# Enhancing Decision Space Diversity in Evolutionary Multiobjective Algorithms

Ofer M. Shir[1], Mike Preuss[2], Boris Naujoks[2], and Michael Emmerich[3]

[1] Department of Chemistry, Princeton University
Frick Lab, Princeton NJ 08544, USA
oshir@Princeton.EDU

[2] Technische Universität Dortmund, Lehrstuhl für Algorithm Engineering
44221 Dortmund, Germany
{mike.preuss,boris.naujoks}@uni-dortmund.de

[3] Natural Computing Group, LIACS, Leiden University
Niels Bohrweg 1, 2333 CA Leiden, The Netherlands
emmerich@liacs.nl

**Abstract.** In multi-criterion optimization, Pareto-optimal solutions that appear very similar in the objective space may have very different pre-images. In many practical applications the decision makers, who select a solution or preferred region on the Pareto-front, may want to know different pre-images of the selected solutions. Especially, this will be the case when they would like to present alternative design candidates in later stages of a multidisciplinary design process.

In this paper we extend an existing CMA-ES niching framework, which has been previously applied successfully to multi-modal optimization, to the multi-criterion domain for boosting decision space diversity. At the same time, we introduce the concept of space aggregation for diversity maintenance in the aggregated spaces, i.e. search/decision and objective space. Empirical results on synthetic multi-modal bi-criteria test problems with known efficient sets and Pareto-fronts demonstrate that the diversity in the decision space can be significantly enhanced without hampering the convergence to a precise and diverse Pareto front approximation in the objective space of the original algorithm.

## 1 Introduction

Pareto-optimization aims at solving optimization problems with multiple, possibly conflicting, objective functions [1]. The general approach is to find non-dominated solution sets and, especially in continuous spaces, approximate true Pareto-fronts of the problem. It is important in the context of this paper to distinguish between the *Pareto-front* and the *efficient set*. While the former denotes the set of non-dominated points in the objective space, the latter refers to the set of vectors in the search space that are pre-images of the points in the Pareto-front under the mapping of the vector-valued objective function. At the same time, multiple points in the efficient set may be projected onto the same point on the Pareto-front. Moreover, unless certain continuity assumptions on

M. Ehrgott et al. (Eds.): EMO 2009, LNCS 5467, pp. 95–109, 2009.

the objective functions hold, there is no evidence that neighboring points on the Pareto-front stem from the same region of the decision space. This scenario is illustrated in Figure 1. Therefore, attaining a set of solutions that covers the entire Pareto-front does not necessarily guarantee obtaining a set that yields a good coverage of the decision set. Moreover, diversity of an approximation set to the Pareto front in the objective space does not necessarily imply diversity of solutions in its corresponding efficient set approximation, though the latter is desirable.

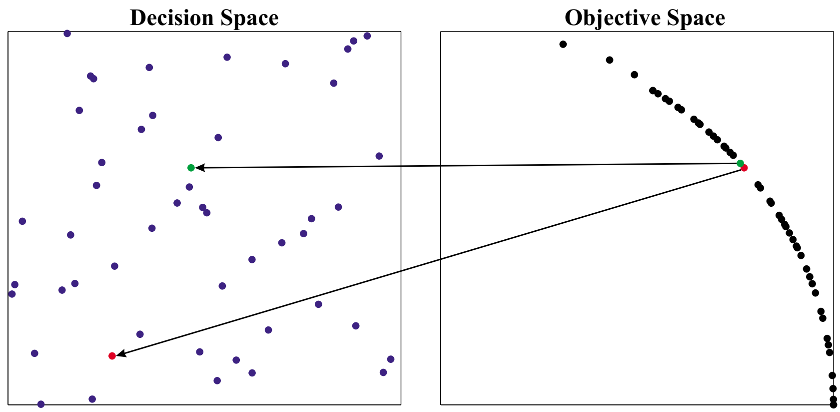


**Fig. 1.** Diversity for *decision making*: Illustrative example for a scenario where two adjacent points on the Pareto front are mapped onto two points in two completely different regions in the decision space. Units and scales are arbitrary.

### 1.1 Motivation

Indeed, it has been pointed out recently that not only high diversity of solutions in the objective space but also high diversity of solutions in the efficient set can be of interest for decision makers [2,3]. We choose to furthermore motivate this idea with the following two examples:

- Firstly, let us consider the problem of finding molecules with certain properties that can serve as drug candidates in a de novo drug discovery process [4,5]. Clearly, the approximation of different target properties can be formulated as a multi-objective optimization task. However, once a set of molecules has been found that has a good spread over the Pareto front, there may still be molecules that violate some constraints that had not been considered by the expert. In such cases, alternative solutions with similar properties, would be of interest.
- A second example is multidisciplinary optimization processes in the automotive or aerospace industries [6,7], that follow a restricted design process workflow. Here, different development teams focus on different aspects of a design and come up with a set of solutions that are favorable from the point of view of their discipline to discuss these solutions with experts from other

disciplines until a consensus design is found. Also in this case it would be desirable for the decision maker to identify different possible solutions that map to the preferred region on the Pareto front, as the objectives of the other disciplines cannot be evaluated *a priori*.

### 1.2 Background

Up to date, there has been very few work that addressed search space diversity in Evolutionary Multiobjective Optimization (see, e.g., [2,6,8,9]). Apart from this, current benchmarks do not consider this issue in the way performance is evaluated. This paper presents a wholehearted attempt to increase decision space diversity in existing state-of-the-art Evolutionary Multiobjective Optimization Algorithms (EMOA).

Based on related studies in multi-modal optimization, the modification of the selection criteria alone is not sufficient to boost diversity in the decision space. This is due to the fact that *Evolutionary Algorithms* (EAs) tend to lose their population diversity for several reasons, such as genetic drift, fast takeover, and disruptive recombination [10]. This problem is typically addressed by *Niching methods*, an extension of EAs to multi-modal optimization [11,12]. These methods allow for parallel convergence into multiple good solutions. Niching has been traditionally investigated within Genetic Algorithms (GAs) [11], but recently there were several studies of niching in Evolution Strategies (ES), especially as combined with the Covariance Matrix Adaptation ES (CMA-ES; See, e.g., [13]). The obtained ES-based niching techniques proved to be robust and efficient strategies for identifying multiple global optima in degenerate landscapes, and were successfully applied to synthetic as well as to real-world high-dimensional problems [14].

