Maintaining Population Diversity in Evolution Strategy for Engineering Problems

Roman Dębski, Rafał Dreżewski, Marek Kisiel-Dorohinicki

Department of Computer Science
AGH University of Science and Technology, Kraków, Poland
{drezew,doroh}@agh.edu.pl

Abstract. In the paper three new mechanisms for maintaining population diversity in (μ, λ)-evolution strategies are introduced: deterministic modification of standard deviations, crowding, and elitism. The proposed mechanisms are experimentally verified with the use of optimal shape designing of rotating variable-thickness annular elastic discs problem.
Keywords: evolution strategies, maintaining population diversity, optimal shape design

1 Introduction

Many engineering tasks lead to the global optimization problems, which in most cases cannot be solved by traditional methods. In such a situation one can use (meta)heuristic computational techniques such as *evolutionary algorithms*. This term covers a wide range of search and optimization methods, based on analogies to Darwinian model of evolutionary processes. Particularly interesting from the engineering problems point of view are evolution strategies. They are most often used for solving continuous optimization problems and are distinguished by a real-valued representation, Gaussian mutation with auto-adaptation as the main variation operator, and deterministic selection scheme [2, 8].

The paper discusses the application of evolution strategies to optimal shape designing of a rotating variable-thickness annular disc—the optimal shape means here the one corresponding to the maximal elastic carrying capacity of the disc [6]. The main goal of the paper is to present some modifications of "classical" evolution strategies mainly focused on maintaining population diversity and thus protecting the searching process from getting stuck in a local extrema. Maintaining population diversity seems to be the problem of vast importance in this case.

The paper is organized as follows. Classical evolution strategies and proposed mechanisms for maintaining population diversity (i.e. deterministic modification of standard deviations, crowding and elitism) are described in section 2. Section 3 presents the optimization problem: the design of rotating variable-thickness annular elastic disc, proposed representation of the solutions, and the model used to evaluate their quality (fitness). Selected experimental results with the proposed (μ, λ)-evolution strategy with additional mechanisms conclude the work.

2 Evolution Strategies

Evolution strategies (ES) were developed by Rechenberg and Schwefel in the 1960s at the Technical University of Berlin. The first applications were aimed at hydrodynamical problems like shape optimization of a bended pipe and drag minimization of a joint plate [3]. ES is a special instance of an evolutionary algorithm characterized by real-valued vector representation, Gaussian mutation as main variation operator, self-adaptation of mutation rate, and deterministic selection mechanisms.

2.1 Classical Approach

Algorithmic framework of contemporary evolution strategies may be described with the use of following notation [7]:

- $(\mu + \lambda)$-ES generates λ offspring from μ parents and selects the μ best individuals from $\mu + \lambda$ (parents and offspring) individuals $(1 \leq \mu \leq \lambda)$,
- (μ, λ)-ES denotes an ES that each time step generates λ offspring from μ parents and selects the μ best individuals only from λ (offspring) individuals $(1 \leq \mu \leq \lambda)$.

The individuals in a population consist of the objective variables vector $\boldsymbol{x}$ and a vector of *strategy parameters* $\boldsymbol{\sigma}$, where σ_i denotes the standard deviation used when applying a zero-mean Gaussian mutation to the i-th component in parent vector. These parameters are incorporated into the representation of individual in order to obtain evolutionary *self-adaptation* of an ES [8, 1]. The mutation operator changes strategy parameters according to:

$$\sigma_i' = \sigma_i \exp(\tau_0 N(0,1) + \tau N_i(0,1)) \tag{1}$$

and the objective variables (a simplified case of uncorrelated mutations):

$$x_i' = x_i + N(0, \sigma_i') \tag{2}$$

where the constant $\tau \propto \frac{1}{\sqrt{2\sqrt{n}}}$, $\tau_0 \propto \frac{1}{\sqrt{2n}}$, $N(0,1)$ is a standard Gaussian random variable sampled once for all n dimensions and $N_i(0,1)$ is a standard Gaussian random variable sampled for each of the n dimensions.

If the number of parents $\mu > 1$, the objective variables and internal strategy parameters can be recombined with usual recombination operators, for example *intermediate recombination* [4], which acts on two parents $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ and creates an offspring $\boldsymbol{x'}$ as the weighted average:

$$x_i' = \alpha x_{1i} + (1 - \alpha) x_{2i} \tag{3}$$

where $\alpha \in [0, 1]$ and $i = 1, \ldots, n$. The same may be applied to standard deviations:

$$\sigma_i' = \alpha \sigma_{1i} + (1 - \alpha) \sigma_{2i} \tag{4}$$

It is not necessary to apply the same recombination operator for objective variables and standard deviations. For example one can use *discrete recombination* for standard deviations and *intermediate recombination* for objective variables.

