technische universität
dortmund

Boosting Decision-Space Diversity in Multi-Objective Optimization unsing Nichcing-CMA and Aggregation

Ofer M. Shir, Mike Preuss,
Boris Naujoks, Michael Emmerich

Algorithm Engineering Report
TR08-2-004
Sept 2008
ISSN 1864-4503

Faculty of Computer Science
Algorithm Engineering (Ls11)
44221 Dortmund / Germany
http://ls11-www.cs.uni-dortmund.de/

Boosting Decision-Space Diversity in Multi-Objective Optimization using Niching-CMA and Aggregation

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

[1] Natural Computing Group, Leiden University
Niels Bohrweg 1, 2333 CA Leiden, The Netherlands
{oshir,emmerich}@liacs.nl,
http://natcomp.liacs.nl

[2] Universität Dortmund, Lehrstuhl für Algorithm Engineering
44221 Dortmund, Germany
{mike.preuss,boris.naujoks}@uni-dortmund.de,
http://ls11-www.cs.uni-dortmund.de

Abstract. Two solutions that occupy (almost) the same space in the objective space may have pre-images in the decision space that are essentially different. For a decision maker it can be interesting to know both pre-images in such cases. However, most Pareto optimization algorithms focus on diversity in the objective space only and thus will likely obtain only one solution. In this paper we propose a method aiming for approximation sets that possess a high diversity in objective space *as well as* decision space. The method integrates aggregation of the two spaces into an existing CMA-niching framework to yield a multi-objective algorithm. Based on a study on synthetic multimodal problems we discuss the aggregation and niching concept and assess the performance of the new approach. We conclude that considering the aggregated space by itself is not sufficient for attaining high diversity in the decision space, but it is rather a *bridge* for niching to multi-objective optimization.

1 Introduction

In multi-objective optimization we are interested in solving problems with many, possibly conflicting objectives. A common approach to solve these problems is to generate a diverse set of non-dominated solutions in the objective space and let the decision maker choose one of the solutions from it. Various algorithms have been proposed for this task, including many evolutionary algorithms [1].

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 a decision maker [2, 3]. For instance, if decision makers select a favorite point on the Pareto front, it might be interesting for them to find different possibilities to realize this solution. Hence, if there are two different pre-images of the selected point on the Pareto front in the efficient set (cf. Fig.

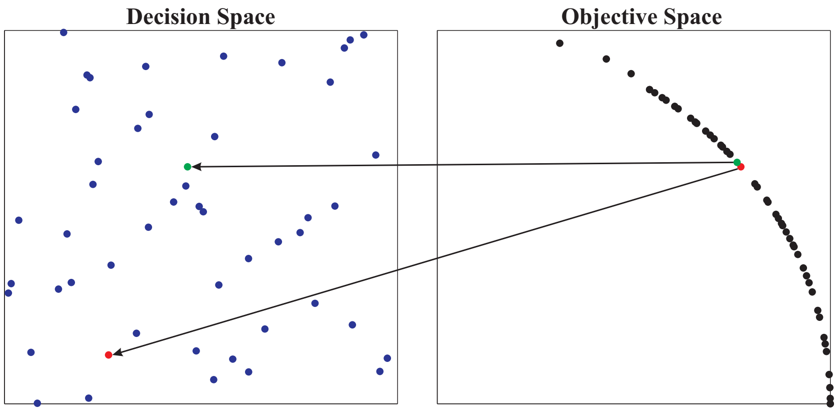

Fig. 1: *Diversity for* decision making*: Illustrative example for a scenario where two adjacent points on the Pareto front are mapped onto two points in two completely different regions in the decision space. Units and scalability are arbitrary.*