### 1.3 Overview

The new approach reported in this paper introduces two conceptual changes to the selection strategy of EMOA: The *first* is the employment of an aggregated diversity measure that takes into account the local density of solutions in the decision space with the local density in the objective space. However, aggregation alone would not be sufficient to prevent fast takeover and drift effects from occurring. These effects are already known to cause a rapid loss of diversity in ordinary EA/EMOA in early stages of the evolution, where Pareto domination rather than contribution to diversity is still the governing criterion for selection. Therefore, we consider the introduction of *dynamic niching* using *resource sharing*, also referred to as the *dynamic niching framework*, as the second element, due to counteract the aforementioned effects.

As a proof of concept for the new approach, we shall present in this paper empirical results on synthetic multi-modal bi-criteria test problems with known efficient sets and Pareto-fronts. We will demonstrate that diversity in the decision space can be significantly enhanced without hampering the convergence to a diverse Pareto front approximation in the objective space of the original algorithm.

As reference methods, we will report on the performance of the *multi-criterion* version of the CMA-ES (referred to in our notation as CMA-MO) [15], as well as the NSGA-II and its derived variants [16,2] on the same test problems.

The paper is organized as follows: In section 2 we discuss related work. The algorithmic approach is outlined in section 3. Then, in section 4, the proposed scheme is evaluated on test problems. Finally, in section 5 we summarize our findings and suggest directions for future research.

## 2 Related Work

We review here several related studies to our work. Due to the crossing-branches nature of our work, these treat the topics of *niching* and *multi-objective* optimization.

Niching techniques have been already used in the multi-objective optimization arena, earlier. Horn et al. introduced a niching technique for multi-objective optimization, known as the *niched-Pareto GA (NPGA)* [17]. The algorithm was a variant of the *fitness sharing* niching method, whereas the *niching distance metric was set to consider the objective space only*. Selection was based on so-called *Pareto domination tournaments* or on the minimal niche count, otherwise. The NPGA was a classical example of using an existing single-objective niching technique, in a straightforward manner, for multi-objective optimization – only by redefining the niching distance metric and the selection mechanism. However, its kernel was the simple GA and it lacked any self-adaptation mechanism.

A multi-objective approach aiming for a good diversity in decision as well as in objective space was the GDEA, as introduced by Toffolo and Benini [9]. GDEA invoked two selection criteria, non-dominated sorting as the primary one and a metric for decision space diversity as the secondary one.

Another approach, the so-called *Omni-Optimizer* [2], extended the classical NSGA-II [16] by considering the diversity in the decision space additionally. Its selection is performed with a changing secondary selection criterion, targeting either the decision or the objective space diversity in each generation.

An EMOA approach designed for maintaining diversity in both spaces is the KP1, as proposed by Chan and Ray [8]. Here, two criteria to measure the diversity of solutions in the corresponding spaces are defined and applied in each generation. These are the dominated hypervolume of each individual for the objective space and a neighborhood counting approach for the decision space.

A more structural analysis of the correlation between decision and objective space in multi-objective optimization has been introduced lately [3,18], while focusing on defining different test functions and analyzing the algorithmic behavior on them.

## 3 The Algorithmic Approach

Before introducing the new framework we would like to review some of its components, and in particular the extension of the CMA-ES into multi-modal domains by means of a specific niching technique.

The CMA-ES (see, e.g., [13]), is a derandomized ES variant that has been successful in treating correlations among object variables by efficiently learning matching mutation distributions. Explicitly, in generation $g$, $\lambda$ offspring are generated by means of Gaussian sampling:

$$\boldsymbol{x}_k^{(g)} \sim \mathcal{N}\left(\langle \boldsymbol{x} \rangle_W^{(g-1)}, {\sigma^{(g-1)}}^2 \mathbf{C}^{(g-1)}\right) \quad k = 1, \ldots, \lambda \tag{1}$$

The best $\mu$ search points out of these $\lambda$ offspring undergo weighted recombination and become the parent of the following generation, denoted by $\langle \boldsymbol{x} \rangle_W^{(g)}$. The covariance matrix $\mathbf{C}^{(g)}$ is initialized as the *unity matrix* and is learned during the course of evolution, based on cumulative information of successful past mutations (the so-called *evolution path*). The global step-size, $\sigma^{(g)}$, is updated based on information extracted from *principal component analysis* of $\mathbf{C}^{(g)}$ (the so-called *conjugate evolution path*). For more details we refer the reader to [13].

A *niching framework* for $\left(1 \overset{+}{,} \lambda\right)$ derandomized-ES kernels subject to a fixed niche radius has been introduced recently (see, e.g., [14]). This framework considers $q$ search points, which carry their defining strategy parameters (referred to as *CMA-Sets* or *D-Sets*), and correspond to sub-populations operating in different parts of the search space (niches). The niches and their representatives are re-formed in each generation using the dynamic peak identification (DPI) routine [14]. It takes into account both the ranked fitness of the individuals as well as the spatial distance between them; For the spatial selection, a niche radius must be defined *a priori* [14]. Individuals that belong to the same niche are located in a hyper-sphere, defined by that radius, around the central individual, namely the peak individual. Unlike previous CMA-Niching ES, this study will introduce multiple parents in each niche, subject to $(\mu_W, \lambda)$ selection with weighted recombination according to the standard formulas [19]. Sizing the niche population is done with $\lambda = 4 + \lfloor 3 \cdot \ln(n) \rfloor$, $\mu = \lfloor \frac{\lambda}{2} \rfloor$, with $n$ as the search space dimensionality, following the recommendation in [19] (for further argumentation see also [13]). To this end, we choose to define the additional selected offspring as the set of at most $\lfloor \frac{\lambda}{2} \rfloor - 1$ individuals that are within niche radius from the peak individual and share its same parent. This way, it is guaranteed that the ES mutation distribution evolves continuously. Since the value of $\mu$ may vary over time, other auxiliary coefficients must be updated accordingly, such as the recombination weights. Algorithm 1 summarizes the Niching-CMA routine.

