-the-art algorithms in finding an equidistant approximation to the Pareto front.

The electrostatic energy optimization concept offers many research questions that could be explored in subsequent studies. It would be especially of interest to formulate a coherent proof for Conjecture 1. In this light, it would make sense to further assess the evolutionary stability of ESPEA for arbitrary scalarization functions. Exact algorithms [17] for minimizing $U(S)$ could also be investigated. For this, set-based gradient of the $U(S)$ function needs to be computed [30]. Subsequent studies could also analyze ESPEA’s performance on many-objective problems. Although one of ESPEA’s great advantages is that it is parameter-free, it could make sense to employ a user-supplied parameter to strengthen or weaken the bias induced by the scalarized preference value. Apart from these, testing on real-world applications (such as [1]) is also intended.

6. REFERENCES

- [1] Allering, F., Premm, M., Shukla, P.K., Schmeck, H.: Electrical load management in smart homes using

- evolutionary algorithms. In: *Evolutionary Computation in Combinatorial Optimization*, pp. 99–110. Springer (2012)
- [2] Bertsekas, D.: *Nonlinear programming*. Athena scientific Belmont (1999)
- [3] Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research* **181**(3), 1653–1669 (2007)
- [4] Branke, J., Deb, K., Dierolf, H., Osswald, M.: Finding knees in multi-objective optimization. In: *PPSN VIII*, pp. 722–731. Springer (2004)
- [5] Braun, M.A., Shukla, P.K., Schmeck, H.: Preference ranking schemes in multi-objective evolutionary algorithms. In: R.H. Takahashi, K. Deb, E.F. Wanner, S. Greco (eds.) *EMO, LNCS*, vol. 6576, pp. 226–240. Springer (2011).
- [6] Carrasqueira, P., Alves, M.J., Antunes, C.H.: An improved multiobjective electromagnetism-like mechanism algorithm. In: *Applications of Evolutionary Computation*, pp. 627–638. Springer (2014)
- [7] Das, S., Suganthan, P.N.: Differential evolution: A survey of the state-of-the-art. *Evolutionary Computation, IEEE Transactions on* **15**(1), 4–31 (2011)
- [8] Deb, K.: *Multi-objective optimization using evolutionary algorithms*, vol. 16. John Wiley & Sons (2001)
- [9] Deb, K., Agrawal, R.B.: Simulated binary crossover for continuous search space. *Complex Systems* **9**(3), 1–15 (1994)
- [10] Deb, K., Goyal, M.: A combined genetic adaptive search (GeneAS) for engineering design. *Computer Science and Informatics* **26**, 30–45 (1996)
- [11] Deb, K., Gupta, S.: Understanding knee points in bicriteria problems and their implications as preferred solution principles. *Engineering optimization* **43**(11), 1175–1204 (2011)
- [12] Deb, K., Jain, H.: An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, part I: Solving problems with box constraints. *Evolutionary Computation* (2013)
- [13] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on* **6**(2), 182–197 (2002)
- [14] Deb, K., Thiele, L., Laumanns, M., Zitzler, E.: *Scalable test problems for evolutionary multiobjective optimization*. Springer (2005)
- [15] Durillo, J.J., Nebro, A.J.: jMetal: A java framework for multi-objective optimization. *Advances in Engineering Software* **42**, 760–771 (2011)
- [16] Emmerich, M.T., Deutz, A.H.: Test problems based on Lamé superspheres. In: *Evolutionary Multi-Criterion Optimization*, pp. 922–936. Springer (2007)
- [17] Fischer, A., Shukla, P.K.: A levenberg–marquardt algorithm for unconstrained multicriteria optimization. *Operations Research Letters* **36**(5), 643–646 (2008)
- [18] Geoffrion, A.M.: Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications* **22**, 618–630 (1968)
- [19] Halliday, D., Resnick, R., Walker, J.: *Fundamentals of physics extended*. John Wiley & Sons (2010)
- [20] Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society* pp. 263–291 (1979)
- [21] Li, H., Zhang, Q.: Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II. *Evolutionary Computation, IEEE Transactions on* **13**(2), 284–302 (2009)
- [22] Marler, R.T., Arora, J.S.: Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization* **26**(6), 369–395 (2004)
- [23] Nash, J.F.: The bargaining problem. *Econometrica: Journal of the Econometric Society* pp. 155–162 (1950)
- [24] Pareto, V.: *Cours d’économie politique*. Librairie Droz (1896)
- [25] Shukla, P., Braun, M., Schmeck, H.: Theory and algorithms for finding knees. In: R.C. Purshouse, P.J. Fleming, C.M. Fonseca, S. Greco, J. Shaw (eds.) *EMO, LNCS*, vol. 7811, pp. 156–170. Springer Berlin Heidelberg (2013).
- [26] Shukla, P.K., Braun, M.A.: Indicator based search in variable orderings: Theory and algorithms. In: R.C. Purshouse, P.J. Fleming, C.M. Fonseca, S. Greco, J. Shaw (eds.) *EMO, LNCS*, vol. 7811, pp. 66–80. Springer (2013).
- [27] Shukla, P.K., Cipold, M.P., Bachmann, C., Schmeck, H.: On homogenization of coal in longitudinal blending beds. In: *GECCO*, pp. 1199–1206. ACM (2014)
- [28] Shukla, P.K., Doll, N., Schmeck, H.: A theoretical analysis of volume based pareto front approximations. In: *GECCO*, pp. 1415–1422. ACM (2014)
- [29] Shukla, P.K., Hirsch, C., Schmeck, H.: A framework for incorporating trade-off information using multi-objective evolutionary algorithms. In: *PPSN XI*, pp. 131–140. Springer (2010)
- [30] Sosa Hernández, V.A., SchÄijtzte, O., Emmerich, M.: Hypervolume maximization via set based newtons method. In: A.A. Tantar, et al. (eds.) *EVOLVE, Advances in Intelligent Systems and Computing*, vol. 288, pp. 15–28. Springer (2014).
- [31] Tiwari, S., Fadel, G., Deb, K.: AMGA2: improving the performance of the archive-based micro-genetic algorithm for multi-objective optimization. *Engineering Optimization* **43**(4), 377–401 (2011)
- [32] Tsou, C.S., Kao, C.H.: An electromagnetism-like meta-heuristic for multi-objective optimization. In: *CEC*, pp. 1172–1178. IEEE (2006)
- [33] Van Veldhuizen, D.A., Lamont, G.B.: Evolutionary computation and convergence to a pareto front. In: *Late breaking papers at the genetic programming 1998 conference*, pp. 221–228. Citeseer (1998)
- [34] Zitzler, E., Deb, K., Thiele, L.: Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary computation* **8**(2), 173–195 (2000)