2.2 Maintaining Population Diversity Mechanisms

An evolutionary algorithm works properly (in terms of searching for a global solution) if the population consists of individuals different enough, i.e. the so-called diversity in the population is preserved. Yet many algorithms tend to prematurely loose this useful diversity and, as a result, there is possibility that population gets stuck in some local extrema instead of searching for a global one. To avoid this undesirable behavior in classical ES the mechanism of self-adaptation, as described above, was proposed. Yet this mechanism proves often not sufficient for very complex multi-modal problems.

In [5] the additional mutation operator was introduced. It improved obtained results, however it turned out that this additional mutation was still insufficient in the case of highly multi-modal problems of shape designing. The new mechanisms for maintaining population diversity in ES introduced in this paper include:

M1: the new mechanism of modifying standard deviation. Standard deviation is changed deterministically, in the following way:

$$\sigma'(\phi) = \sigma_{max} - (\sigma_{max} - \sigma_{min})\frac{(\phi - \phi_{min})^2}{(\phi_{max} - \phi_{min})^2} \tag{5}$$

where:

$\sigma_{max}, \sigma_{min}$ are predefined values of maximal and minimal mutation standard deviation;

ϕ_{max}, ϕ_{min} are maximal and minimal fitness found in all past generations.

The above function causes that the standard deviation of "poor" solutions is increased, so their children can "jump" in the solution space with greater probability. There are of course different function with the same characteristic and the one presented above was chosen arbitrary.

M2: the mechanism of crowding. The next generation population of μ individuals is generated from the population of λ individuals in the following way. The individuals in the offspring population are sorted on the basis of their fitness values. The best individual from the offspring population is added to the next generation population. Then the next individual is added to the population only if its Gaussian norm based distance from all k individuals already added to the population is greater than:

$$k \cdot d_{min} \cdot \left(1 - \left(\frac{t}{t_{max}}\right)^2\right), \quad \text{for } t = 1, 2, \ldots t_{max}. \tag{6}$$

where k is the number of individuals already present in the new generation population, d_{min} is the minimal distance between two individuals, t_{max} is the predefined maximal number of generations, and t is the actual number of generation. The third component of the above equation $\left(1 - (\frac{t}{t_{max}})^2\right)$ is decreasing as the number of generation increases (the analogy to "temperature" in simulated annealing technique). This mechanism works in such a way that "good" individuals can be more "crowded" than the "poor" ones

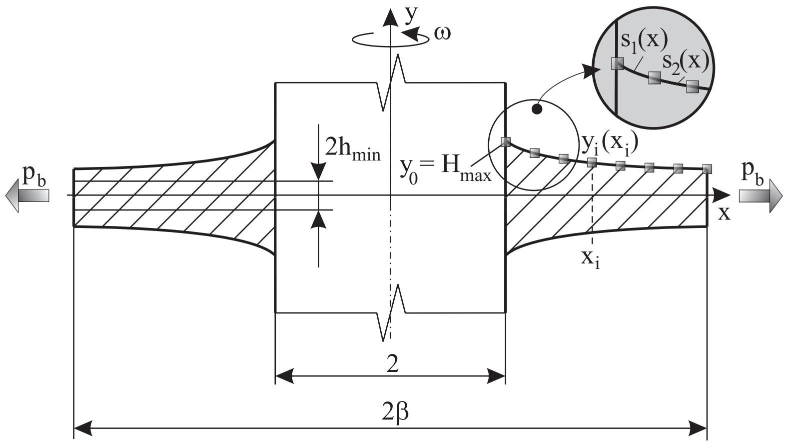

Fig. 1. Annular disc under consideration and its profile representation

but this restriction is loosen as the current number of generation increases. Thanks to this we can promote population diversity—individuals located in basins of attraction of "poor" local minima can not be crowded but they can be moved to the next generation population of μ individuals.

M3: elitism mechanism. The best individual from the population of μ parents is always moved to the population of λ offspring.

3 The optimization problem

The most important assumptions about the physical model of the discussed design problem are as follows [6]:

1. We consider an annular elastic disc of variable thickness $h = h(r)$ rotating with constant angular velocity ω and subject to uniform traction p_b at the outer radius b. The disc is clamped at the inner radius a.
2. The classical theory of thin discs with small gradient dh/dr is assumed and hence the stresses t_{ry} and s_y are neglected[1].
3. The material is linear-elastic with Young's modulus E, Poisson's ratio ν and subject to the Huber-Mises-Hencky (H-M-H) yield condition.
4. The small-strain theory is adopted.

