Application of Parallel Mixed-Integer Evolution Strategies with Mutation Rate Pooling

Martin Schütz* **Joachim Sprave†**

University of Dortmund
Department of Computer Science
Chair of Systems Analysis
44221 Dortmund

Abstract

Two parallel evolution strategies (PESs) for mixed-integer optimization are presented. Based on the well-known *migration-* and *neighboorhood-*model and using a mutation rate pool for strategy parameters, both algorithms are applied to the synthesis of *multilayer optical coatings* (MOCs), especially for the design of an antireflection coating for germanium infrared optics. Even though PESs are no tailored heuristics, i.e., need no optical knowledge or a special starting design, the optimization results confirm theoretical considerations that predict a relationship between the optical thickness and the best achievable mean reflectance of the MOC under consideration. Confirming theoretical results and beating a number of other solutions presented in literature clearly demonstrate the robustness of the presented alogrithms.

1 Introduction

Most so-called *real world problems* in the field of parameter optimization do not fit into the framework of synthetic test functions which are usually used to assess optimization methods, not even in the field of evolutionary algorithms (EAs) where the search for deceptive problems is nearly as old as the methods themselves. There are three properties of problems which seem to be quite often typical for technical optimization and are not covered by traditional EA[1] approaches: constraints, mixed-integer parameter spaces, and variable-dimension, such that the optimization (minimization) problem under consideration can be formalized as follows:

Definition 1 (Minimization problem) *Minimize $f(\vec{x}, \vec{d})$ subject to:*

$$g_i(\vec{x}, \vec{d}) \geq 0 \qquad \forall i \in \{1, \ldots, m\} \tag{1}$$

$$h_j(\vec{x}, \vec{d}) = 0 \qquad \forall j \in \{1, \ldots, l\} \tag{2}$$

$$\begin{aligned} \vec{x} &= (x_1, \ldots, x_{n_x}) \in X \subseteq \mathbb{R}^{n_x}, \\ \vec{d} &= (d_1, \ldots, d_{n_d}) \in Y \subseteq \mathbb{Z}^{n_d} \\ & \quad n_x,\, n_d \in \mathbb{N} \textit{ not fixed}, \\ & \quad f,\, g_i,\, h_j : X \times Y \to \mathbb{R} \end{aligned}$$

□

The restrictions formulated in (1) are called *inequality constraints* whereas those in (2) are named *equality constraints.* A point $(\vec{x}, \vec{d})$ which fulfills (1) and (2) is called *feasible.* The main difference between the fixed- and the variable-dimensional optimization problem is the fact that n_x and n_d may vary during optimization in the non-fixed case.

In contrast to constraints which can be handled with the help of lethal mutations (see [21], for other methods concerning constraint handling the reader is refered to [11]), concepts for solving mixed-integer and/or variable-dimensional problems are nearly unavailable, which leads to the need of designing new adequate algorithms. Bäck and Schütz [3] introduced two more reasons which should motivate further work:

1. It is not sufficient to embed $\mathbb{Z}$ in $\mathbb{R}$ using the standard ES with rounding the object variables.

2. Disconnecting $\mathbb{Z}$ and $\mathbb{R}$ assumes separability between discrete and continuous optimization, which is not given for any problem instance.

A typical example for the problems indicated above is the optimal design of *multilayer optical coatings* (MOCs). The optimization of MOCs by means of an evolution strategy (ES) was discussed in [3]. It was shown that EAs

*schuetz@ls11.informatik.uni-dortmund.de
†sprave@ls11.informatik.uni-dortmund.de
[1] It is presumed that the reader is familiar with concepts and notations of evolutionary algorithms, otherwise [4] is recommended as a starting point for further reading. The reader is referred to [2] for introductionary reading into the field of evolution strategies.

which can be classified as synthesis methods are appropriate methods to evolve such designs, i.e., can compete with refinement methods, but the results could not beat the best solutions found by traditional refinement methods, i.e., local optimization of start designs which were created using expert knowledge and intuition.

