FACTA UNIVERSITATIS (NIŠ)
SER.: ELEC. ENERG. vol. 22, no. 2, August 2009, 159-174

Multi-Objective Optimization Using Evolution Strategies

Michael Jaindl, Alice Köstinger, Christian Magele, and Werner Renhart

Abstract: The present paper gives an overview of different versions of Evolution Strategies, namely the (1+1) Evolution Strategy, the Higher Order $(\mu/\rho,\ \lambda)$ Evolution Strategy and the Niching $[\kappa(\mu/\rho,\lambda)]$ Evolution Strategy, and how these methods can be applied to problems in Electrical Engineering. Significant features of the algorithms implemented by the authors are presented. Finally, results are discussed on three electromagnetic optimization problems.

Keywords: Multi-objective optimization, multi-modal optimization, stochastic optimization, evolution strategy.

1 Introduction

OPTIMIZATION both with deterministic and stochastic methods has been applied successfully over the last decades in all fields of engineering, including electrical engineering. When solving practical electromagnetic problems where no hypothesis can be made a priori, deterministic methods often converge to one of the function's local minima. Therefore, if there is no or very little knowledge about the behavior of the objective function, about the presence of local minima or the location of feasible and nonfeasible regions in the multidimensional parameter space it is advisable to start the optimization process with a stochastic strategy. Stochastic methods choose their path from one or more initial configurations to the final solution(s) through the parameter space using some "random factor" and they accept deterioration in the objective function during the iteration process. This fact enables them to act more "global" than their deterministic counterparts. High stability in convergence even when there is little or no knowledge about the behaviour

Manuscript received on June 14, 2009.

The authors are with Graz University of Technology, Institute for Fundamentals and Theory in Electrical Engineering, Kopernikusgasse 24/3, A-8010 Graz, Austria (e-mail: Christian.Magele@TUGraz.at).

of the objective(s) and a rather simple implementation complete the set of advantages. The main drawback, however, is the high number of function evaluations. But this drawback was extremely weakened by the recent enormous increase in computer power. Furthermore, the availability of parallel computer environments has promoted their applicability to real world problems in electrical engineering, since all higher order Evolution Strategies are intrinsicly parallel. In the following sections the main features of Evolution Strategies will be discussed and each of the presented methods will be applied to an electromagnetic optimization problem. The problems and the results will not be discussed in complete detail.

2 Evolution Strategies

Evolution Strategies use the principles of organic evolution as rules for the optimum seeking process [1]. Just a few statistical processes acting on and within populations and species are responsible for the vast majority of life. These processes are *reproduction, mutation, competition and selection* [1].

3 (1+1) Evolution Strategy

The (1+1) Evolution Strategy [2] is a simple mutation-selection scheme (Fig. 1) called two membered Evolution Strategy. The ”population” consists of one parent only, determined by a certain parameter configuration with a respective number of stepsizes, one for each parameter. The quality of this configuration is determined by the given objective function $f(\boldsymbol{p})$. The basic evolutionary operation is **mutation**, the generation of one descendant d by means of adding a normally distributed random vector to the parameter values of the parent p. This distribution makes smaller changes more likely than larger ones.

$$\boldsymbol{p}_d^k = \boldsymbol{p}_p^k + \boldsymbol{\nu}(0, \sigma^k(\alpha)) \tag{1}$$

The ”fitter” of both individuals, obtained by evaluating the objective function, serves as the ancestor of the following iteration (**selection**). After a certain number of functions calls (= generations) it is essential to adjust the stepwidth (e.g. after $10 * n$ function calls, where n is the number of optimization parameters). At first the ratio of successful mutations over all mutations $p(pos)$ is evaluated. If $p(pos)$ is less than a specified threshold, all stepsizes σ are increased by dividing through a stepsize factor sf, in case $p(pos)$ is larger than a specified threshold, all stepsizes σ are decreased by multiplying by a step size factor sf. In the classic implementation [2] the strategy parameter $p(pos)$ is set to $p(pos) = 0.2$, the stepsize factor sf is set

