

Species Separation by a Clustering Mean towards Multimodal Function Optimization

CATALIN STOEAN, MIKE PREUSS, AND RUXANDRA STOEAN

ABSTRACT. Present paper introduces a new evolutionary technique for multimodal real-valued optimization which uses a clustering method for separating the individuals within a population into species that are each connected to different optima from the search space. It is applied for a set of benchmark functions both for uni- and multimodal optimization and it proves to be very efficient as regards both the accuracy of the obtained results and the costs regarding the fitness evaluation calls that are spent.

2000 Mathematics Subject Classification. Primary 68T20; Secondary 91C20, 92D15.

Key words and phrases. Evolutionary Algorithms, Clustering, Function Optimization.

1. Introduction

Trying to find multiple solutions that lie within a continuous landscape is not a straightforward task for an evolutionary heuristic. A canonical evolutionary algorithm (EA) is very fit for finding the global optimum within a landscape that does not exhibit other local peaks. When there is only one optimum within the search space, the population of candidate solutions converges fast towards that peak (genetic drift). This fast convergence of the entire population of individuals turns against the algorithm when the landscape of the search space contains multiple hills, as the entire population once more tends to agglomerate towards one hill that is not necessarily the best one of them; when this happens, it is said that the search is blocked into a local optimum.

In order to tackle multimodal landscapes, the genetic drift has to be avoided or at least to be postponed as long as possible. In other words, the task of the evolutionary methods that deal with this type of problems is to maintain the diversity within the population as long as possible but, at the same time, head towards the "hottest" spots of the landscape. The main idea that lies behind most multimodal optimization techniques is to divide the population into sets of individuals that resemble one another, usually by gathering in the same subpopulation the individuals that are closer than a given radius with respect to a specific metric.

In this paper, a new method for multimodal optimization is introduced that does not use a radius for separating the subpopulations (or species), but it employs a clustering method instead. In order to verify the separation provided by the clustering technique, an additional helping mechanism is used for determining whether the various species are connected to different hills in the landscape or not.

The paper continues with a brief survey of other evolutionary methods that were tailored for problems that have multiple solutions which is then followed by the presentation of the proposed technique. The next section comprises the experimental

analysis on the application of the proposed technique for the optimization of a set of benchmark functions and the paper encloses with the concluding ideas.

2. Previous Approaches for Multimodal Optimization

The main goal of an EA that deals with a multimodal optimization problem is to avoid the premature convergence, i.e. blocking into a local optimum. This task is achieved by maintaining diversity within the population for a longer period of time and in this way a better exploration of the search space is performed.

Numerous techniques have been proposed for improving performance when dealing with multimodal problems. One of the most used evolutionary heuristics for multimodal optimization is represented by *niching* [5]. The preservation of diversity is conducted through the separation of the population according to niches; these are formed by joining individuals that lie closer than a given radius to one another. The main drawback of this method is represented by the existence of the threshold whose value very much depends on the problem to be solved. In this respect, Deb and Goldberg proposed a manner of computing the value for the radius that leads to the formation of subpopulations [2]; it is regularly embraced in most experimental designs for such EA methodologies. It uses the radius of the smallest hypersphere containing feasible space, which is given by (1).

$$r = \frac{1}{2} \sqrt{\sum_{i=1}^D (x_i^u - x_i^l)^2}. \quad (1)$$

In (1), D represents the number of dimensions of the problem at hand and x_i^u and x_i^l are the upper and the lower bounds of the i -th dimension. Knowing the number of optima N_G and being aware that each niche is enclosed by a D -dimensional hypersphere of radius r , the niche radius σ_{share} can be estimated as (2).

$$\sigma_{share} = \frac{r}{\sqrt[D]{N_G}}. \quad (2)$$

Note, however, that the employment of equation (2) for computing the value of the radius presumes that one must know in advance how many optima there exist within the fitness landscape.

Other methods like the island or diffusion models [4] that restrict the way individuals interact have been imagined for the same purpose of keeping diversity more and better within the population. Although they do not have a radius threshold, these do not focus on the particularities of the landscape, but rather on the restrictions upon individuals when recombination is applied.