1), both of them are of potential interest for the decision maker.
The proposed approach can be modeled as a ranking criterion as follows: Let $\mathcal{A}$ denote an approximation set on which we would like to establish a ranking, $\mathbf{x_A}$ and $\mathbf{x_B}$ two solutions in $\mathcal{A}$. Normally, a solution $\mathbf{x_A}$ is preferred to a solution $\mathbf{x_B}$ if $\mathbf{x_A}$ has a better dominance rank than $\mathbf{x_B}$ in $\mathcal{A}$, e.g. with respect to non-dominated sorting. Given that $\mathbf{x_A}$ and $\mathbf{x_B}$ share the same dominance rank in $\mathcal{A}$, $\mathbf{x_A}$ is preferred to $\mathbf{x_B}$, if and only if (iff) it contributes more to the diversity of the approximation set in the objective space. In the proposed selection principle, $\mathbf{x_A}$ remains preferable to $\mathbf{x_B}$, if $\mathbf{x_A}$ has a better dominance rank than $\mathbf{x_B}$ in $\mathcal{A}$. However, given that $\mathbf{x_A}$ and $\mathbf{x_B}$ share the same dominance rank in $\mathcal{A}$, then $\mathbf{x_A}$ is preferred to $\mathbf{x_B}$, iff it contributes more to the diversity in the **aggregated space**, i.e. the combined objective and decision space. This principle can be instantiated in different ways, depending on the diversity measure defined on the aggregated space.

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 [4]. This problem is addressed by *Niching methods*, an extension of EAs to multimodal optimization [5, 6] These methods allow for parallel convergence into multiple good solutions. Niching has been traditionally investigated within Genetic Algorithms (GA) [5]. Recently, it became popular in Evolution Strategies (ES), especially as combined with the Covariance Matrix Adaptation (CMA) ES as the state-of-the-art ES [7, 6]. Moreover, Igel et al. proposed a *multi-criterion* (MC) version of the CMA ES [8]. In this work, we suggest to employ the aggregation of spaces *and* niching within this algorithm.

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 approach 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)* [9]. 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 [10]. 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 [11] 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 [12]. 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, 13], while focusing on defining different test functions and analyzing the algorithmic behavior on them.

3 The Algorithmic Approach

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

The CMA-ES [14], is a derandomized ES variant that has been successful in treating correlations among object variables by efficiently learning matching

Algorithm 1 (μ_w, λ)-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 $\mu = \lfloor \frac{\lambda}{2} \rfloor$ fittest individuals of niche j with $Parent(peak(j))$
  Apply weighted recombination on $\boldsymbol{x}_w$ and $\boldsymbol{z}_w$ w.r.t. those individuals
  Inherit the CMA-Set of $peak(j)$ and update it w.r.t. $\boldsymbol{x}_w$ and $\boldsymbol{z}_w$
end for
if N_{DPS}=size of $DPS < q$ **then**
  Generate $q - N_{dps}$ new search points, reset CMA-Sets
end if

mutation distributions. Explicitly, given an initial search point $\boldsymbol{x}^{(0)}$, λ offspring are generated by means of Gaussian variations:

$$\boldsymbol{x}^{(g+1)} \sim \mathcal{N}\left(\langle \boldsymbol{x} \rangle_W^{(g)}, {\sigma^{(g)}}^2 \mathbf{C}^{(g)}\right) \tag{1}$$

Here, $\mathcal{N}(\boldsymbol{m}, \mathbf{C})$ denotes a normally distributed random vector with mean $\boldsymbol{m}$ and a covariance matrix $\mathbf{C}$. The best μ search points out of these λ offspring undergo weighted recombination and become the parent of the next generation, denoted by $\langle \boldsymbol{x} \rangle_W$. The covariance matrix $\mathbf{C}$ 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 [14].

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., [7]). 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 reformed in each generation using the dynamic peak identification (DPI) routine [7]. 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 needs to be defined a-priori [7]. In contrast to previous CMA-Niching ES, this study will introduce multiple parents in each niche. 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 μ may vary over time, other auxiliary coefficients must be updated accordingly, such as the recombination weights. As for the value of λ, we propose to set it to its recommended default value: $\lambda = 4 + \lfloor 3 \cdot \ln(n) \rfloor$, with n as the search space dimensionality. Algorithm 1 summarizes the niching with CMA-ES routine.

The proposed routine uses the CMA-niching 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:

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

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.* 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.

3.1 Estimation of 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 ρ for specific cases. The following derivations are strictly limited to $2D$ decision or objective spaces, but we believe that they could be generalized to n-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 considering the aggregated space, and 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 ρ, 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):

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

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 [15].