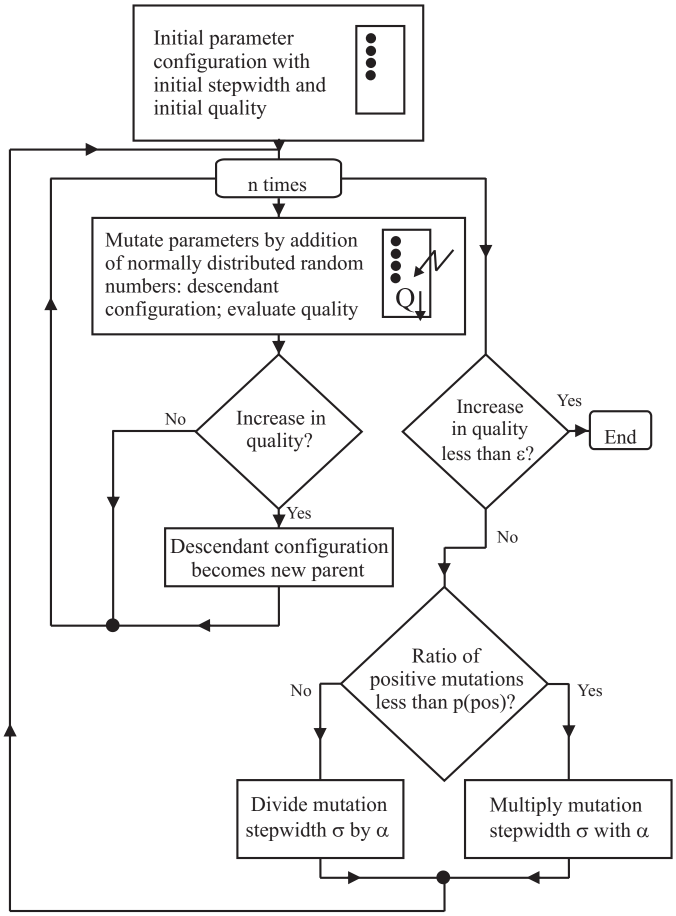

Fig. 1. Flowchart of the (1+1) ES.

to $sf = 0.85$. The iterative process is ended when a prescribed stopping criterion (maximum number of function calls, proper ε criterion) is met.

3.1 Optimal design of an electromagnet

The (1+1) Evolution Strategy will be applied to design the pole face of an electromagnet (Fig. 2) to adjust the magnetic field in the region of interest in a prescribed way. Fig. 3 shows a zoomed view of the problem together with the four trial variables p_1 to p_4.

The electromagnet is made from non-linear magnetic material, the forward problem is solved using the finite element method package EleFAnT2D [3]. The task is to obtain a homogeneous field of $|B_z| = 0.02$ T in the region of interest, which is right in the middle of the air gap and covers 80% of the area of the pole. After 240 iterations the (1+1) Evolution Strategy ended in the solution presented in Fig. 5, the diagrams of the initial and the optimized magnetic fields are given in Fig. 6 and Fig. 7.

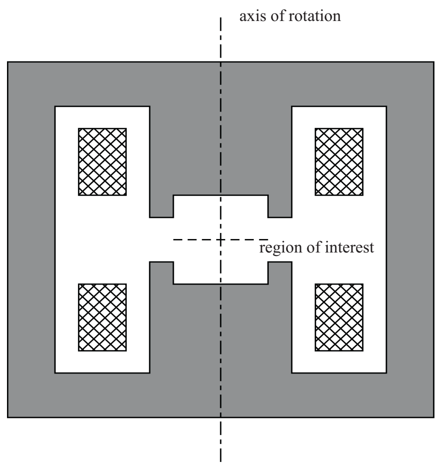

Fig. 2. Full Model of the Electromagnet.

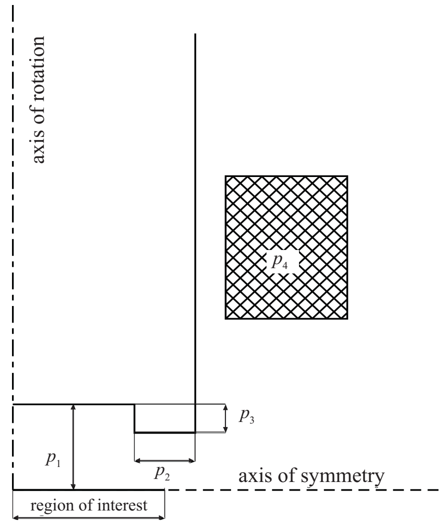

Fig. 3. Zoomed View of the Region of Interest.

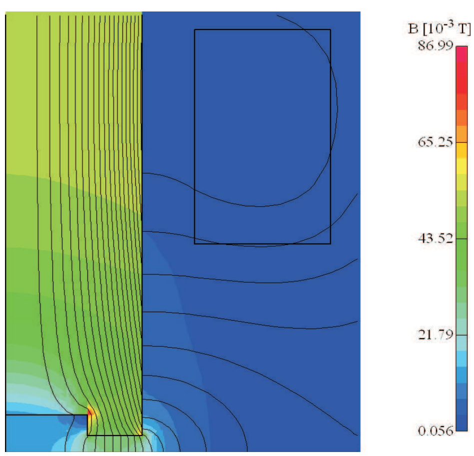

Fig. 4. FEM Solution of the initial configuration.