The proposed multi-objective routine uses the Niching-CMA scheme as it is, with the following modifications:

– ranking of individuals is based upon non-dominated sorting.
– distance between niches is calculated in the aggregated space.
– the estimation of the niche radius is adjusted.

Given the $n$-dimensional decision vector of individual $k$, $\boldsymbol{x}_k = (x_{k,1}, ..., x_{k,n})$, with its assigned objective $d$-dimensional vector, $\boldsymbol{f}_k = (f_{k,1}, ..., f_{k,d})$, and given the equivalent decision and objective vectors of individual $l$, $(\boldsymbol{x}_l,\ \boldsymbol{f}_l)$, the distance between individuals $k, l$ is defined as follows:

**Algorithm 1.** $(\mu_w, \lambda)$-CMA-ES Niching with Fixed Niche Radius

```
for i = 1, ..., q search points do
    Generate λ samples based on the CMA-Set of individual i
end for
Evaluate fitness of the population
Compute the Dynamic Peak Set (DPS) with the DPI Routine
for j = 1..q elements of DPS do
    Identify at most μ = ⌊λ/2⌋ fittest individuals of niche j with Parent(peak(j))
    Apply weighted recombination on x_w and z_w w.r.t. those individuals
    Inherit the CMA-Set of peak(j) and update it w.r.t. the variations carried out
end for
if N_dps=size of DPS < q then
    Generate q − N_dps new search points, reset CMA-Sets
end if
```

$$d_{k,l} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{k,i} - x_{l,i})^2 \; + \; \frac{1}{d}\sum_{j=1}^{d}(f_{k,j} - f_{l,j})^2} \tag{2}$$

It is implicitly assumed that decision parameters and objective function values are scaled within a common order of magnitude. In order to *select individuals* based on multiple objectives, the selection mechanism was modified. As outlined before, the niches are identified based on their ranked quality, which is implemented here by means of *non-dominated sorting* [16]. Following this, the routine will proceed as usual: Starting with rank 0, a greedy identification of the niches will be carried out, considering the distance with respect to the aggregated objective and decision spaces. If not all $q$ niches are populated, the routine will proceed to rank 1, and so on.

**Comparison.** The uniqueness of the proposed approach with respect to the mainstream EMOA lies in two main aspects: Firstly, the employment of a *single selection phase*, rather than two, and secondly, the consideration of space aggregation for the sake of diversity measurement. Moreover, this method differs from the CMA-MO algorithm in its ES mechanism: Unlike the elitist single-parent $(1 + \lambda)$-kernel of the CMA-MO, the proposed scheme employs a *comma* multi-parent $(\mu_W, \lambda)$-kernel, which may be advantageous in certain environments.

**Setting a Default Value for the Niche Radius.** Since our method aims to approximate the Pareto front by populating it with a uniform distribution of $q$ niches, we can estimate the niche radius, whenever the aim is to distribute the niches evenly across the search space. The following derivations are valid for $2D$ objective spaces, but we believe that they could be generalized to $d$-dimensional spaces. Consider a connected Pareto front, and assume that we can define its *length*, denoted by $l_{FRONT}$. Also, let the diameter of the Pareto set be denoted by $l_{SET}$. Upon demanding a uniform distribution of niches, one may write:

$$2 \cdot \rho \cdot q = \sqrt{l_{FRONT}^2 + l_{SET}^2} \tag{3}$$

*Simplified Model.* One can consider a simplified model for providing an upper and a lower bounds for $\rho$, by taking into account only the objective space. For this purpose let us consider the *Nadir* objective vector, denoted here as $\boldsymbol{\zeta}^{(\mathcal{N})} = (f_{1,\mathcal{N}}, f_{2,\mathcal{N}})^T$. In the general $d$-dimensional objective space, the *Nadir* objective vector is defined as the vector with the *worst objective values of all Pareto optimal solutions* (as opposed to the worst objective values of the entire space):

$$\zeta_i^{(\mathcal{N})} = \max\left\{ f_i \,\middle|\, (f_1, \ldots, f_i, \ldots, f_d)^T \in \mathcal{F}_N \right\}. \qquad (4)$$

The Nadir objective vector can be computed for $d = 2$ by employing single-objective optimization. For $d > 2$, heuristics are available, but the problem is considered to be computationally hard [20].

Without loss of generality, assume that the objectives $\{f_1,\ f_2\}$ are assigned with values in the intervals $\{[f_{1,min}, f_{1,\mathcal{N}}],\ [f_{2,min}, f_{2,\mathcal{N}}]\}$, respectively. The length of the assumably-connected Pareto front has the following lower and upper bounds:

$$\begin{aligned} l_{FRONT,min} &= \sqrt{\left((f_{1,\mathcal{N}} - f_{1,min})^2 + (f_{2,\mathcal{N}} - f_{2,min})^2\right)} \\ l_{FRONT,max} &= |f_{1,\mathcal{N}} - f_{1,min}| + |f_{2,\mathcal{N}} - f_{2,min}| \end{aligned} \qquad (5)$$

Hence, upon assuming a uniformly spaced population of the $q$ niches along the front, one can derive

$$\frac{\sqrt{\left((f_{1,\mathcal{N}} - f_{1,min})^2 + (f_{2,\mathcal{N}} - f_{2,min})^2\right)}}{2 \cdot q} \leq \rho \leq \frac{|f_{1,\mathcal{N}} - f_{1,min}| + |f_{2,\mathcal{N}} - f_{2,min}|}{2 \cdot q} \qquad (6)$$

*The General Case.* For the general case, we choose to define the default values as the diameters of the decision or the objective spaces, respectively:

$$r_{SET} = \sqrt{\sum_{i=1}^{n} (x_{i,max} - x_{i,min})^2} \qquad r_{FRONT} = \sqrt{\sum_{j=1}^{d} (f_{j,max} - f_{j,min})^2} \qquad (7)$$

and thus

$$\rho = \frac{\sqrt{\sum_{i=1}^{n} (x_{i,max} - x_{i,min})^2 + \sum_{j=1}^{d} (f_{j,max} - f_{j,min})^2}}{2 \cdot q} \qquad (8)$$

The niche radius is essentially a crucial parameter of this method, and its estimation or tuning is critical for the algorithmic success.