The profile of the disc is represented by the 3^{rd} order spline built on equidistant nodes (see fig. 1), where a dimensionless radius $x = r/a$ was introduced.

[1] In this paper we use t and r symbols to denote the stresses instead of τ and σ, which are usually used in the literature, because the latter ones are already used in the evolution strategies description

After introducing basic equations one may formulate the optimization problem by defining a decision variables vector, a feasible region and an objective function. The decision variables vector:

$$\boldsymbol{Y} = (y_1, y_2, \ldots, y_n) \in M \subset \mathbf{R}^n \tag{7}$$

represents the shape of the disc in n equidistant points.

The feasible region:

$$M = \{\boldsymbol{Y} \in \mathbf{R}^n \mid k_d \cdot h_{min} \le y_j \le k_g \cdot H_{max} \quad \forall j = 1, \ldots, n\} \tag{8}$$

assumes that the disc is clamped at the inner radius having there fixed thickness ($y_0 = H_{max}$) and that the disc can be neither too thin (not thinner than $k_d \cdot h_{min}$) nor too thick (not thicker than $k_g \cdot H_{max}$). Objective function is described by the following formula:

$$\Phi = \left\{ c \left[\frac{1}{\beta - 1} \int_1^\beta s_i(x)\, dx \right] + (1 - c)\sqrt{p^2 + \Omega^2} \right\} \rightarrow max \tag{9}$$

where $0 \le c \le 1$ makes it possible to set the importance of each of the two criteria taken into account. The first of them (with the multiplier c) is connected with the equalization of the stress intensity and the second one with the external loadings (it is worth noting that if $c = 0$ this criterion becomes a simple maximization of elastic carrying capacity). Such a generalization is very helpful in estimating the limit carrying capacity or decohesive carrying capacity.

4 Experimental results

Optimal shapes in the meaning of criterion (9) connected with different ratio Ω/p were presented in [6]. Below only the analysis of the proposed mechanisms is discussed.

All the results of experiments presented in this section were obtained for the following values of the systems' parameters (the definitions of these parameters can be found in [5]): $\beta = 2$, $H_{max}/h_{min} = 5$, $p/\Omega = 0$, $k_d = 0.9$, $k_g = 1.1$, $\nu = 0.3$, $s_0/E = 0.001$, $c = 1$, $\mu = 15$, $\lambda = 100$, $\sigma_{min} = 10^{-6}$, $\sigma_{max} = 1$, $d_{min} = 0.13$.

The results for the basic version of the algorithm are presented in the figure 2. It can be observed that the population diversity is not maintained very well in this case. Usually it is quite quickly reduced after 20–30 (sometimes after about 100–150) generations. In the first case there is usually an individual with high fitness value (as compared to other individuals within the population) within the population but with very low standard deviation values. In such a case the population is quite quickly composed of the clones of such individual, the algorithm loses the population diversity and the abilities to explore the search space—usually the results obtained in such a case are very poor.

In the second case (when the population diversity falls down after 100–150 time steps) the algorithm locates the basin of attraction of one of the local

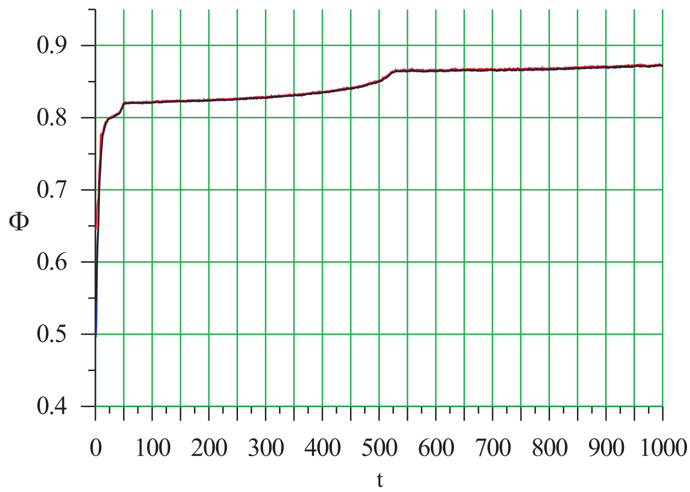

Fig. 2. Minimum, average and maximum value of the objective function for the base algorithm (with no additional mechanisms)

maxima. The population diversity is very low but the standard deviations are such that there is still the possibility of locating basins of attraction of other maxima (the values of standard deviations are analyzed in [6, 5]). Sometimes the algorithm locates the basin of attraction of quite "good" maxima but generally results obtained are not satisfying.