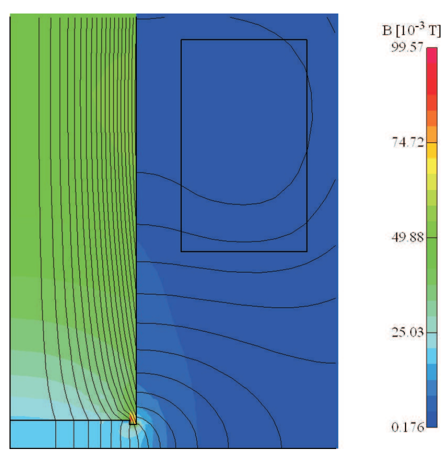

Fig. 5. FEM Solution of the optimal configuration.

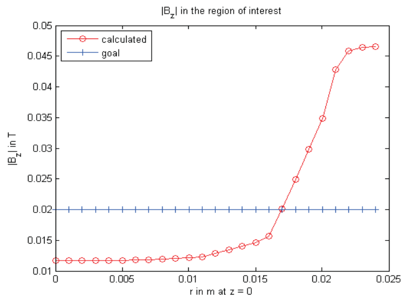

Fig. 6. Diagram of the magnetic field in the region of interest of the initial configuration.

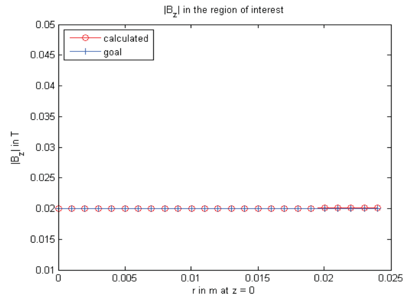

Fig. 7. Diagram of the magnetic field in the region of interest after 240 iterations using the (1+1) ES.

4 Higher Order Evolution Strategies

The (1+1) scheme does not take **population** into account. A first step to overcome this drawback is to introduce more than one descendant. This leads to a multi-membered Evolution Strategy, a $(1+\lambda)$ strategy featuring λ descendants [4]. In this strategy, both the parent and the descandents have to undergo selection. Since another vital characteristic of biological evolution can be found in the **limited lifetime** of an individual, in higher order $(1,\lambda)$ Evolution Strategies the parent configuration is no longer subject to the selection process. The consideration of μ parents rather than one allows the imitation of sexual reproduction by introducing **recombination** of ρ parents prior to the mutation step, leading to a $(\mu/\rho,\lambda)$ strategy [4].

5 $(\boldsymbol{\mu}/\boldsymbol{\rho},\boldsymbol{\lambda})$ Evolution Strategy

Fig. 8 displays the flow chart of the $(\mu/\rho,\lambda)$ Evolution Strategy. A number of ρ individuals is selected at random to form the first generation, the respective stepsize are assigned and all fitness values are evaluated.

Recombination is performed using arithmetic crossover. A number of ρ configurations from the current population is chosen to produce ρ children using (2) and (3). This procedure is repeated until a completely new, still unmutated population exists.

$$\boldsymbol{d}^1_{unmutated} = a\boldsymbol{p}^i + (1-\alpha)\boldsymbol{p}^j \tag{2}$$

$$\boldsymbol{d}^2_{unmutated} = (1-\alpha)\boldsymbol{p}^i + a\boldsymbol{p}^j \tag{3}$$

The parents participating in this process are selected under consideration of their fitness. Applying a roulette wheel selection or a tournament selection fitter