## 4 Experimental Analysis

Our aim is to provide a *proof of concept* for the proposed aggregation approach: Concerning the achieved decision space diversity of the generated results, an originally single-objective method enhanced by the aggregation scheme shall generally be competitive to any multi-objective algorithm designed for that purpose, and superior to any standard multi-objective algorithm. We therefore focus our experimental procedure on landscapes with interesting decision space characteristics, that is, functions with several pre-images for certain points in the efficient set (non-injective functions).

### 4.1 Test Functions: Non-injective Artificial Landscapes

The following set of bi-objective functions is considered in order to test the algorithmic performance. Not many more test problems with these characteristics are known to us, however the chosen four still have very different properties.

1. **Omni-Test by Deb.** Deb et al. constructed a bi-criteria multi-global landscape for testing their Omni-Optimizer [2]. Explicitly, it reads:

$$f_1(\boldsymbol{x}) = \sum_{i=1}^{n} \sin(\pi x_i) \longrightarrow \min, \qquad f_2(\boldsymbol{x}) = \sum_{i=1}^{n} \cos(\pi x_i) \longrightarrow \min \quad (9)$$

   where $\forall i\ x_i \in [0, 6]$. We consider $n = 5$.
2. **EBN.** The EBN family of functions [21] introduced a very basic set of test-problems for multi-objective algorithms. Explicitly, it reads:

$$f_1^{(\gamma)}(\boldsymbol{x}) = \left(\sum_{i=1}^{n} |x_i|\right)^{\gamma} \cdot n^{-\gamma} \rightarrow \min, \quad f_2^{(\gamma)}(\boldsymbol{x}) = \left(\sum_{i=1}^{n} |x_i - 1|\right)^{\gamma} \cdot n^{-\gamma} \rightarrow \min \quad (10)$$

   The EBN problems are attractive in the context of efficient set approximation, as the pre-images of points in the objective space are not single points, but rather line segments on the diagonals of $[0, 1]^n$, excepting the extremal points $(0, 1)^T$ and $(1, 0)^T$ [22]. Each point in $[0, 1]^n$ is efficient. In our study we consider the case of a linear Pareto front, $\gamma = 1$, with $n = 10$.
3. **"Two-on-One".** This test-case was originally introduced in an interesting study of the Pareto-optimal set [18], large parts of which have two pre-images. It is a two-dimensional function, with a $4^{th}$-degree polynomial with two minima as $f_1$ versus the sphere function as $f_2$:

$$\begin{aligned} f_1(x_1, x_2) &= x_1^4 + x_2^4 - x_1^2 + x_2^2 - g x_1 x_2 + h x_1 + 20 \longrightarrow \min \\ f_2(x_1, x_2) &= (x_1 - k)^2 + (x_2 - l)^2 \longrightarrow \min \end{aligned} \quad (11)$$

   We consider the asymmetric case, with $g = 10$, $h = 0.25$, $k = 0$, and $l = 0$ (case number 3 as reported in [18]).
4. **Lamé Superspheres.** We consider a multi-global instantiation of a family of test problems introduced by Emmerich and Deutz [23], the Pareto fronts of which have a spherical or super-spherical geometry. In contrast to the EBN problem, the set of pre-images of a point on the Pareto front for this instance is finite, and solutions are placed on equidistant parallel line-segments, with

**Table 1.** Hypervolume of the resulting Pareto fronts of the 5 different algorithms on the 4 test-cases: average and standard-deviation over 30 runs

| Hypervolume | Niching-CMA | CMA-MO | NSGA-II | NSGA-II-Agg. | Omni-Opt. |
|---|---|---|---|---|---|
| Omni-Test | $30.27 \pm 0.05$ | $\mathbf{30.43 \pm 0.002}$ | $30.17 \pm 0.034$ | $29.81 \pm 0.2$ | $29.72 \pm 0.20$ |
| EBN | $3.295 \pm 0.038$ | $\mathbf{3.489 \pm 0.001}$ | $3.30 \pm 0.082$ | $2.848 \pm 0.173$ | $2.058 \pm 0.064$ |
| Two-on-One | $173.44 \pm 0.14$ | $\mathbf{174.52 \pm 0.005}$ | $172.59 \pm 1.53$ | $171.58 \pm 2.1$ | $168.24 \pm 7.72$ |
| Superspheres | $3.172 \pm 0.037$ | $\mathbf{3.205 \pm 0.007}$ | $3.203 \pm 0.001$ | $3.109 \pm 0.108$ | $2.481 \pm 0.375$ |

**Table 2.** Decision-space diversity, as defined in Eq. 13, of the 5 different algorithms on the 4 test-cases: average and standard-deviation over 30 runs. See also Figure 2.