The crowding technique [3] was also developed as a method of maintaining diversity: New obtained individuals replace only similar individuals in the population. However, there are some issues that have arisen around this methodology, as it suits only a limited set of multimodal problems and the value for the population size parameter very much depends on the considered problem and the number of attraction basins that exist.

The method that is currently introduced separates the population into species and it does not employ a radius parameter for that, instead, more efficiently, it connects each species to a different optimum by taking advantage of the fitness landscape and vicinity.

3. Nearest-Better Algorithm for Multimodal Optimization

Proposed Nearest Better Separation Evolutionary Algorithm (NBSEA) begins with a random population of candidate solutions in the definition interval of the function to be optimized. The recently created Nearest-Better Clustering (NBC) algorithm [7] is applied to this set of points in order to separate them into subsets of solutions within which all follow the same peak. In order to correct the separation in such a manner that there are no two or more species that target the same optimum, a mechanism that verifies whether two points are in the basin of the same peak or not is used: It checks if any best individual from a species lies within the same attraction basin as another dominating individual from a different subpopulation. If so, the two species are unified into a single one. Usual genetic operators are then applied to the entire population and then the dominating individuals from each species are reinserted into the population in case they disappeared because of the variation. In order to clarify each step, the following subsections will present every separate tool and finally the entire NBSEA is assembled in subsection 3.3.

3.1. Nearest-Better Clustering. The clustering mechanism is based on a "nearest neighbor" principle, but also uses topological information in addition to location of points. For each point, it searches for the closest neighbor that is better in terms of fitness. The basic idea behind NBC is that the best individuals in different basins of attraction are more distant from each other as compared to the mean distance of all individuals to their nearest better neighbors. Every individual connects to its nearest better neighbor, in terms of (Euclidean) distance. The longest edges – those higher than $\phi \cdot \text{mean}(\text{lengths of all edges})$ – are removed and the prototypes for each cluster are represented by those individuals that do not connect to others. Clusters are thus formed around these (dominating) individuals, which subsequently represent the species masters for NBSEA.

The NBC technique is thoroughly described in Algorithm 1. The approach has only one additional parameter to be tuned, with 2 being a good default value [7].

3.2. Detect Multimodal Mechanism. A mechanism to verify if two points lie within the same attraction basin or not has been introduced within the Multinational Algorithm in [8]. It takes a number of interior points (this being the only parameter of the mechanism) between the two points it is applied for and, if there exists any interior point with the fitness smaller than the worst of the two initial points, it is concluded that they follow different peaks, while otherwise they are said to be situated on the same hill. The method is described in detail in Algorithm 2. In a nutshell, the algorithm returns *true* if the two points follow different optima and *false* if they lie within the same attraction basins.

3.3. Nearest-Better Separation Evolutionary Algorithm. The steps that are followed by the currently proposed algorithm are presented in Algorithm 3. After the initialization of the population, the evolution progress within a repetitive process: the NBC method (described in Algorithm 1) is applied in order to separate the population into subpopulations. The selection operator is then applied to the entire population and it is afterwards followed by the variation operators, recombination and mutation. If two individuals from the same species recombine, the offspring is considered to belong to the same group. In other recombination cases and in the situation when an

Algorithm 1 The Nearest-Better Clustering Algorithm

Require: A population of individuals x .**Ensure:** Clusters of individuals: number, grouping, prototypes.

```

for  $i = 1$  to  $pop\_size$  do
  compute distances from  $x[i]$  to all individuals
end for
for  $i = 1$  to  $pop\_size$  do
  find nearest individual that is fitter than  $x[i]$ , i.e.  $x[j]$ ;
  if found then
     $edge[i] = j$ ;
  else
     $edge[i] = 0$ ;
  end if
end for
 $m = avg_{i=1}^{pop\_size}(distance(x[i], x[edge[i]]))$ ;
for  $i = 1$  to  $pop\_size$  do
  if  $edge[i] \neq 0$  and  $distance(x[i], x[edge[i]]) > \phi \cdot m$  then
     $edge[i] = 0$ ;
  end if
end for
return the prototypes –  $x[i]$ , where  $edge[i] = 0$  – and membership to clusters –  $x[i] \in cluster[j]$ , where  $edge^n[i] = prototype[j]$ ,  $i = 1, 2, \dots, pop\_size$ ,  $j = 1, 2, \dots, noOfClusters$ ;