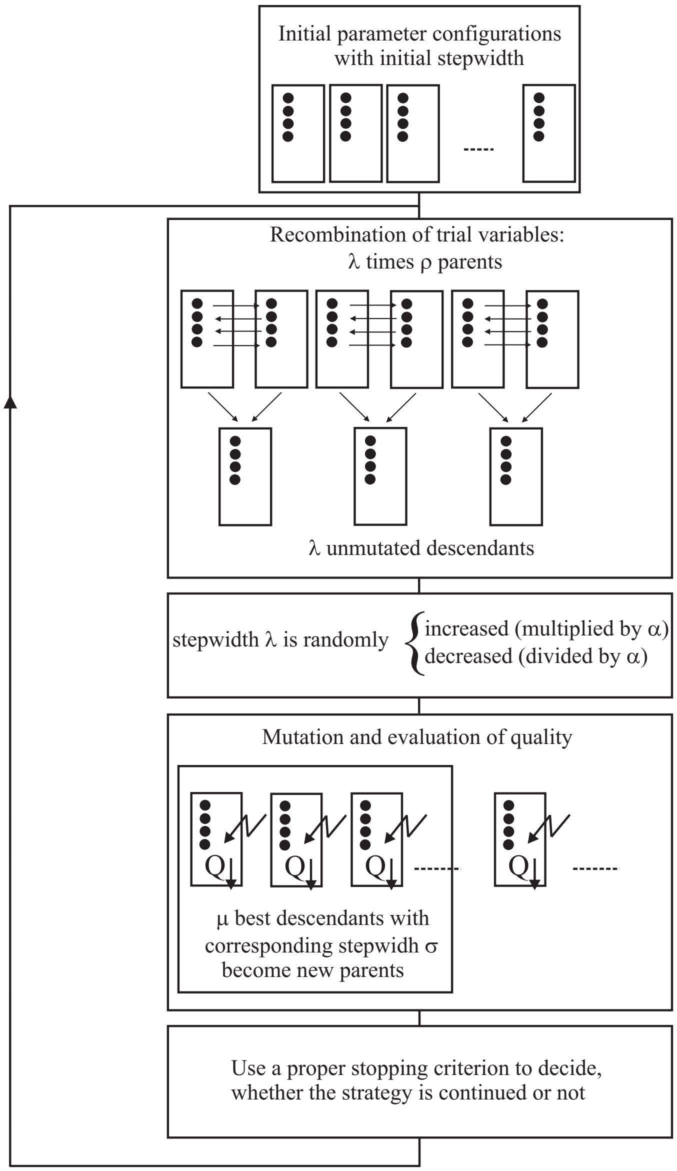

Fig. 8. Flowchart of the $(\mu/\rho, \lambda)$ ES.

individuals will have a remarkably greater chance to be chosen (**implicit** or **mating selection**).

Then all individuals have to undergo mutation. Unlike in the (1+1) Evolution Strategy, all stepsizes are adapted from generation to generation. Each descendant has inherited the stepsizes from the parents, which are slightly decreased or increased at random using (4) prior to mutation.

$$\begin{aligned}\sigma^{(i)}_{descendant} &= (\sigma^{(i)}_{descendant,inherited}) * \alpha \quad \text{or} \\ \sigma^{(i)}_{descendant} &= (\sigma^{(i)}_{descendant,inherited}) / \alpha \end{aligned} \tag{4}$$

Finally, ρ **new** parents are selected form the λ descendants using **explicit** or **environmental** selection to form the next generation. The whole process is repeated until a prescribed stopping criterion is met. Among others the crowding behaviour can be used to decide whether to stop or not. To do so, the level of crowding of the current population is evaluated using a ”mean” individual $\bar{x}$

$$\bar{x} = \frac{1}{\lambda}\sum_{i=1}^{\lambda} x_i \tag{5}$$

and the crowding radius cr, where λ is the number of descendants and n is the number of parameters

$$cr = \frac{1}{\lambda n}\sum_{j=1}^{n_p}\sum_{i=1}^{\lambda}\left(\frac{x_{j,i} - \bar{x}_j}{\bar{x}_j}\right)^2. \tag{6}$$

If cr falls under a certain value, the optimal solution is found.

5.1 Model of a magneto-rheologic clutch

Fig. 9 shows an axi-symmetric sectional drawing of the model to be optimized. The coil is responsible for the generation of the magnetic field. The main body and the primary disks are rigidly coupled, while the secondary disks make up the part of the system to be coupled. The space between the disks is filled with a magneto-rheologic fluid. Magneto-rheologic fluids exhibit a distinct enlargement of their viscosity when being exposed to a magnetic field. The reason for this behaviour can be found in the countless micro particles dissolved in a base liquid, which are distributed randomly within the base liquid in the absence of a magnetic field. However, as soon as the fluid is exposed to a magnetic field, the micro-sized magnetic particles tend to form rigid chains and hence enlarge the viscosity of the fluid remarkably. In case of the clutch under consideration it looks as if the two disks connected by a gap of fluid were coupled rigidly.

The total of 13 parameters are summarized in Fig. 10.

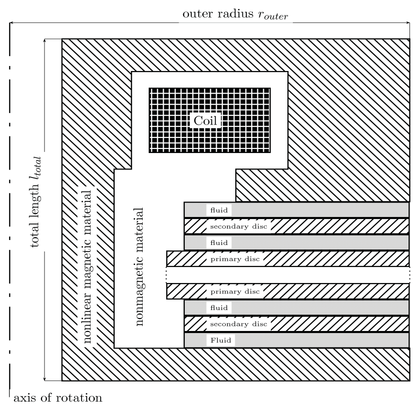

Fig. 9. Model of a magneto-rheologic clutch.

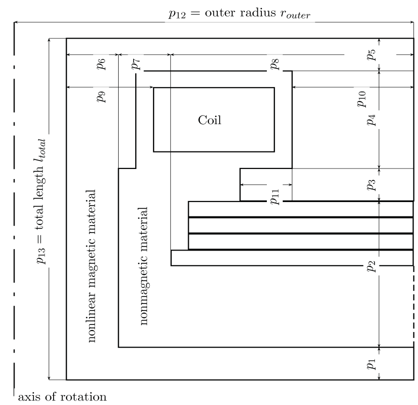

Fig. 10. Trial variables of the model to be optimized.