The additional operators presented in [5] had not greatly improved the results, and there still had been the tendency to lose the population diversity during optimization.

It was observed that the mechanism $M1$ (the deterministic standard deviation modification) applied separately led to worse final results as compared to the base algorithm. The main problem was related to the lack of convergence—the algorithm with the deterministic standard deviation modifications resembled random walk.

The use of the mechanism of crowding ($M2$) results in better results when the diversity of the population is considered but additionally causes that the algorithm is not "stable"—there are chaotic changes of the fitness values within the population (see fig. 3). When we additionally introduce the elitism mechanism ($M2$) the chaotic fluctuations of the fitness values disappear—the algorithm works "stable" (fig. 5). The application of all three mechanisms together (see fig. 5) results in maintaining population diversity, there are no chaotic changes in fitness values of the individuals, and additionally the possibility of dominating the whole population by the copies of the individual with small fitness value and small values of standard deviations is reduced. Also in this case the average results obtained are the best.

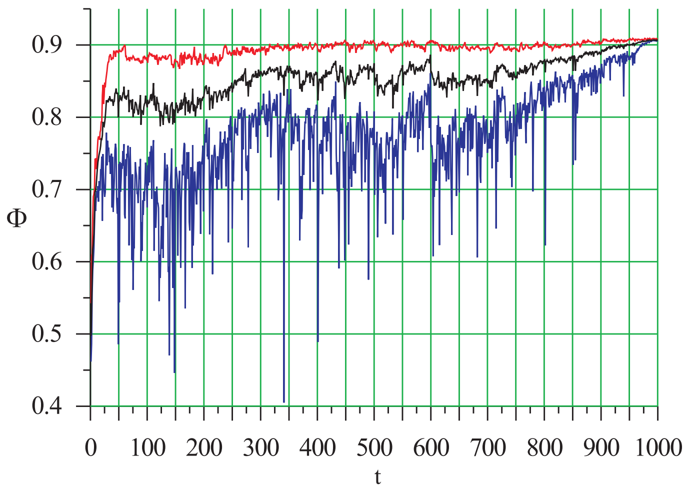

Fig. 3. Minimum (bottom line), average (middle line) and maximum (top line) value of the objective function for mechanism $M2$ (crowding)

5 Concluding remarks

Many tasks related to optimal designing cannot be solved by the use of classical methods (e.g. calculus of variations) because of various reasons. It is the case for example when there is no strict mathematical model of a problem (e.g. the mapping between the decisive variables and the objective function is unknown). In such a situation the optimization process is performed as a sequence of evaluations of possible solutions. When the domain space make the complete searching impossible one can use some heuristic methods to control the algorithm of candidate solutions selecting, like evolutionary algorithms or simulated annealing. Both the techniques in their classic forms usually do not work correctly in problems with many local extrema. In case of evolutionary algorithms one of the key problems is related to maintaining population diversity.

In the paper the following three modifications to the classic (μ, λ)-ES were discussed: deterministic modification of standard deviations, crowding and elitism. The analysis was based on the results of the shape optimization of rotating variable-thickness annular elastic disc. The results clearly show that only the simultaneous use of all these three mechanisms help to maintain population diversity and, in consequence, lead to a more stable searching processes and finally—better solutions. Future research could concentrate on the further verification of the proposed mechanisms. Other engineering problems should be considered for this purpose.

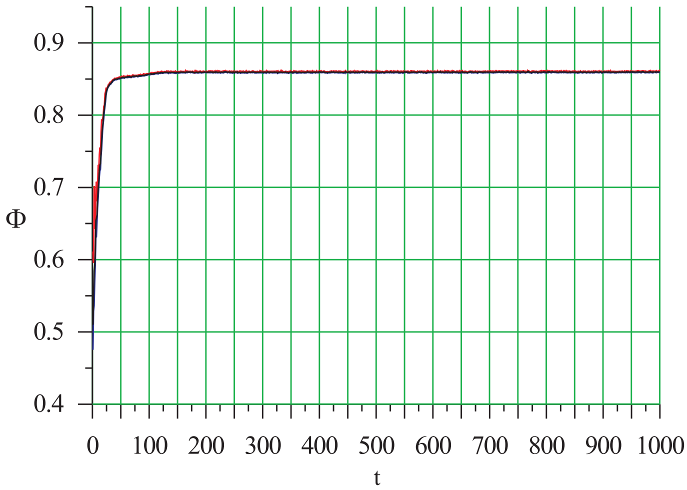

Fig. 4. Minimum, average and maximum value of the objective function for mechanism $M3$ (elitism)