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 a lower bound of

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

and an upper bound of

$$l_{FRONT,max} = |f_{1,\mathcal{N}} - f_{1,min}| + |f_{2,\mathcal{N}} - f_{2,min}| \,. \qquad (6)$$

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

$$\frac{\sqrt{\left((f_{1,\mathcal{N}} - f_{1,min})^2 + (f_{2,\mathcal{N}} - f_{2,min})^2\right)}}{2 \cdot q} \leq \rho \leq \frac{|f_{1,\mathcal{N}} - f_{1,min}| + |f_{2,\mathcal{N}} - f_{2,min}|}{2 \cdot q} \qquad (7)$$

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:

$$r_{SET} = \sqrt{\sum_{i=1}^{n} (x_{i,max} - x_{i,min})^2} \qquad (8)$$

$$r_{FRONT} = \sqrt{\sum_{j=1}^{m} (f_{j,max} - f_{j,min})^2} \qquad (9)$$

And thus

$$\rho = \frac{\sqrt{\sum_{i=1}^{n} (x_{i,max} - x_{i,min})^2 + \sum_{j=1}^{m} (f_{j,max} - f_{j,min})^2}}{2 \cdot q} \qquad (10)$$

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

4 Numerical Simulations

Our aim is to provide a *proof of concept* for the proposed approach. We therefore focus our experimental procedure on landscapes with interesting decision space characteristics.

4.1 Test Functions: Artificial Landscapes

The following set of bi-objective functions is considered in order to test the algorithmic performance:

1. **Omni-Test by Deb et al.** As mentioned earlier, Deb constructed a bi-criteria multi-global landscape for testing his Omni-Optimizer [2]. Explicitly, it reads:
$$f_1(\boldsymbol{x}) = \sum_{i=1}^{n} \sin(\pi x_i) \longrightarrow \min \qquad f_2(\boldsymbol{x}) = \sum_{i=1}^{n} \cos(\pi x_i) \longrightarrow \min \tag{11}$$
where $\forall i\ x_i \in [0, 6]$.
2. **EBN** The EBN family of functions [16] introduced a very basic set of test-problems for multi-objective algorithms. Explicitly, it reads:
$$\begin{aligned} f_1^{(\gamma)}(\boldsymbol{x}) &= \left(\sum_{i=1}^{n} |x_i|\right)^{\gamma} \cdot n^{-\gamma} \longrightarrow \min \\ f_2^{(\gamma)}(\boldsymbol{x}) &= \left(\sum_{i=1}^{n} |x_i - 1|\right)^{\gamma} \cdot n^{-\gamma} \longrightarrow \min \end{aligned} \tag{12}$$
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 $(1, 0)^T$ (for the proof see [17]).
In our study we shall consider the case of $\gamma = 1$.
3. **"Two-on-One"** This test-case was originally introduced in an interesting study of the Pareto-optimal set [13]. 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:
$$\begin{aligned} f_1(x_1, x_2) &= x_1^4 + x_2^4 - x_1^2 + x_2^2 - cx_1x_2 + dx_1 + 20 \longrightarrow \min \\ f_2(x_1, x_2) &= (x_1 - k)^2 + (x_2 - l)^2 \longrightarrow \min \end{aligned} \ . \tag{13}$$
We consider the asymmetric case, with $c = 10$, $d = 0.25$, $k = 0$, and $l = 0$ (case number 3 as reported in [13]).
4. **Lamé Superspheres** We consider a multi-global instantiation of a family of test problems introduced by Emmerich and Deutz [18], 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, each of them being a pre-image of a local Pareto front.
Let $d = \frac{1}{n-1} \sum_{i=2}^{n} x_i$, and $r = \sin^2(\pi \cdot d)$,
$$f_1 = (1 + r) \cdot \cos(x_1) \longrightarrow \min \qquad f_2 = (1 + r) \cdot \sin(x_1) \longrightarrow \min \tag{14}$$
with $x_1 \in \left[0, \frac{\pi}{2}\right]$, and $x_i \in [1, 5]$ for $i = 2 \ldots n$.