| Diversity | Niching-CMA | CMA-MO | NSGA-II | NSGA-II-Agg. | Omni-Opt. |
|---|---|---|---|---|---|
| Omni-Test | $\mathbf{0.247 \pm 0.061}$ | $0.042 \pm 0.028$ | $0.191 \pm 0.085$ | $0.207 \pm 0.065$ | $0.0301 \pm 0.002$ |
| EBN | $\mathbf{0.484 \pm 0.007}$ | $0.424 \pm 0.010$ | $0.412 \pm 0.023$ | $0.357 \pm 0.027$ | $0.012 \pm 0.010$ |
| Two-on-One | $\mathbf{0.296 \pm 0.012}$ | $0.113 \pm 0.002$ | $0.183 \pm 0.102$ | $0.162 \pm 0.088$ | $0.093 \pm 0.032$ |
| Superspheres | $\mathbf{0.412 \pm 0.022}$ | $0.115 \pm 0.019$ | $0.224 \pm 0.046$ | $0.307 \pm 0.049$ | $0.0729 \pm 0.060$ |

integer distances to each other, each of them being a pre-image of a local Pareto front. Let $\xi = \frac{1}{n-1}\sum_{i=2}^{n} x_i$, and $r = \sin^2(\pi \cdot \xi)$,

$$f_1 = (1+r) \cdot \cos(x_1) \longrightarrow \min \qquad f_2 = (1+r) \cdot \sin(x_1) \longrightarrow \min \tag{12}$$

with $x_1 \in \left[0, \frac{\pi}{2}\right]$, and $x_i \in [1,5]$ for $i = 2 \ldots n$. We consider here $n = 4$.

### 4.2 Experiment

For presentation of the experimental results, we adhere to the structured reporting scheme suggested in [24], starting with the scientific question to answer.

**Research Question.** Does aggregation-niching boost decision space diversity?

**Pre-Experimental Planning.** Within first test runs, we found that a Pareto front of size 50 provides a meaningful compromise between speed and solution quality, especially for the purpose of visually examining the resulting solution sets. Most of the considered algorithms ran into stagnation after less than 50.000 evaluations, so that we chose this limit for the following experiment.

In order to assess the diversity in decision space, we set up and tested a corresponding quantifier. Given a population of size $\mu_N$, we define the population diversity of the Pareto set as the mean value of the $\frac{\mu_N(\mu_N-1)}{2}$ Euclidean distances between all individuals, normalized by the diameter $R$ of the decision space:

$$D = \frac{2}{R \cdot \mu_N(\mu_N - 1)} \cdot \sum_{A \neq B} \|\boldsymbol{x}_A - \boldsymbol{x}_B\| \tag{13}$$

**Table 3.** Calculation of the U-Test for the 4 landscapes for the 5 different algorithms. The tables contain calculations for both performance criteria: p-values for the diversity measure are presented in the upper-right part of the table; p-values for the hypervolume measure are presented in the lower-left part. Highlighted values indicate where the null hypothesis cannot be rejected at the 5% significance level (no difference).

**Omni-Test**

| p-values | Niching-CMA | CMA-MO | NSGA-II | NSGA-II-Agg. | Omni-Opt. |
|---|---|---|---|---|---|
| Niching-CMA | | 5.49e-11 | 0.0117 | 0.0199 | 3e-11 |
| CMA-MO | 3.02e-11 | | 5.19e-07 | 5.07e-10 | 0.0138 |
| NSGA-II | 6.01e-08 | 3.02e-11 | | **0.684** | 7.66e-08 |
| NSGA-II-Agg. | 3.02e-11 | 3.02e-11 | 3.02e-11 | | 1.94e-10 |
| Omni-Opt. | 3e-11 | 3e-11 | 3e-11 | 3e-11 | |

**EBN**

| p-values | Niching-CMA | CMA-MO | NSGA-II | NSGA-II-Agg. | Omni-Opt. |
|---|---|---|---|---|---|
| Niching-CMA | | 3.02e-11 | 3.02e-11 | 3.02e-11 | 3e-11 |
| CMA-MO | 3.02e-11 | | 0.017 | 3.69e-11 | 3e-11 |
| NSGA-II | **0.971** | 3.02e-11 | | 2.23e-09 | 3e-11 |
| NSGA-II-Agg. | 3.02e-11 | 3.02e-11 | 3.02e-11 | | 3e-11 |
| Omni-Opt. | 3e-11 | 3e-11 | 3e-11 | 3e-11 | |

**Two-on-One**

| p-values | Niching-CMA | CMA-MO | NSGA-II | NSGA-II-Agg. | Omni-Opt. |
|---|---|---|---|---|---|
| Niching-CMA | | 3.02e-11 | 6.36e-05 | 5.46e-09 | 3.02e-11 |
| CMA-MO | 3.02e-11 | | 0.00868 | **0.865** | **0.0701** |
| NSGA-II | 0.000377 | 9.51e-06 | | **0.122** | 4.94e-05 |
| NSGA-II-Agg. | 0.000141 | 8.48e-09 | 0.0451 | | 0.00907 |
| Omni-Opt. | 3.02e-11 | 3.02e-11 | 3.02e-11 | 3.02e-11 | |

**Super-Spheres**

| p-values | Niching-CMA | CMA-MO | NSGA-II | NSGA-II-Agg. | Omni-Opt. |
|---|---|---|---|---|---|
| Niching-CMA | | 3.02e-11 | 3.02e-11 | 2.37e-10 | 3.02e-11 |
| CMA-MO | 6.72e-10 | | 9.91e-11 | 3.02e-11 | 5.86e-06 |
| NSGA-II | 1.61e-06 | 8.48e-09 | | 2.19e-07 | 2.22e-09 |
| NSGA-II-Agg. | 0.00152 | 8.99e-11 | 3.02e-11 | | 4.97e-11 |
| Omni-Opt. | 3.02e-11 | 3.02e-11 | 3.02e-11 | 3.02e-11 | |

**Task.** We demand that the aggregation enhanced algorithms perform better than their non-aggregating counterparts in terms of diversity. Statistically, they should be better in at least 3 of 4 cases (U-test 5% level). Furthermore, they should perform as well as multi-objective algorithms specifically designed for keeping decision space diversity high (not worse at 5%) while keeping the hypervolume metric performance at a competitive level (this task is secondary and therefore not specified in detail).