References

1. T. Bäck, D. Fogel, D. Whitley, and P. Angeline. Mutation. In T. Bäck, D. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*. IOP Publishing and Oxford University Press, 1997.
2. T. Bäck and H.-P. Schwefel. An overview of evolutionary algorithms for parameter optimisation. *Evolutionary Computation*, 1(1), 1993.
3. T. Bäck and H.-P. Schwefel. Evolutionary computation: An overview. In *Proceedings of the Third IEEE Conference on Evolutionary Computation*, Piscataway NJ, 1996. IEEE Press.
4. L. Booker, D. Fogel, D. Whitley, and P. Angeline. Recombination. In T. Bäck, D. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*. IOP Publishing and Oxford University Press, 1997.
5. R. Dębski, R. Dreżewski, and M. Kisiel-Dorohinicki. Preserving diversity in evolution strategy for shape design of rotating elastic disc. In *Proceedings of the 6th Conference on Evolutionary Algorithms and Global Optimization (KAEiOG 2003)*. University of Zielona Góra, 2003.
6. R. Dębski, G. Woźniak, R. Dreżewski, and M. Kisiel-Dorohinicki. Evolutionary strategy for shape design of rotating variable-thickness annular elastic disc. In *Proceedings of the Symposium on Methods of Artificial Intelligence (AI-METH 2002)*, Gliwice, 2002. Silesian University of Technology.
7. G. Rudolph. Evolution strategies. In T. Bäck, D. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*. IOP Publishing and Oxford University Press, 1997.
8. H.-P. Schwefel. *Numerical Optimization of Computer Models*. John Willey & Sons, Ltd., Chichester, 1981.

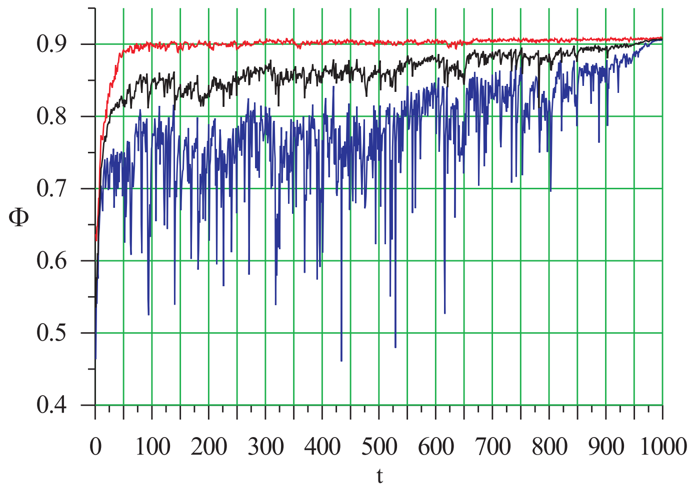

Fig. 5. Minimum (bottom line), average (middle line) and maximum (top line) value of the objective function for combined mechanisms $M2$ and $M3$ (crowding and elitism)

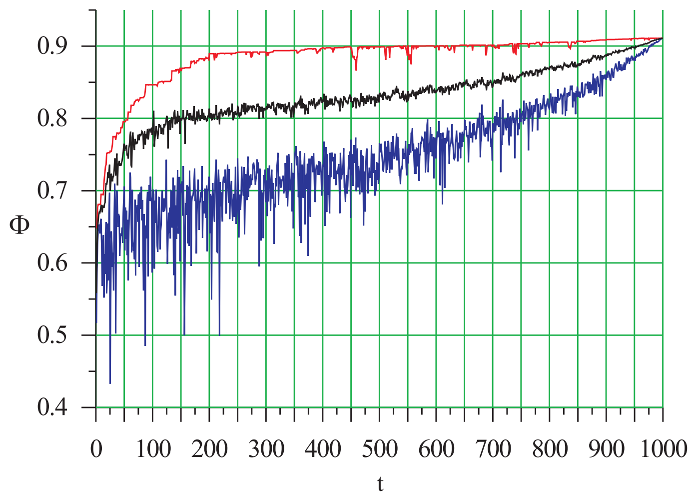

Fig. 6. Minimum (bottom line), average (middle line) and maximum (top line) value of the objective function for combined mechanisms $M1$, $M2$ and $M3$ (the deterministic standard deviation modification, crowding and elitism)