Table 1: *Hypervolume of the resulting Pareto fronts of the* 4 *different algorithms on the* 4 *test-cases: average and standard-deviation over* 20 *runs.*

Hypervolume	Niching-CMA	NSGA-II	NSGA-II-Agg	Omni-Opt.
Omni-Test	30.27 ± 0.05	30.17 ± 0.034	29.80 ± 0.23	29.75 ± 0.18
EBN	3.283 ± 0.042	3.289 ± 0.088	2.87 ± 0.182	2.064 ± 0.057
Two-on-One	173.4 ± 0.26	173.7 ± 1.56	172.7 ± 1.78	150.2 ± 28.6
Superspheres	3.176 ± 0.038	3.203 ± 0.001	3.117 ± 0.080	2.457 ± 0.372

4.2 Modus Operandi

We carried out numerical simulations on the bi-criteria landscapes introduced in the previous section in order to test the algorithmic performance of the proposed method. We chose to apply three additional algorithms as reference methods: the NSGA-II [11], the Omni-Optimizer [2], and a variant of the NSGA-II which considers an aggregated space in the crowding calculations (referred to in our notation as *NSGA-II-Agg*). The latter routine is meant to assess the importance of the aggregation concept for attaining decision space diversity. The idea was to approximate the Pareto front by means of $q = 50$ points, and allocate a fixed number of $NumEval_{\max} = 50,000$ function evaluations per run. We are aware that these are not the optimal settings for the reference methods; The Omni-Optimizer, for instance, was reported in [2] to employ a population of $1,000$ individuals. However, our goal here is also to exploit the advent of modern derandomized Evolution Strategies, which offer optimization with minimal settings.

In order to assess the boost of diversity in the decision space, we introduce a corresponding quantifier. Given the population size μ_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 d of the decision space:

$$D = \frac{2}{d \cdot \mu_N(\mu_N - 1)} \cdot \sum_{A \neq B} \|\boldsymbol{x}_A - \boldsymbol{x}_B\| \tag{15}$$

4.3 Numerical Observation

We present the numerical results by means of plots of typical runs of the resulting approximated Pareto-set and Pareto-front (i.e., all the non-dominated individuals of the last generation; cf. Fig. 2 to 5). The plots present the outcome of the different algorithms both in the decision and the objective spaces for each landscape. 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 presents the calculations of the S-metric, as a performance criterion in the objective space, averaged over 20 runs. Moreover, Table 2 presents the calculations of the decision space diversity, as defined in Eq. 15, averaged over 20 runs.

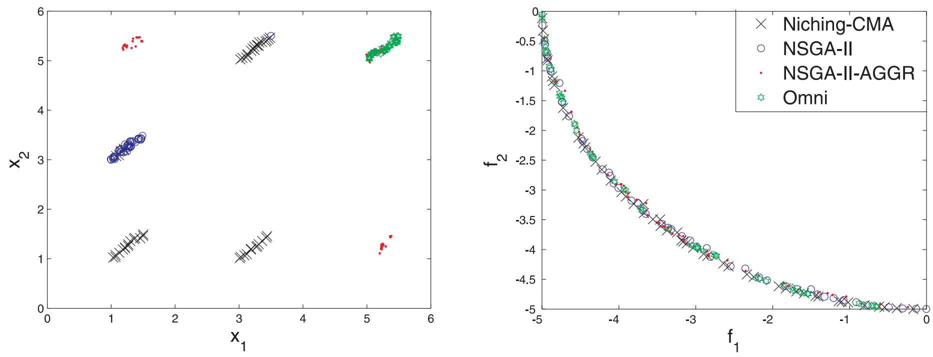

Fig. 2: *5D Omni-Test landscape (Eq. 11): Final populations of the four routines (see legend). Left: Decision space; Right: Objective space.*

Generally speaking, the proposed algorithm performs in a satisfying manner, obtaining good Pareto-sets with high diversity in the decision space, which are mapped to a well-approximated Pareto-fronts. In terms of the performance criterion in the objective space, the S-metric (hypervolume), Niching-CMA and the NSGA-II performed equally well, while the NSGA-II with aggregation and the Omni-Optimizer typically performed slightly worse. Regarding the diversity in the decision space, the proposed algorithm accomplished its goal: it attained higher decision space diversity in comparison to the other methods on all landscapes. This result can also be clearly observed in the decision space plots. 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 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* (see [13]), 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