**Setup.** We ran the proposed aggregation-enhanced niching method (Niching-CMA) against four reference methods: The CMA-MO [15], the NSGA-II [16], the Omni-Optimizer [2], and a variant of the NSGA-II which considers an aggregated space in the crowding calculations (referred to as *NSGA-II-Agg*). The latter routine is created from the standard NSGA-II in order to assess the importance

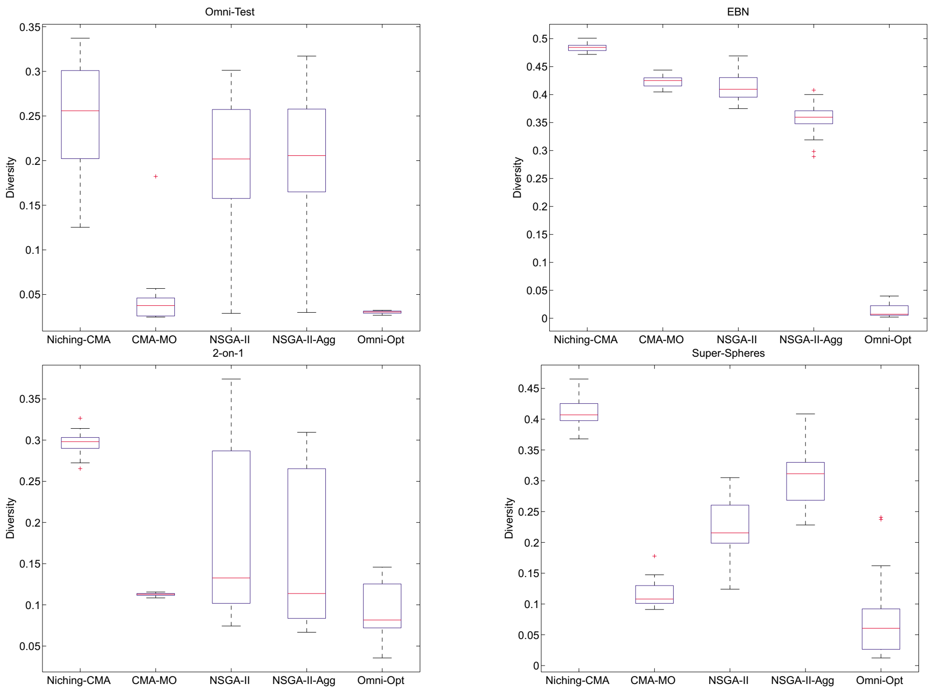


**Fig. 2.** Measured diversities in 30 runs of each of the 5 algorithms on Omni-test (upper left), EBN (upper right), the Two-on-one (lower left), and Super-Spheres (lower right) test problems

of the aggregation concept for attaining decision space diversity. All 5 methods are run on all 4 test problems of section 4.1 with 30 repeats each. We are aware that the enforced small populations may not be optimal for all algorithms; the Omni-Optimizer, for instance, was reported in [2] to employ a population of $1,000$ individuals. However, apart from these settings, we rely on default values.

**Experimentation/Visualization.** Figures 3 and 4 show typical outcomes of the resulting approximated Pareto-sets and Pareto-fronts. Note that the decision space is represented by plotting $x_1$ versus $x_2$, except for the Superspheres test-case where $x_1$ is plotted versus $\frac{1}{(n-1)} \cdot \sum_{i=2}^{n} x_i$.

Table 1 provides the S-metric results, following $2D$ hypervolume calculations for test-cases 1-4 with reference points $\{(1,1),(2,2),(35,7),(2,2)\}$, respectively; Table 2 presents the calculations of the decision space diversity as defined in Eq. 13. Figure 2 contains the box-plots for the latter table. Furthermore, Table 3 presents the p-values for Mann-Whitney U-Tests for both the hypervolume as well as the diversity criterion, between all 5 algorithms on all 4 test problems.

**Observations.** In the Omni-Test landscape, Niching-CMA performed very well, while typically obtaining 4 Pareto subsets, in comparison to one or two subsets for each of the other routines. In the EBN landscape, Niching-CMA attained

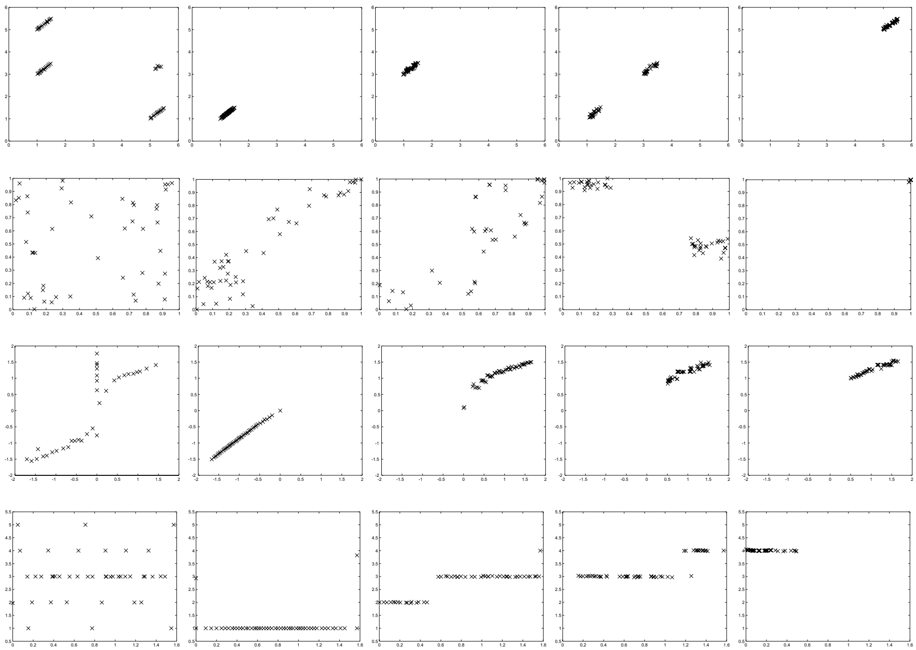

**Fig. 3.** Final populations of the 5 algorithms in the decision spaces of the 4 different landscapes. Note that the decision space is represented by plotting $x_1$ versus $x_2$, except for the Superspheres test-case where $x_1$ is plotted versus $\frac{1}{(n-1)} \cdot \sum_{i=2}^{n} x_i$. Columns, from left to right, present the algorithms in the following order: Niching-CMA, CMA-MO, NSGA-II, NSGA-II-Agg, Omni-Opt. First row presents the Omni-Test problem, followed by the EBN, 2-on-1, and Superspheres.