```

Algorithm 2 The Detect Multimodal mechanism for two individuals x and y

Require: Two individuals x and y .**Ensure:** *TRUE* if x and y track different optima and *FALSE* otherwise.

```

 $i = 1$ ;
 $found = FALSE$ ;
while  $i < number\ of\ gradations$  and not  $found$  do
  for  $j = 1$  to  $number\ of\ dimensions$  do
     $interior_j = x_j + (y_j - x_j) \cdot gradation_i$ ;
  end for
  if  $f(interior) < \min(f(x), f(y))$  then
     $found = TRUE$ ;
  end if
end while
return  $found$ ;

```

individual is obtained via mutation, the offspring are considered *free*, meaning that they are not attached to any subpopulation.

As the selection and variation operators may omit and/or alter the master individuals of some species, a process of integration of the seeds that were previously detected occurs at this stage. Before inserting an individual, it is checked whether it already exists in the population, as it is introduced only in the case when there is no other copy of it. When such a seed is reinserted, it replaces the worst individual in the species of the current seed or, in case its subpopulation was eliminated, it substitutes the worst member of the population that is not a master individual.

Algorithm 3 Structure of the Nearest Better Algorithm

Require: A search/optimization problem**Ensure:** The set of seeds

```

1: begin
2: Initialize population;
3:  $t = 0$ ;
4: while stop condition is not met do
5:   Identify species seeds using Algorithm 1;
6:   if  $t > 0$  then
7:     Reconsider old seeds;
8:   end if
9:   Apply mating selection;
10:  Apply recombination;
11:  Apply mutation;
12:  Integrate the seeds into resulting population;
13:  Assimilate the free individuals to the species;
14:   $t = t + 1$ ;
15: end while
16: return the set of seeds
17: end

```

The free individuals are then integrated to the existing species; the seeds are taken ascending accordingly to the (Euclidean) distance to each free individual and it is tried, using the *detect – multimodal* mechanism, to see if they follow the same peak or not. If the affirmative case, a free individuals is assigned to the seed's species that targets the same optimum as they do. When a free individual does not follow the same peak as any of the existing seeds, it is set as a new master individual and it creates its new own species.

When the first generation passes, immediately after seeds determination, the masters from the previous generation are checked against the newly found ones, once again in ascendent order according to the distance, as the ones that are closer to each other are more likely to follow the same optimum. If there are seeds from the previous generation that follow different peaks (*detect – multimodal = true*), they are added to the current population and further set as seeds.

The algorithm may stop after a fixed number of fitness evaluations is consumed (this is the case used in the experiments of the current paper), or after a fixed number of generations, or when the peaks are found with a desired accuracy.

4. Experimental Results

Experiments are driven on a set of functions that includes one with only one global optimum in order to test if the method can still cope with such problems, three trap functions (two uni-dimensional and one bi-dimensional) and a bi-dimensional one that exhibits 10 optima disposed on an irregular landscape. The aim of the tests is to observe how the proposed approach behaves under a very low budget of fitness evaluations and how the results improve when the investment rises. The results regard the peaks accuracies, the number of found peaks and the number of basins that are detected within the given landscape.

TABLE 1. Considered Benchmark Functions

Function	Optima
$F1(x, y) = x^2 + y^2, -5.12 \leq x, y \leq 5.12$	1
$F2(x) = \begin{cases} \frac{160}{15}(15-x) & \text{for } 0 \leq x < 15 \\ \frac{200}{5}(x-15) & \text{for } 15 \leq x \leq 20 \end{cases}$	2
$F3(x) = \begin{cases} \frac{160}{10}x & \text{for } 0 \leq x < 10 \\ \frac{160}{5}(15-x) & \text{for } 10 \leq x < 15 \\ \frac{200}{5}(x-15) & \text{for } 15 \leq x \leq 20 \end{cases}$	2
$F4(x) = \begin{cases} 80(2.5-x) & \text{for } 0 \leq x < 2.5 \\ 64(x-2.5) & \text{for } 2.5 \leq x < 5 \\ 64(7.5-x) & \text{for } 5 \leq x < 7.5 \\ 28(x-7.5) & \text{for } 7.5 \leq x < 12 \\ 28(17.5-x) & \text{for } 12.5 \leq x < 17.5 \\ 32(x-17.5) & \text{for } 17.5 \leq x < 22.5 \\ 32(27.5-x) & \text{for } 22.5 \leq x < 27.5 \\ 80(x-27.5) & \text{for } 27.5 \leq x \leq 30 \end{cases}$	5
$F5(x) = \sin^6(5\pi x), 0 \leq x \leq 1$	5
$F6(x, y) = (0.3x)^3 - (y^2 - 4.5y^2)xy - 4.7\cos(3x - y^2(2+x))\sin(2.5\pi x)$ $-0.9 \leq x \leq 1.2, -1.2 \leq y \leq 1.2$	10

