

# Optimal Number of Evolution Strategies Mutation Step Sizes in Dynamic Environments

Lutz Schönemann

Systems Analysis Group, Department of Computer Science, University of Dortmund, Germany

lutz.schoenemann@cs.uni-dortmund.de

**Categories and Subject Descriptors:** G.1.6: Numerical Analysis: Optimization

**General Terms:** Algorithms

**Keywords:** Evolution strategies, number mutation step sizes, dynamic environments

## 1. INTRODUCTION

Evolutionary algorithms (EA) mimic nature's mechanisms of evolution and have proven as good problem solvers for parameter optimization problems of the type  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . In dynamic optimization problems the function definition is ascribed to the static function by

$$f_d(x, t) := f(x - x_t^*) .$$

The optimum  $x_t^*$  moves during the optimization process and therefore depends on the time  $t$ .

In the case of dynamic environments the main task is not to find one good solution. Instead, the algorithm must follow the moving target with a small distance. For this purpose many evolutionary algorithms consists of self-adaptive mechanisms. This means, that they can adjust their parameter settings during the evolutionary run. In evolution strategies (ES) [1] the adaptation is done on a set of strategy parameters which influence the variation of the object variables. One variant of ES consists of up to  $n$  mutation step sizes. On the one hand, using a separate step size for every single coordinate the adaptation of the algorithm to the problem at hand is regarded to work better. On the other hand, the higher the number of strategy parameters the higher is the time needed for adaptation [2]. Therefore, one must find the silver bullet between a fast and a good adaptation by choosing an appropriate number of step sizes.

Several options are lying between the two extremes of one and  $n$  mutation step sizes. A frequent choice is to choose two mutation step sizes. In this case, usually the first step size is used for one coordinate and the second step size is used for the others. Mostly, it is not known in advance which of the  $n$  coordinates should be varied by the separate step

size. This difficulty holds for static as well as for dynamic environments.

In this study we take a first step to investigate the influence of the used number of different mutation step sizes in dynamic environments. To make things easier we concentrate on the two extremes of one and  $n$  mutation step sizes. Depending on the obtained results it is reserved to further studies to investigate other choices.

## 2. EXPERIMENTAL INVESTIGATIONS

Our experimental investigations are conducted on three test functions. Every test function is a representative of a class of similar optimization problems. We use the

- well-known sphere model as a unimodal function,
- Ackley function as a multimodal function of moderate difficulty and
- Rastrigin function as a representative of the class of multimodal functions of high difficulty.

For the case of dynamic environments various types of dynamics are known. In our study we restrict the dynamics to a continuously changing environment. More exact, the moving frequency is  $\Delta g = 1$  meaning that the optimum moves before every generation. We call this *pseudo-continuous movement*. The optimum moves in one direction with a constant severity. It is hoped that the algorithm could adapt to the mean severity and follows the optimum with a small distance.

For every test run the optimum is relocated by one of three moving types:

- I: The optimum moves only in one (the first) coordinate, all other coordinates remain constant.
- II: The optimum moves in all  $n$  coordinates. In every coordinate the moving strength is equal.
- III: Here, the optimum moves in all coordinates, too. But the covered distance is different in every coordinate. The covered distance in every coordinate increases from the first coordinate to the last one.

The results for the different moving types are comparable because the total severity  $s$  is always the same.

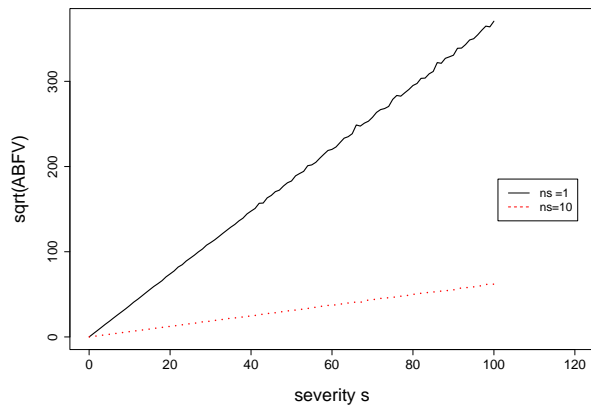
The experiments are conducted with a standard parameterization of a (15,100)-ES. The first population is initialized in the optimum. The initial mutation step sizes match the expected distance of the moving optimum in the next

step. Hence, no time is needed to reach the tracking phase in which the ES follows the optimum within a certain mean distance. Therefore, the used run time of 1000 generations is sufficient for informative results.