a quasi-uniform distribution in the decision space. In the "Two-on-One" landscape, the proposed algorithm managed to explore both branches of the so-called *propeller-shaped Pareto-set* (for more details see [18]), while the other algorithms typically explored either one of the two branches. In the Super-Spheres landscape, Niching-CMA performed extremely well, while obtaining a good distribution of typically 3 Pareto subsets. The other methods, nevertheless, usually obtained a single Pareto subset. This is clearly observed in the fourth row of Figure 3, where the final population of the these algorithms is mostly concentrated along a single line, corresponding to a single Pareto subset.

**Discussion.** Generally speaking, the proposed algorithm performs in a satisfying manner, obtaining good Pareto-sets with high diversity in the decision space, which are mapped onto well-approximated Pareto-fronts. In terms of the performance criterion in the objective space, the S-metric (hypervolume), CMA-MO did best on all test problems, whereas Niching-CMA and NSGA-II performed slightly worst and equally well, and NSGA-II with aggregation and Omni-Optimizer showed slightly worse performance. Regarding the diversity in the decision space, the proposed algorithm accomplished its goal: It attained

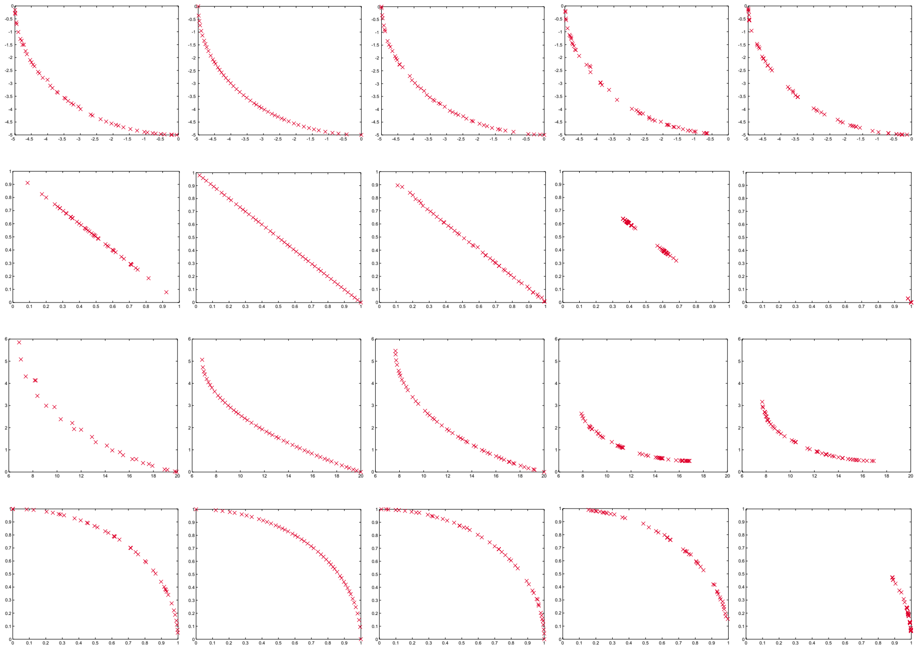

**Fig. 4.** Final populations of the 5 algorithms in the objective spaces of the 4 different landscapes, $f_1$ plotted versus $f_2$. Columns, from left to right, present the algorithms in the following order: Niching-CMA, CMA-MO, NSGA-II, NSGA-II-Agg, Omni-Opt. Rows from top to bottom: Omni-Test problem, EBN, 2-on-1, and Superspheres.

higher decision space diversity in comparison to the other methods on all landscapes. The CMA-MO, the S-metric winner, did not attain high decision space diversity; This is not a surprising result, as it is not meant to target this goal.

It should be noted that introducing the aggregation component into the NSGA-II did improve the attained decision space diversity to some extent on two landscapes, but did not have a considerable contribution. We may conclude that considering the aggregated space by itself does not seem to be sufficient for attaining high diversity in the decision space. We rather consider it as a *bridge* for niching to multi-objective domains. The Omni-Optimizer performed comparably poor in terms of the attained decision space diversity, and it is likely due to being hampered by the small population size.

## 5 Summary and Outlook

This paper addressed the topic of decision space diversity in the framework of Evolutionary Multi-Objective Algorithms. After providing the reader with the motivation for this study, and reviewing the existing work done on this topic, we outlined a new approach which aims at tackling multi-criterion problems while boosting diversity in the efficient set. The proposed algorithm relied on an existing CMA-based niching technique, which required adjustments in the selection scheme and the diversity measure. Due to the fact that it is a niche-radius based method, we proposed a way to choose a default value for this parameter. The algorithm was applied to a test-bed of non-injective artificial bi-criteria landscapes of various dimensions, and compared to the multi-objective CMA as well as to the classical GA-based EMOA: NSGA-II and its variants. The observed numerical results were satisfying, and provided us with the desired proof of concept for the proposed method. Furthermore, we concluded that employing space aggregation solely does not seem to be sufficient for attaining decision space diversity, and that niching could be the required bridging mechanism for multi-objective optimization. It should be noted that the GA-based methods performed poorly, likely due to the small population sizes that are typically employed by ES-based algorithmic kernels. Future research will be needed to test the approach on higher dimensional objective spaces and to explore various possibilities for parametrization and instantiation of the proposed approach.

## Acknowledgments

The authors would like to thank Christian Igel for his dedicated support in using the Shark Library for the CMA-MO calculations. Mike Preuss gratefully acknowledges support from the Deutsche Forschungsgemeinschaft (DFG), grant no. RU 1395/3-2 "Ein Verfahren zur Optimierung von aus Mehrkomponenten bestehenden Schiffsantrieben".