4.1. Test Functions. Table 1 presents the list of the functions that are considered for the experiments. Sphere function ($F1$) is taken for two variables and it has only one optimum. Two-Peak Trap Function ($F2$) has one global optimum and a local one, the latter situated just at the border of the interval; Central Two-Peak Trap Function ($F3$) also has two optima, one global and one local, but the latter is located toward the middle of the interval. Five-Uneven-Peak Trap Function ($F4$) has two global optima near the interval border and they are separated by three local ones. Deb's Function ($F5$) has five global optima to be found and, finally, Waves function ($F6$) is the most complex of the entire list as it is asymmetric and some peaks are difficult to find as they lie on the border or on flat hills. The functions are also considered for testing in papers like [6] or [8].

4.2. Task. Apply NBSEA on the functions in Table 1 and observe how it behaves for a very limited number of fitness evaluations (10^3) up to 10^4 , under 30 different parameter settings from which the most prolific is outlined.

4.3. Experimental Setup. There are three measures considered for each function in the test suite and each is computed as the average value over 30 repeated runs :

- (1) *Peak ratio* gives the ratio between the number of found peaks and the actual number of existing peaks. A peak is considered to be found if there is at least one candidate solution that gets closer than $\epsilon = 10^{-1}$ to the actual peak.
- (2) *Peak accuracy* refers to the summed difference in module between the fitness values of the desired optima and of the individuals that are closest to them.
- (3) *Found basins* refers to the number of attraction basins of the desired optima that have at least one individual inside.

For each involved parameter, 30 Latin Hypercube Samples (LHS) [1] are considered for generating equally distant values within reasonable intervals: Population size is taken between 10 and 200, mutation and recombination probabilities in $[0, 1]$, the

number of interior points between 1 and 10, while the mutation strength is dependant on the domains of definition, i.e. between $[0, 1]$ for $F5$ and $F6$ and $[0, 4]$ for the rest.

Concerning the used operators, tournament selection, intermediate recombination and mutation with normal perturbation are employed. The stop condition is considered for 10^3 fitness evaluation calls, $2 \cdot 10^3$, $5 \cdot 10^3$, $7 \cdot 10^3$ and 10^4 , in turn.

4.4. Results and Visualization. Obtained results for the first three functions in the best parameter configuration are illustrated in Figure 1. The detection of the correct number of basins is perfect for all three cases, even for the lowest number of fitness evaluations. For $F1$, the peak is detected with the desired accuracy even for the lowest budget, while for $F2$ and $F3$ they improve together with the increase in the number of fitness evaluations. Third column measures how close the candidate solutions get in accuracy to the desired peaks: When dealing with a unimodal problem ($F1$), NBSEA behaves very well as even for 10^3 evaluation calls the optimum is detected with an accuracy of $1e-04$.

Figure 2 reports the results of the NBSEA applied for the other three functions. For $F4$ and $F5$ the number of detected basins is perfect for all considered budgets. The figure from the left center demonstrates that for $F5$ all 5 peaks are detected with the desired accuracy even for 10^3 fitness evaluations. $F4$ has sharper peaks than the sinusoidal $F5$, therefore it is harder to get closer to them, as it can be seen from the top left graphic of the figure.