Table 2: *Decision-space diversity, as defined in Eq. 15, of the 4 different algorithms on the 4 test-cases: average and standard-deviation over 20 runs.*

Diversity	Niching-CMA	NSGA-II	NSGA-II-Agg	Omni-Opt.
Omni-Test	0.256 ± 0.060	0.205 ± 0.079	0.222 ± 0.070	0.030 ± 0.002
EBN	0.483 ± 0.008	0.410 ± 0.023	0.356 ± 0.028	0.011 ± 0.010
Two-on-One	0.295 ± 0.01	0.136 ± 0.036	0.116 ± 0.031	0.106 ± 0.054
Superspheres	0.413 ± 0.024	0.239 ± 0.049	0.307 ± 0.046	0.062 ± 0.056

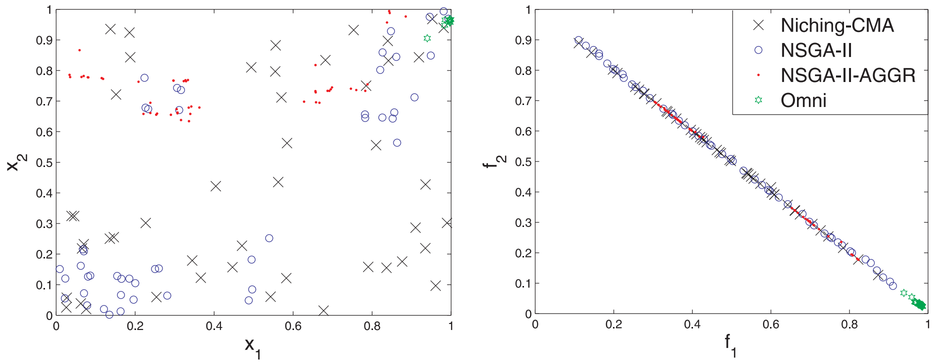

Fig. 3: *10D EBN landscape (Eq. 12): Final populations of the four routines (see legend). Left: Decision space; Right: Objective space.*

clearly observed in Figure 5, where the final population of the these algorithms is mostly concentrated along a single line, corresponding to a single Pareto subset. Hence, in multi-globality terms, Niching-CMA clearly outperformed the other methods on these landscapes.

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 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 This is likely due to being hampered by the small population size.

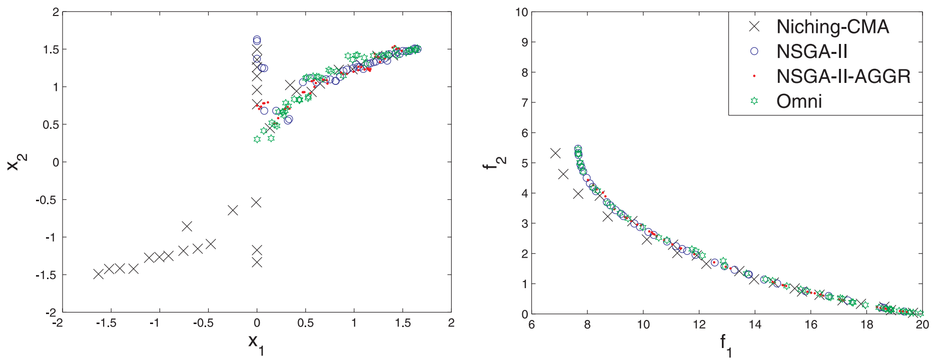

Fig. 4: *2D Two-on-One landscape (Eq. 13): Final populations of the four routines (see legend). Left: Decision space; Right: Objective space.*