## References

1. Coello, C.A.C., Lamont, G.B., Van Veldhuizen, D.A.: Evolutionary Algorithms for Solving Multiobjective Problems. Springer, Berlin (2007)
2. Deb, K., Tiwari, S.: Omni-optimizer: A Procedure for Single and Multi-objective Optimization. In: Coello Coello, C.A., Hernández Aguirre, A., Zitzler, E. (eds.) EMO 2005. LNCS, vol. 3410, pp. 47–61. Springer, Heidelberg (2005)
3. Rudolph, G., Naujoks, B., Preuss, M.: Capabilities of EMOA to Detect and Preserve Equivalent Pareto Subsets. In: Obayashi, S., Deb, K., Poloni, C., Hiroyasu, T., Murata, T. (eds.) EMO 2007. LNCS, vol. 4403, pp. 36–50. Springer, Heidelberg (2007)
4. Nicolaou, C., Brown, N., Pattichis, C.: Molecular optimization using computational multi-objective methods. Current Opinion in Drug Discovery and Development 10, 316–324 (2007)
5. Kruisselbrink, J.W., Bäck, T., IJzerman, A.P., van der Horst, E.: Evolutionary algorithms for automated drug design towards target molecule properties. In: GECCO 2008: Proceedings of the 10th annual conference on Genetic and evolutionary computation, pp. 1555–1562. ACM, New York (2008)
6. Parmee, I.C., Cvetković, D., Watson, A.H., Bonham, C.R.: Multiobjective satisfaction within an interactive evolutionary design environment. ECJ 8(2), 197–222 (2000)

7. Schütze, O., Vasile, M., Coello, C.C.: Approximate solutions in space mission design. In: Rudolph, G., Jansen, T., Lucas, S., Poloni, C., Beume, N. (eds.) PPSN 2008. LNCS, vol. 5199, pp. 805–814. Springer, Heidelberg (2008)
8. Chan, K.P., Ray, T.: An Evolutionary Algorithm to Maintain Diversity in the Parametric and the Objective Space. In: International Conference on Computational Robotics and Autonomous Systems (CIRAS), Centre for Intelligent Control, National University of Singapore (2005) ISSN: 0219-6131
9. Toffolo, A., Benini, E.: Genetic Diversity as an Objective in Multi-Objective Evolutionary Algorithms. Evolutionary Computation 11(2), 151–167 (2003)
10. Preuss, M., Schönemann, L., Emmerich, M.: Counteracting genetic drift and disruptive recombination in $(\mu + /, \lambda)$-EA on multimodal fitness landscapes. In: Genetic and Evolutionary Computation Conference (GECCO), pp. 865–872. ACM Press, New York (2005)
11. Mahfoud, S.W.: Niching Methods for Genetic Algorithms. PhD thesis, University of Illinois at Urbana Champaign (1995)
12. Shir, O.M.: Niching in Derandomized Evolution Strategies and its Applications in Quantum Control. PhD thesis, Leiden University, The Netherlands (2008)
13. Hansen, N., Ostermeier, A.: Completely Derandomized Self-Adaptation in Evolution Strategies. Evolutionary Computation 9(2), 159–195 (2001)
14. Shir, O.M., Bäck, T.: Niching with Derandomized Evolution Strategies in Artificial and Real-World Landscapes. Natural Computing: An International Journal (2008)
15. Igel, C., Hansen, N., Roth, S.: Covariance Matrix Adaptation for Multi-objective Optimization. Evolutionary Computation 15(1), 1–28 (2007)
16. Deb, K.: Multi-Objective Optimization Using Evolutionary Algorithms. Wiley, New York (2001)
17. Horn, J., Nafpliotis, N., Goldberg, D.E.: A Niched Pareto Genetic Algorithm for Multiobjective Optimization. In: Conference on Evolutionary Computation (CEC), pp. 82–87. IEEE Service Center, Piscataway (1994)
18. Preuss, M., Naujoks, B., Rudolph, G.: Pareto Set and EMOA Behavior for Simple Multimodal Multiobjective Functions. In: Runarsson, T.P., Beyer, H.-G., Burke, E.K., Merelo-Guervós, J.J., Whitley, L.D., Yao, X. (eds.) PPSN 2006. LNCS, vol. 4193, pp. 513–522. Springer, Heidelberg (2006)
19. Hansen, N., Kern, S.: Evaluating the CMA Evolution Strategy on Multimodal Test Functions. In: Yao, X., Burke, E.K., Lozano, J.A., Smith, J., Merelo-Guervós, J.J., Bullinaria, J.A., Rowe, J.E., Tiňo, P., Kabán, A., Schwefel, H.-P. (eds.) PPSN 2004. LNCS, vol. 3242, pp. 282–291. Springer, Heidelberg (2004)
20. Ehrgott, M.: Multicriteria Optimization, 2nd edn. Springer, Berlin (2005)
21. Emmerich, M., Beume, N., Naujoks, B.: An EMO algorithm using the hypervolume measure as selection criterion. In: Coello Coello, C.A., Hernández Aguirre, A., Zitzler, E. (eds.) EMO 2005. LNCS, vol. 3410, pp. 62–76. Springer, Heidelberg (2005)
22. Emmerich, M.: Single- and Multi-objective Evolutionary Design Optimization Assisted by Gaussian Random Field Metamodels. PhD thesis, University of Dortmund, Germany (2005)
23. Emmerich, M., Deutz, A.: Test problems based on lamé superspheres. In: Obayashi, S., Deb, K., Poloni, C., Hiroyasu, T., Murata, T. (eds.) EMO 2007. LNCS, vol. 4403, pp. 922–936. Springer, Heidelberg (2007)
24. Preuss, M.: Reporting on Experiments in Evolutionary Computation. Technical Report CI-221/07, University of Dortmund, SFB 531 (2007)