Waves function $F6$ is no doubt the most complex case of the considered test suite. Nevertheless, the number of found basins gets close to the actual solution even for the cheapest considered stop criterion that finds in average a bit above 9 solutions. Some optima have very narrow attractions basins, while others are large, so it would be impossible for a typical evolutionary technique for multimodal optimization that relies on fixed radius size to distinguish the subpopulation correctly, i.e. to have each one connected to a different optimum. However, as the last line of Figure 2 demonstrates, NBSEA behaves very well even for a small number of evaluation calls and without a radius threshold that is additionally hard to be tuned. As a corresponding parameter, NBSEA has the number of interior points that is a positive integer number taken between 1 and 10 in the undertaken experiments; in the best parameter configurations of $F6$, it was chosen 1 for 10^3 evaluations, 3 for 10^4 , while for the other stop criteria it was taken 2. As it can be observed, when there are more fitness evaluation calls, the number of interior points can be taken higher; the highest it is, the more accurate the separation between subpopulation is, but at the same time more fitness evaluation calls are consumed as more interior points are evaluated.

Naturally, if more than 10^4 evaluation calls are considered, the quality of results increases. However, NBSEA is meant to be an economical technique that is suitable for real-world optimization problems. In average, the runtime of NBSEA takes 0.04 seconds for 10^3 fitness evaluation calls on a computer with an Intel Core 2 Extreme, 2 x 2.80 GHz processor and a RAM memory of 4 GB.

5. Conclusions and Future Directions

A new evolutionary technique for multimodal optimization, Nearest Better Algorithm, that employs a topological clustering method for separating the population into species is herein proposed.

The NBSEA gives very promising results, especially for very economic investments regarding the number of fitness evaluation calls. Indeed, it can be observed that when