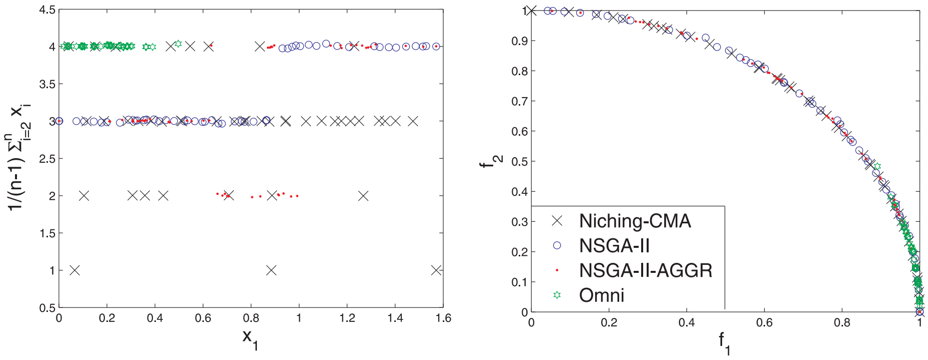

Fig. 5: *4D Super-Spheres landscape (Eq. 14): Final populations of the four routines (see legend). Left: Decision space; Right: Objective space.*

5 Summary and Outlook

The constructed algorithm required adjustments in the selection scheme and the diversity measure. Due to the fact that it is niche-radius based, we proposed a way to approximate this parameter. The algorithm was applied to a test-bed of conventional artificial bi-criteria landscapes, of various dimensions, and compared to the classical GA-based EMOAs: the NSGA-II as well as the Omni-Optimizer algorithms. The observed numerical results were satisfying, and provided us with the desired proof of concept for the proposed method. 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 parameterization and instantiation of the proposed approach.

References

1. Coello, C.A.C., Lamont, G.B., VanVeldhuizen, 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: Evolutionary Multi-Criterion Optimization (EMO), Springer, Berlin (2005) 47–61
3. Rudolph, G., Naujoks, B., Preuss, M.: Capabilities of EMOA to Detect and Preserve Equivalent Pareto Subsets. In: Evolutionary Multi-Criterion Optimization (EMO), Springer, Berlin (2007) 36–50
4. 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), ACM Press, New York (2005) 865–872

5. Mahfoud, S.W.: Niching Methods for Genetic Algorithms. PhD thesis, University of Illinois at Urbana Champaign (1995)
6. Shir, O.M.: Niching in Derandomized Evolution Strategies and its Applications in Quantum Control. PhD thesis, Leiden University, The Netherlands (2008)
7. Shir, O.M., Bäck, T.: Niching with Derandomized Evolution Strategies in Artificial and Real-World Landscapes. Natural Computing: An International Journal (2008)
8. Igel, C., Hansen, N., Roth, S.: Covariance Matrix Adaptation for Multi-objective Optimization. Evolutionary Computation **15**(1) (2007) 1–28
9. Horn, J., Nafpliotis, N., Goldberg, D.E.: A Niched Pareto Genetic Algorithm for Multiobjective Optimization. In: Conference on Evolutionary Computation (CEC), IEEE Service Center, Piscataway, NJ (1994) 82–87
10. Toffolo, A., Benini, E.: Genetic Diversity as an Objective in Multi-Objective Evolutionary Algorithms. Evolutionary Computation **11**(2) (2003) 151–167
11. Deb, K.: Multi-Objective Optimization Using Evolutionary Algorithms. Wiley, New York (2001)
12. 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.
13. Preuss, M., Naujoks, B., Rudolph, G.: Pareto Set and EMOA Behavior for Simple Multimodal Multiobjective Functions. In: Parallel Problem Solving from Nature (PPSN), Springer, Berlin (2006) 513–522
14. Hansen, N., Ostermeier, A.: Completely Derandomized Self-Adaptation in Evolution Strategies. Evolutionary Computation **9**(2) (2001) 159–195
15. Ehrgott, M.: Multicriteria Optimization. second edn. Springer, Berlin (2005)
16. Emmerich, M., Beume, N., Naujoks, B.: An EMO algorithm using the hypervolume measure as selection criterion. In: Evolutionary Multi-Criterion Optimization (EMO), Springer, Berlin (2005) 62–76
17. Emmerich, M.: Single- and Multi-objective Evolutionary Design Optimization Assisted by Gaussian Random Field Metamodels. PhD thesis, University of Dortmund, Germany (2005)
18. Emmerich, M., Deutz, A.: Test Problems Based on Lame Superspheres. In: Evolutionary Multi-Criterion Optimization (EMO), Springer, Berlin (2007) 922–936