To compare the results for different strategies and moving types we use the performance measure **average best function value (ABFV)**. Abridging, the ABFV is a robust measure for the average fitness of a mean run and is described in more detail in [3]. In the case of minimization, a smaller ABFV means a better behavior.

The first test function used here is the well-known sphere model. This is a simple function with one local optimum, which is therefore also the global one. In the static case the consideration of this function serves for measuring the convergence velocity of an algorithm.

If the optimum moves following moving type I the usage of one mutation step size seems to be a very inappropriate choice, because the ES must adapt one step size to the moving severity whereas for the other step sizes a value near zero is optimal. Hence, the results of such a strategy are very poor compared to an ES with  $n$  mutation step sizes (Fig. 1). Notice, that on the  $y$ -axis the square root of the

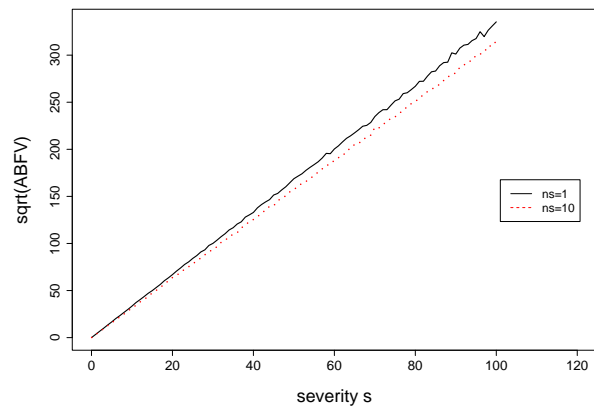


**Figure 1: ABFV of an (15, 100)-ES with  $n_\sigma = \{1, 10\}$  mutation step size/s on the dynamic 10-dimensional sphere. The optimum moves every generation in one coordinate (moving type I) with a total severity  $s$ .**

ABFV is plotted. The linear run of the curve proves that the results perfectly depend quadratically on the moving severity  $s$ . The ES is able to follow the moving optimum for the moving severities tested here. Random samples assist the assumption that this assertion is true even for much higher severities. But the distance may be too high.

If the optimum moves in all dimensions with an equal severity only one step size is necessary. Because an ES with  $n$  step sizes must adapt  $n$  strategy parameters instead of only one, one could assume that it would perform worse than an ES with one mutation step size. The experiments revealed that the curves of both strategies are nearly the same. This means that the ES with  $n$  step sizes behaves not worse than the ES with only one step size.

Figure 2 shows the results for moving type III. Here, the ES with  $n$  step sizes performs slightly better. But the differences are smaller than expected. One reason may be that in this situation the trade off to adapt  $n$  step sizes to an value significant greater than zero is so high that the benefits of



**Figure 2: ABFV of an (15, 100)-ES with  $n_\sigma = \{1, 10\}$  mutation step size/s on the dynamic 10-dimensional sphere. The optimum moves every generation in all coordinates (moving type III) with a total severity  $s$ .**

the different step sizes is compensated. If this holds we expect that for much higher problem dimensions the situation changes meaning that the ES with one step size performs better. Indeed, this could be observed for  $n = 30$  dimensions.

From the point of characteristic of the curves the results for the three moving types on the Rastrigin function are comparable to the ones on the sphere function. For both ES the ABFV increases quadratically with the moving severity. Due to the problem hardness the ABFV for the Rastrigin function are worse than the ones on the sphere. Again, for moving type I the ES with  $n$  step sizes is superior to the ES with only one step size. As it was already seen on the sphere model, there exist no significant differences between the results of the two ES variants when the optimum moves following type II and III.

The overall behavior of the two ES variants on the Ackley function is the same. In contrast to the two other functions the ABFV does not depend quadratically on the moving severity. Anew, it could be observed that the ES with  $n$  mutation step sizes performs better than or as good as the ES with one mutation step size.

## Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the *Collaborative Research Center "Computational Intelligence" (SFB 531)*.

## 3. REFERENCES

- [1] H.-G. Beyer and H.-P. Schwefel. Evolution strategies: A comprehensive introduction. *Natural Computing*, 1(1):3–52, 2002.
- [2] G. Rudolph. On correlated mutations in evolution strategies. In R. Männer and B. Manderick, editors, *Proc. Parallel Problem Solving from Nature 2*, pages 105–114, Amsterdam, 1992. Elsevier.
- [3] L. Schönemann. The impact of population sizes and diversity on the adaptability of evolution strategies in dynamic environments. In *Proc. 2004 Congress on Evolutionary Computation (CEC 2004)*, pages 1270–1277. IEEE Press, 2004.