5.2 Definition of the objective function using fuzzy membership functions

Most real world problems have to treat more than a single objective leading to multi-objective optimization problems. In case of the electromagnetic clutch one has to fulfill the requirement of a total torque of 600Nm while keeping the volume of the steel parts as low as possible. The two conflicting objectives could be termed **best fit** objective and **minimum** objective, respectively. In general, the levels of satisfaction of the two objectives have to be merged into a single number to be used with standard optimization methods. One way to perform this task is to normalize each individual objective using fuzzy membership functions [5], which are then merged into a single function by applying appropriate inference rules. A **best fit** objective (**The overall torque should be as close as possible to 600Nm**) can be easily modeled by a full bell shaped convex function as shown in Fig. 11. If the desired value is reached, μ becomes 1, whereas in all other cases μ takes a value less than 1. On the other hand a minimum objective (**The volume of the steel part should be as small as possible**) can easily be represented by a half bell shaped function as shown in Fig. 12

The designer has to specify the desired values and the ”90% region” in order to evaluate m, l and r. Then the level of satisfaction can be calculated using (7) and (8)

$$\mu_T(\boldsymbol{p}) = \begin{cases} e^{-l(x-m)^2}, & x \leq m \\ e^{-r(x-m)^2}, & x > m \end{cases} \tag{7}$$

$$\mu_V(\boldsymbol{p}) = e^{-r(x-m)^2}, \quad x > m. \tag{8}$$

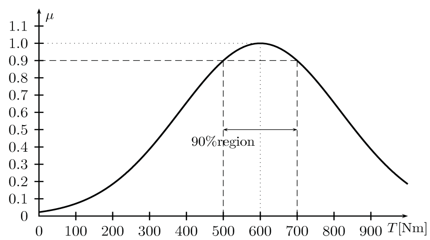

Fig. 11. Fuzzy membership function for the torque objective.

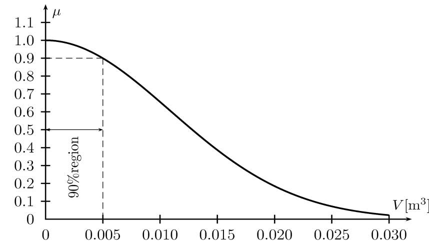

Fig. 12. Fuzzy membership function for the volume objective.

The two fuzzy functions μ_T (7) and μ_V (8) can be transferred into a single scalar value by using an appropriate inference rule given in (9), where more or less emphasis can be put on either objective using the weights w_1 and w_2.

$$f(\boldsymbol{p}) = (w_1 + w_2) - w_1 \mu_T - w_2 \mu_V \tag{9}$$

The objective function $f(\boldsymbol{p})$ (9) has to be minimized with respect to the 13 parameters given in Fig. 10.

5.3 Optimal design of a magneto-rheologic clutch

At first, an initial configuration is presented. It has been taken randomly from the first generation of the (4/2,20) ES and the field plot is shown in Fig. 13.

Using the objective function given in (9) and setting $w_1 = w_2 = 1$, one obtains an optimal configuration shown in Fig. 14.

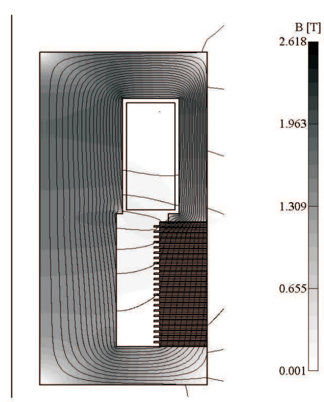

Fig. 13. Field plot of the initial configuration.

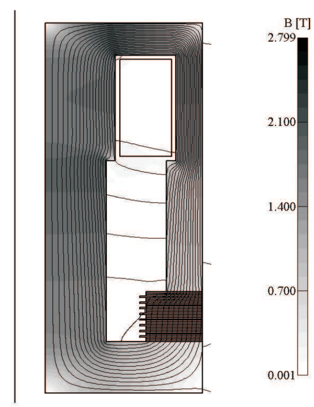

Fig. 14. Field plot of the optimal configuration.

Table 1 summarizes the characteristics of these configuration. Taking into account that the density of steel is approximately 8000 kg/m^3, the weight of the steel parts and the torque to weight ratio is also given.

Table 1. Characteristics of an initial configuration.