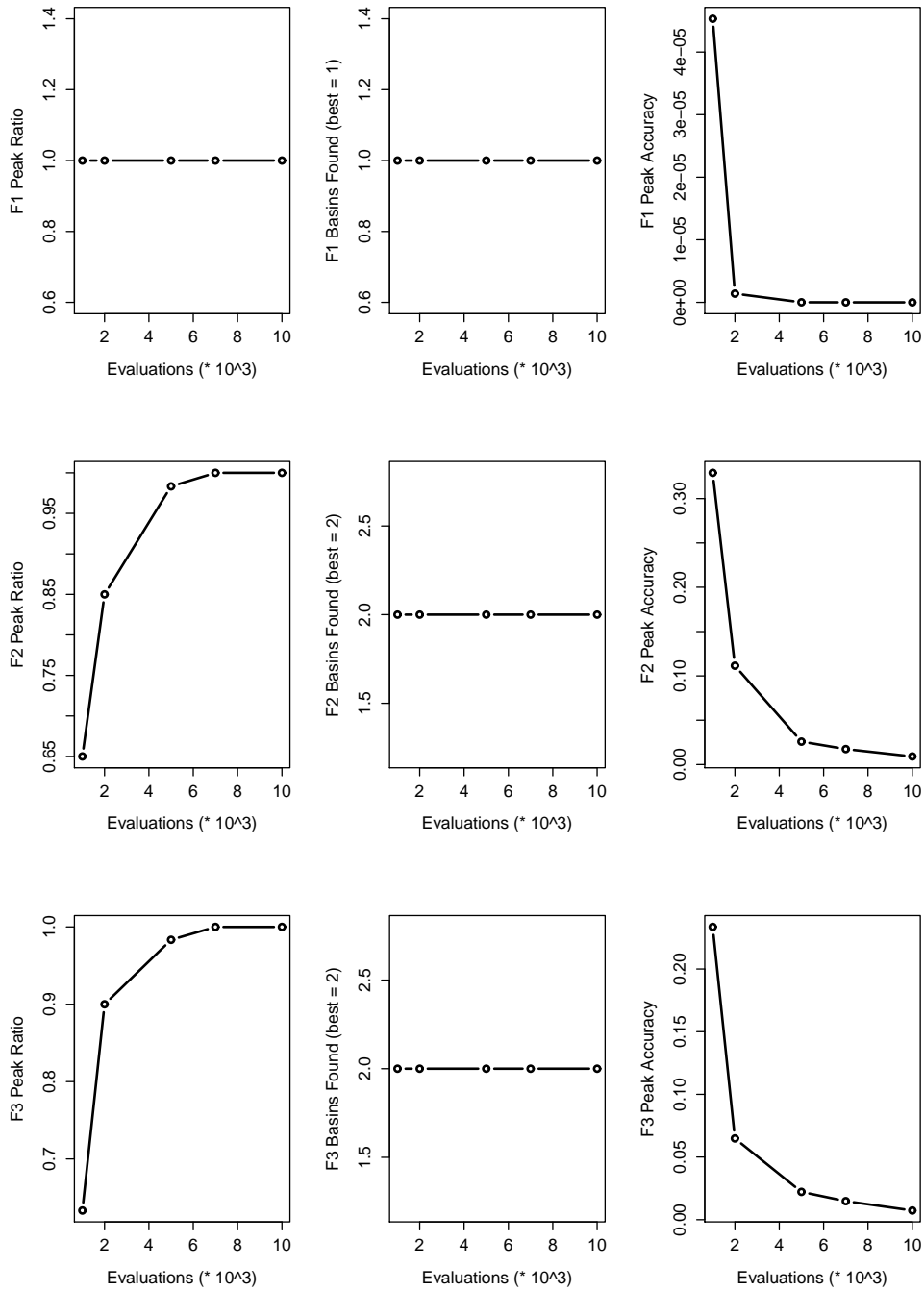


FIGURE 1. Average results over 30 repeats of the best LHS configuration, obtained by NBSEA for $F1$, $F2$ and $F3$. First column contains the peak ratio (best value is 1), the second illustrates the number of found basins (best values are 1 for $F1$ and 2 for $F2$ and $F3$), while the third reports the obtained peak accuracy (best value is 0). Each graph outlines the results obtained from 10^3 up to 10^4 fitness evaluation calls.