One thing that was left out in the first attempt [3] was the handling of constraints. As described later, the so-called *optical thickness* of the resulting filter designs is usually restricted by a lower and an upper bound.

The best quality of a MOC design (in case of non-absorbing MOCs) that can be accomplished depends, roughly spoken, on the maximum optical thickness that is allowed, more precisely ([25], p. 4273):

> "increasing the overall thickness is tantamount to increasing the number of construction parameters available for the solution of the problem. This, in turn, should lead to a lower merit function."

For one special instance of the MOC design, an antireflection coating for germanium infrared optics (for detailed information see [1]), much effort has been performed to exactly determine the relationship between optical thickness and the best achievable average reflection (see [25],[26]), i.e., the quality of the design. As we will see in the excourse to the MOCs basics (section 3), there exists some theory about the predicted shape of this curve.

This paper is an attempt to determine lower bounds experimentally without any assumptions about the characteristics of the resulting curve. To provide that, we needed algorithms that

- are able to handle constraints to search for designs with an arbitrary thickness
- are robust concerning starting points and external control parameters
- are efficiently distributable to make use of parallel computers

Additionally the most important requirement is the capability to find a solution below or at least close to each good design found by other designers in the past, i.e., perform better than standard algorithms.

2 Parallel EAs

Parallel evolutionary algorithms (PEAs) are known to satisfy the requirements given at the end of section 1. Analogous to their non-parallel counterparts, they are probabilistic optimization algorithms imitating the model of natural evolution.

Most parallel EA implementations make use of the natural parallelism given by the population model. Individuals are spread among the processors and the number of individuals per processors normally depends on the computational power of each node and the bandwidth of the communication system. While most of the operators used in EAs are perfectly distributable, selection is not. In most traditional EAs all individuals compete with each other for a place in the population, and every two individuals in a population are potentially mating partners. This is called a *panmictic* population model. Breaking this is essential for a parallel EA with distributed subpopulations.

A very simple approach is the *migration model* [17, 16]. Nearly unmodified EAs are placed on the nodes which occasionally send some individuals to other subpopulations. While subpopulations still remain panmictic, selection can be done independently on each node, even asynchronously, as Rudolph [16] has shown. The migration model defines an interaction graph on subpopulations. The edges of the graph are sometimes called *migration paths*. In addition to a sequential algorithm, there are operators to choose which determine

- which and how many individuals have to leave the population
- how to integrate immigrants

Since the sizes of the subpopulations, the migration frequencies, and the number of migration paths are variable it is possible to balance the need for computational power and communication bandwidth.

Placing single individuals in a graph and defining a neighborhood structure on top of this topology leads to the *neighborhood* [9, 23, 22] or *diffusion model* [18]. The selection operator builds the main difference between both models. Whereas selection in the migration model is a global operator it is locally handled in the diffusion model with respect to the defined neighborhood structure, such that the neighborhood model, which is a better model of nature, seems to perform better on average [18]. Scalability of computational resources can be reached by defining the size of the subgrids on which the individuals are mapped, thus the diffusion model is scalable as the migration model.

3 Synthesis of MOCs

This section should introduce the basics that are necessary to understand the problem covered by the two PESs. According to Furman and Tikhonravov [8] figure 1 presents the basic physical model of multilayer optical coatings which consists of a set of plane parallel isotropic layers separating two homogeneous isotropic media (often these are air and a substrate which can be interpreted as an extremely thick layer). The permittivity and conductivity of the layers depend on one spatial coordinate (here z) perpendicular to the layer-media boundaries. The layers are nonlimited in either directions, i.e., the