Torque [Nm]	Volume [m^3]	Weight [kg]	Torque/Weight [Nm/kg]
167.55	0.000987	7.9	21.21
599.8	0.001296	10.37	58.23

6 Niching [$\kappa(\mu/\rho,\lambda)$] Evolution Strategy

Recombination is usually done by arithmetic crossover, taking the fitness of all parents under consideration into account. Prior to this step the whole generation is clustered using a complete linkage algorithm [6]. In the initial phase of the iteration process a defined number of clusters is applied. Recombination within a cluster is performed with a higher probability than recombination between different clusters Fig. 15. This leads very soon to a certain number of sub-populations, several of them gathering around local solutions. The best solution of each cluster is stored in order to remember all single local solutions. Having a closer look at the clustered population it is possible to detect isolated subpopulations and to adapt the number of clusters κ to the current situation. Since the total size of the population depends on ρ (the number of parents) and κ by $n_{pop} = \rho * \kappa$, a reduced number of clusters helps to save computational effort and speeds up the iteration process. It has also turned out that, in the due course of the optimization process, one local solution after the other is given up, until all individuals gather around the final solution only. The impact of this clustering/cluster-sensitive recombination operator is demonstrated using a multimodal modified Rosenbrock testfunction.

6.1 Modified rosenbrock function

The well known Rosenbrock function is modified (10).

$$\begin{aligned} f(\boldsymbol{p}) = & (100\,x_2 - x_1^2)^2 + (1 - x_1)^2 \\ & - 50((x_1+1)^2 + (x_2-1)^2) \\ & - 10((x_1-1.5)^2 + (x_2-2.5)^2 + e^{(2x_2-5)}) \end{aligned} \tag{10}$$

$f(\boldsymbol{p})$ shows not only one, but three distinct optimal solutions, a global and two local ones as given in table 2. The contour plot can be seen in Fig. 16.

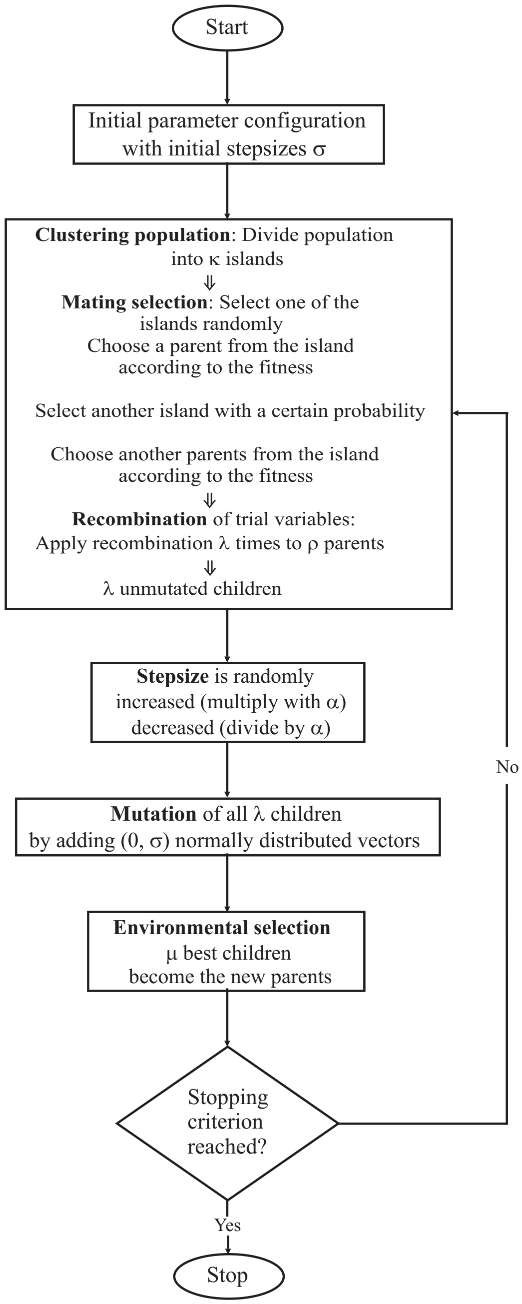

Fig. 15. Flow Chart of the Niching $[\kappa(\mu/\rho,\lambda)]$ Evolution Strategy.

6.2 Detecting local solutions

Fig. 16 shows an intermediate situation where four distinct clusters can be identified. This is done in such a way that no configuration from any another cluster can be found inside the hypershpere of the cluster under investigation. In case such a situation is recognized, the best solutions of each isolated cluster are stored for later use. The number of clusters κ in the NES can now be adjusted referring to the number of isolated clusters.

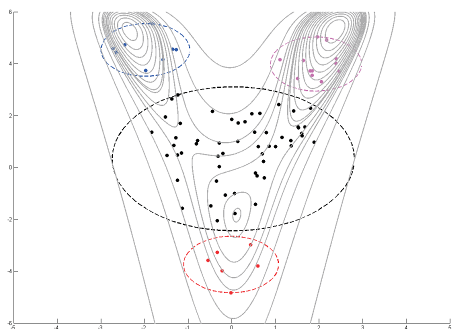

Fig. 16. Four Isolated Clusters.