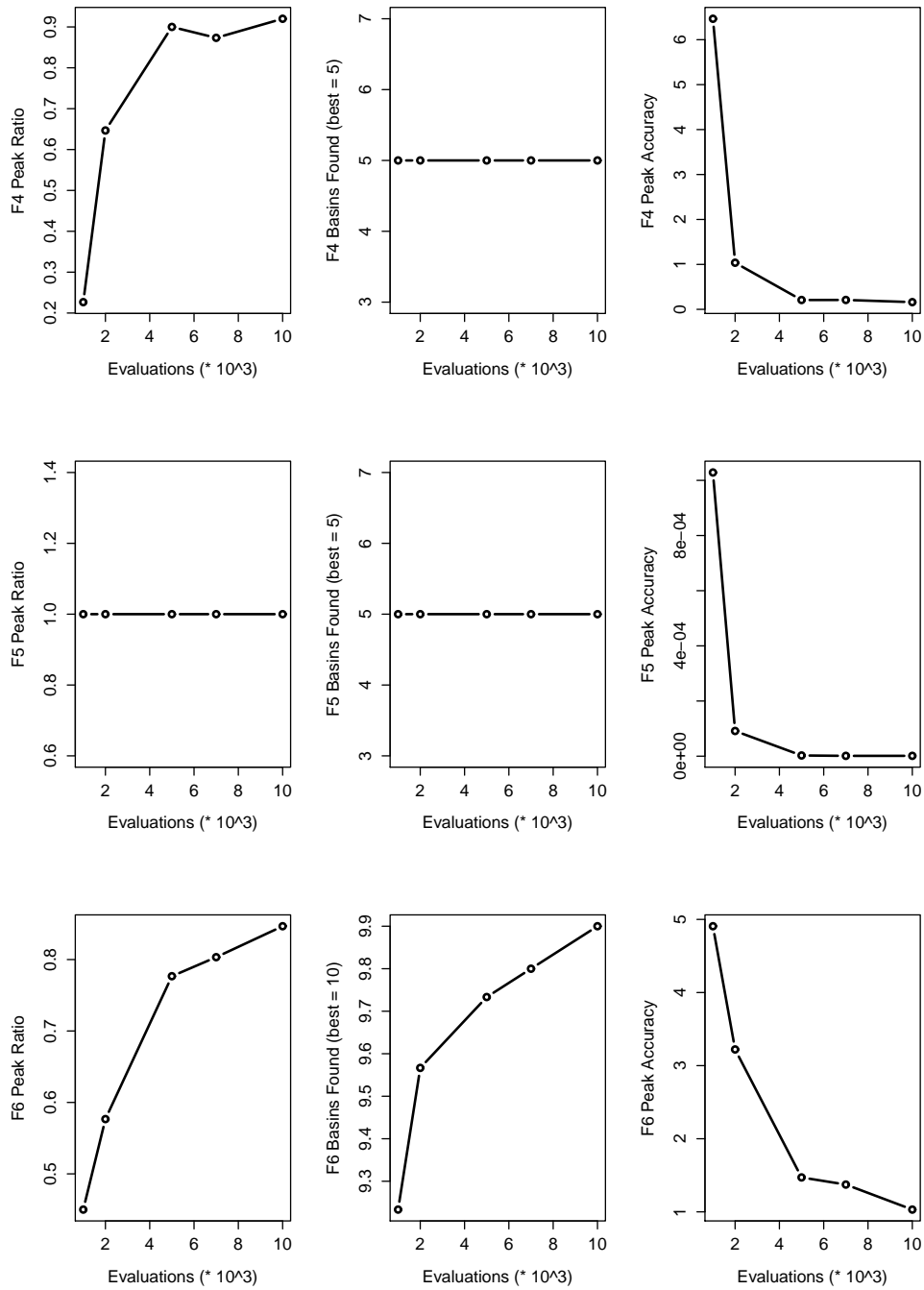


FIGURE 2. Average results over 30 repeats of the best LHS configuration, obtained by NBSEA for $F4$, $F5$ and $F6$. For the number of found basins, i.e. second column, best values are 5, 5 and 10, respectively.

more evaluations are used, the results improve accordingly, especially as regards the peak ratio and the peak accuracy. It has to be noticed that the algorithm gives very accurate results as concerns the number of attraction basins (and optima, accordingly) that a function exhibits, even at very low budgets of fitness evaluations.

It would be however interesting to further investigate next how the ϕ parameter within NBC influences the results and also extend the list of test functions by adding even more complex cases, i.e. functions that possess a large number of spiny, local optima and problems with a higher number of dimensions, and to compare the results with other state-of-the-art evolutionary techniques, all under the same constraints.

References

- [1] T. Bartz-Beielstein, *Experimental Research in Evolutionary Computation – The New Experimentalism*, ser. Natural Computing Series, Berlin, Springer, 2006
- [2] Deb, K., Goldberg, D. E., *An Investigation of Niche and Species Formation in Genetic Function Optimization*, 3rd International Conference on Genetic Algorithms, pp. 42–50, Morgan Kaufman, San Mateo, California 1989
- [3] K. A. DeJong, *An analysis of the behavior of a class of genetic adaptive systems*, Ph.D. dissertation, University of Michigan, Ann Arbor, 1975
- [4] Eiben A. E., Smith J. E., *Introduction to Evolutionary Computing*, Springer-Verlag, Berlin Heidelberg New York, 2003
- [5] Goldberg, D. E., Richardson, J., *Genetic Algorithms with Sharing for Multimodal Function Optimization*, 2nd International Conference on Genetic Algorithms, pp. 41–49, Lawrence Erlbaum Associates, Hillsdale, NJ, 1987
- [6] Li, J. P., Balazs, M. E., Parks, G. T., Clarkson, P. J., *A species conserving genetic algorithm for multimodal function optimization*, *Journal on Evolutionary Computation*, vol. 10, no. 3, pp. 207–234, 2002
- [7] Preuss, M., Schoenemann, L., Emmerich, M., *Counteracting Genetic Drift and Disruptive Recombination in (μ^+, α) -EA on Multimodal Fitness Landscapes*, 7th Annual Conference on Genetic and Evolutionary Computation (GECCO '05), pp. 865–872, ACM Press, 2005
- [8] Ursem, R., K., *Multinational Evolutionary Algorithms*, *Proceedings of the Congress of Evolutionary Computation (CEC-99)*, vol. 3, pp. 1633–1640, 1999

(Catalin Stoean) UNIVERSITY OF CRAIOVA, A. I. CUZA, NO. 13, 200585, CRAIOVA, ROMANIA
E-mail address: `catalin.stoean@inf.ucv.ro`

(Mike Preuss) TECHNISCHE UNIVERSITÄT DORTMUND, OTTO-HAHN, NO. 14, 44227, DORTMUND, GERMANY
E-mail address: `mike.preuss@tu-dortmund.de`

(Ruxandra Stoean) UNIVERSITY OF CRAIOVA, A. I. CUZA, NO. 13, 200585, CRAIOVA, ROMANIA
E-mail address: `ruxandra.stoean@inf.ucv.ro`