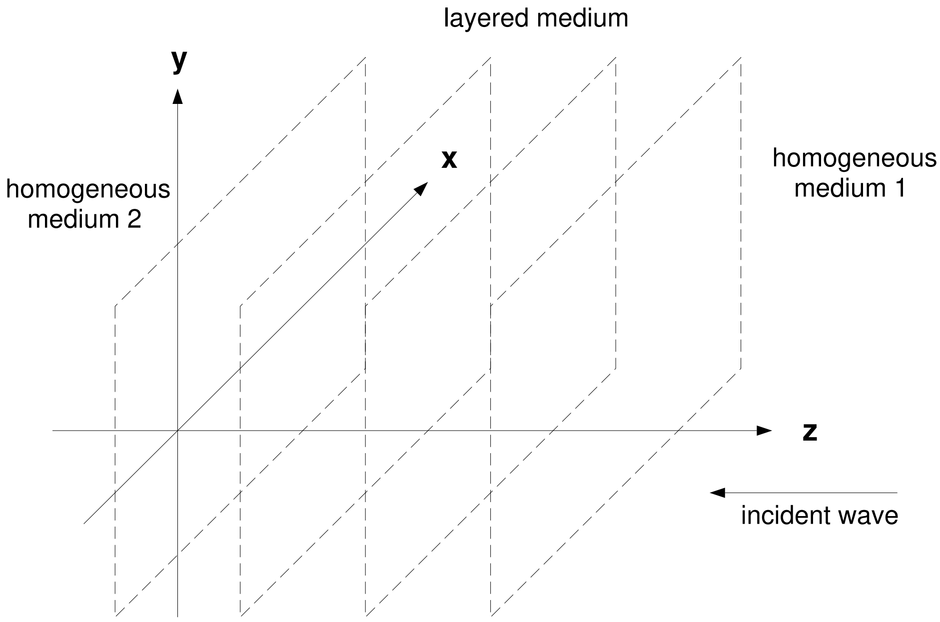

Figure 1: The basic physical model of MOCs.

width of the coating always exceeds the wavelength of the incident wave and the total thickness of the coating by many times. A plane electromagnetic wave enters from the first medium (air) into the multilayer system and is partially reflected or transmitted at the borders between layers of different refractive indices.

Concerning this paper we are interested in the reflection R of a MOC, i.e, our aim is the optimal design of a MOC which transmits 100% of the incident wave in the infrared region of interest. The general problem in the design of MOCs is to find the design parameters (refractive indices and thicknesses[2]) of the system which satisfy the desired spectral (reflectance, transmittance or absorption) properties as good as possible. By making some simplifying assumptions which are common in the theory of MOCs the above optical property R can be fully described by the Maxwell equations which yield the electric vector $\vec{\mathcal{E}}$ and the magnetic vector $\vec{\mathcal{H}}$ of the system under consideration. Distinguishing several special cases a wide range of calculation methods for optical properties, such as the Fresnel formulas, recurrent methods etc., can be identified. For our purposes the widely used *matrix method* for calculating the spectral coefficients, such as transmittance T, reflectance R and absorption A, of the layered media is well suited. Besides its sufficient approximation of the spectral properties one advantage of this method is its structure which can be implemented very easily.

Instead of giving a brief explanation of this method the reader is refered to [8]. In [3] the authors constructed a quick reference of the necessary formulas. In the context of this writing it is sufficient to know that there exists a method which yields to a given MOC and a given wavelength λ the corresponding reflectance value R. A little bit more detailed a MOC using n layers is fully described by a vector of thicknesses $\vec{d} = (d_1, \ldots, d_n)$ and a vector of corresponding refractive indices $\vec{\eta} = (\eta_1, \ldots, \eta_n)$, such that the formula used to calculate the reflectance for a given wavelength is:

$$R(\vec{d}, \vec{\eta}, \lambda) = \frac{4\eta_a\eta_s}{|\eta_a B(\vec{d}, \vec{\eta}, \lambda) + C(\vec{d}, \vec{\eta}, \lambda)|^2} \tag{3}$$

where η_a and η_s describe the refractive index of the outer space and the substrate. B and C are terms depending on $\vec{d}, \vec{\eta}$ and λ encapsulating internal knowledge of the matrix method.