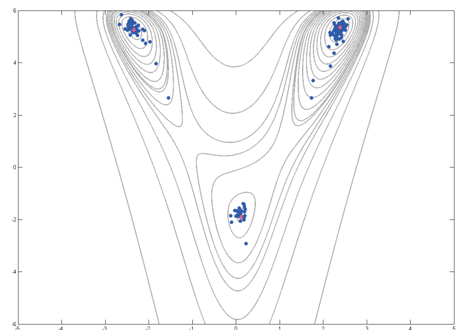

Fig. 17. Best Solutions of Isolated Clusters.

At the end of the optimization process all best isolated solutions, which are shown in Fig. 17 are evaluated again. It turns out that most of these configurations are in the vicinity of the three actual solutions.

The best solutions of each of the three clusters in Fig. 17 are given in Table 3. It can be seen that they correspond very well with the exact values given in table 2.

Table 2. Global and Local Solutions of Modified Rosenbrock Function.

Solution	x_1	x_2	$f(\boldsymbol{p})$
Global	2.38	5.32	-1296.44
Local 1	-2.34	5.29	-955.71
Local 2	0.115	-1.84	-328.75

6.3 Global behaviour of the niching evolution strategy

An additional feature of NES is that they reveal an extended global behaviour compared to a standard Higher Order $(\mu/\rho, \lambda)$ Evolution Strategy. If the modified

Table 3. Evaluated Global and Local Solutions of Modified Rosenbrock Function.

Solution	x_1	x_2	$f(\boldsymbol{p})$
Global	2.39	5.356	-1296
Local 1	-2.34	5.274	-955.6
Local 2	0.117	-1.894	-328.6

Rosenbrock problem is run repeatedly, the $(\mu/\rho,\lambda)$ Evolution Strategy ends up in all of the three possible solutions with a certain probabilty, as shown in Table 4.

Table 4. Probability of Ending up in one of the Possible Solutions.

Solution	%
Global	47
Local 1	17
Local 2	36

On the other hand, if the Niching $[\kappa(\mu/\rho,\lambda)]$ Evolution Strategy is run to solve the problem, the strategy ends up in the global solution every time. This means that the probability of finding the global or a very good local solution is increased applying the Niching Evolution Strategy.

6.4 Front of non dominated solutions

Applying a Niching Higher Order Evolution Strategy and having a closer look at all local solutions obtained with certain weights in (9) it turns out that most of them are on or close to the front of non dominated solutions [7]. Additionally, a certain combination of w_1 and w_2 covers only a part of the front of non dominated solutions. This leads to the aspect to introduce more than one objective function in parallel or in series. Since different w_1 and w_2 put more or less emphasis on one or the other objective, a larger part of the pareto front can be detected.

6.5 Optimal design of the magnetic shunt problem

The 2D topology of the magnetic shunt problem [8], [9] is shown in Fig. 18. The tank wall is made of steel and it is the only part in the model where eddy currents can flow. The magnetic shunt is made up from several slices (the number was set to five) which agree in thickness but can have different widths (p_1 to p_5). The designer is free to subdivide the two dimensional area into appropriate parts for the magnetic shunts, for a slice of copper (which is not taken into account in this paper) as well as for the necessary air regions. Fig. 18 displays the respective trial variables p_6 to p_{10}. A three phase system with 1000A/0º, 1000A/120º and 1000A/240º (from left to right in Fig. 18) is responsible for the resulting magnetic fields.

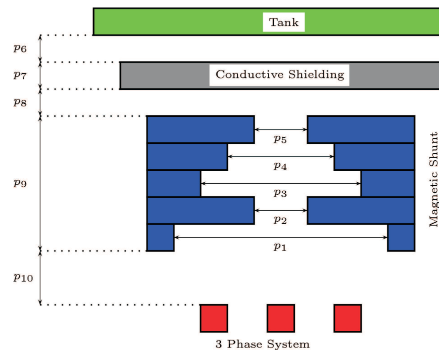

Fig. 18. Topology of the Shunt Problem with Trial Variables p_1 to p_{10}.

A $[\kappa(\mu/\rho,\lambda)]$ NES with $\kappa = 10$ initial clusters, $\mu = 4$ parents with $\rho = 2$ of them being involved in recombination leading to $\lambda = 10$ children is used to solve the above magnetic shunt/shielding problem. The electromagnetic field problems are solved using the 2D Finite Element code ELEFANT2D [3]. In a first stage only magnetic shunts are taken into account. Ten trial variables (Fig. 18) can be adjusted to fix the losses and to minimize the volume of the shunts. In a sequential way, four different objectives functions (9) with different weights are used and given in Table 5.

Table 5. Different Combinations of Weighting Factors.

w_1	w_2
1	1
3	1
5	1
1	3

Each problem was run several times, leading to a an optimal solution for the respective objective function. As an example, the best result obtained with $w_1 = 1$ and $w_2 = 1$ is given in Fig. 19. For comparison, Fig. 20 shows the field plot of a configuration having a massive magnetic shunt with the same overall dimensions as the optimal one.

Assessing the optimized model the volume was decreased by 53% and the eddy

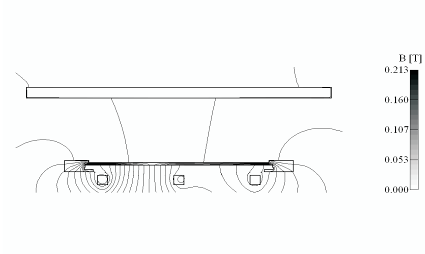

Fig. 19. Optimized Configuration of Magnetic Shunts.

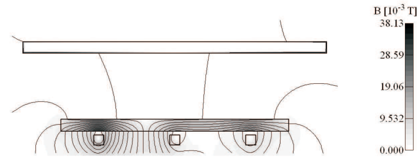

Fig. 20. Configuration with Massive Magnetic Shunt.

current losses by 40% compared to a massive shunt of the same dimensions.

6.6 Non dominated solutions

Among all solutions there are a few with very special characteristics. If the value of one of the objectives of these configurations is improved, at least one of the other objectives is made worse. All these configurations make up the front of non dominated solutions (or the Pareto Optimal Front). To get more inside into the front of non dominated solutions, the best solutions of each of the ten clusters were stored and evaluated. The non dominated ones of each problem (different w_1, w_2) were selected and are plotted in Fig. 21, together with a vast number of solutions obtained in the different optimization runs and the respective front of non dominated solutions. It can be seen that the cluster solutions are quite close to the front of Pareto optimal solutions. Furthermore, in general one single configuration (w_1, w_2) covers only a small part of the front, while the exploitation of all objective functions covers a large range of the Pareto optimal front.

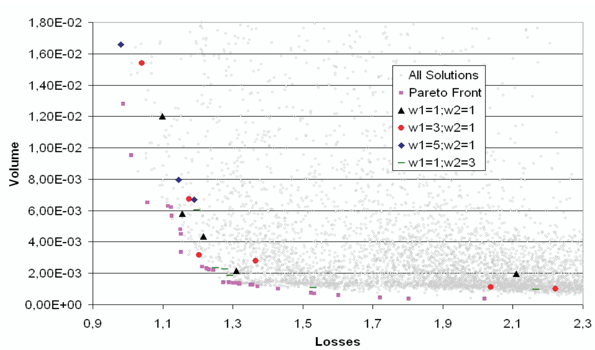

Fig. 21. Front of Non Dominated Solutions.

7 Conclusions

Different versions of Evolution Strategies were applied to solve both single objective and multi objective problems. The objective function was formulated using fuzzy membership functions. Higher Order Niching Evolution Strategies cannot only be used to determine just a single solution but also to find several solutions on or close to the Pareto optimal front.

References

[1] D. Fogel, *Evolutionary Computation*. IEEE Press, 1995.

[2] I. Rechenberg, *Evolutionsstrategie 94*. Stuttgart: Frommann-Holzboog, 1994.

[3] *Computer program package EleFAnT2D*, IGTE TUGraz. [Online]. Available: http://www.igte.tugraz.at

[4] H. Schwefel, *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie*, Birkhaeuser, Basel, 1977.

[5] C. Magele, G. Fürntratt, B. Brandstätter, and K.R.Richter, "Self adaptive fuzzy sets in multi objective optimization using genetic algorithm," *ACES Journal*, vol. 12, no. 2, pp. 26–31, 1997.

[6] O. Aichholzer, F. Aurenhammer, B. Brandstätter, T. Ebner, H. Krasser, C. Magele, and M. Mühlmann, "Evolution strategy and hierarchical clustering," *IEEE Trans. Magn.*, vol. 38, no. 2, pp. 1041 – 1044, 2002.

[7] P. D. Barba, M. Farina, and A. Savini, "An improved technique for enhancing diversity in pareto evolutionary optimization of electromagnetic devices," *COMPEL: Int J for Computation and Maths. in Electrical and Electronic Eng.*, vol. 20, no. 2, 2001.

[8] B. Cranganu-Cretu, J. Smajic, and G. Testin, "Usage of passive industrial frequency magnetic-field shielding for losses mitigation: A simulation approach," in *Proceedings of the Advanced Research Workshop on Transformers (ARWtr 2007)*, University of Vigo, Baiona, Spain, 2007, pp. 325–330.

[9] B. Cranganu-Cretu, J. Smajic, W. Renhart, and C. Magele, "Targeted optimization supported design of industrial devices," *Accepted for Publication in IEEE Transactions on Magnetics*, April 2008.