Before turning to the more concrete optimization problem we want to give some remarks about optimization methods in the field of MOCs. Dobrowolski and Kemp ([6], p. 2876) classify the basic approaches to the design of thin multilayer coatings as

- graphical methods
- analytical methods
- digital design methods

[2]Mostly optical thicknesses are searched for. An optical thickness of a layer is the product of the geometrical thickness and the corresponding refractive index.

where the greatest interest lies on digital design techniques, which can be further classified into *refinement* and *synthesis* methods[3].

According to Dobrowolski ([5], p. 50) a refinement which in some cases makes the use of analysis is defined in the following way:

> "If, for a given problem, a suitable starting design can be identified, it is possible to reach a satisfactory solution by gradually modifying its construction parameters. Such a process is called refinement."

From this citation one can directly conclude that the disadvantage of refinement methods is the need for a suitable starting design of the multilayer. Only with such a starting configuration given, the often used refinement methods, which tend to converge to the minimum that is closest to the starting design, are able to find a good solution.

Dobrowolski characterizes the differences of synthesis methods, to which EAs can be classified, over refinement methods as follows ([5], p. 51):

> "In general, when designing a filter with irregular spectral transmittance or reflectance features, it is difficult not only to select a suitable starting design for refinement, but even to guess the overall thickness of the multilayer required for the solution of the problem. Thin film synthesis methods must be designed to overcome these problems."

This means that synthesis methods mostly generate their own starting design. Usually this step is followed by refinement. As in [6] where a good survey of refinement methods is given, Dobrowolski also presents a good overview of synthesis methods in [5]. A short resume of recent evolutionary approaches is given in [3].

Finally, according to [8] (p. 104)

> "it should be stated that so far there is no universal synthesis method capable to provide optimal solutions of all the problems arising in practice."

For most methods difficulties arise when the number of layers of a coating becomes large, but the starting design is unknown. Furman and Tikhonravov [8] present synthesis methods based on merit function optimization as the presently best methods to overcome those difficulties, which gives the motivation of using this approach in the contest formulated by Aguilera et al. [1].

The task, presented in the design contest, was the design of an antireflection coating for germanium (substrate, refractive index $\eta_s = 4.0$) in the $7.7 - 12.3\mu$m region of the infrared (near the visible region). The designers were asked to use only two coating materials, namely germanium (Ge, refractive index $\eta_H = 4.2$) and zinc sulfide (ZnS, refractive index $\eta_L = 2.2$). For the outer space air with a refractive index $\eta_a = 1.0$ was chosen.

To ensure that all participants solve the same problem, they were asked to minimize the "reflection mean square (RMS)" defect function. Using formula (3) the defect function M can be stated as the mean square error of the system reflectance $R(\vec{d}, \vec{\eta}, \lambda_i)$ to the target reflectance $\tilde{R}(\lambda_i) \equiv 0$ at $m = 47$ equidistant choosen wavelengths λ_i in the region of interest:

$$M(\vec{d}, \vec{\eta}) = \sqrt{\frac{1}{m} \sum_{i=1}^{m} R(\vec{d}, \vec{\eta}, \lambda_i)^2} \,. \tag{4}$$

In order to get an impression of the topological characteristics of this merit function, a three-dimensional plot of RMS-values for a two-layer filter with $\eta_1 = 2.2$ (close to air), $\eta_2 = 4.2$ (close to substrate), and d_1, d_2 varying in the range 0–20 μm is shown in figure 2. The landscape is characterized by parallel "waves," separated by valleys of increasing depth and decreasing width. It is obvious that optimization algorithms may stagnate within a suboptimal valley without ever surmounting the next wave and finding a better but narrower valley. Even within a valley, the landscape is characterized by local optima of different quality, located at the valley's bottom, and the algorithm would have to adapt to the valley direction in order to search for a better optimum within the valley.

The reader should keep in mind that the above visualization is just valid for the two-dimensional case, in which the filter problem seems to be an extremely complex, multimodal optimization problem, especially when no problem-specific knowledge — e.g., a starting design — is provided to the algorithm. From figure 2 one cannot conclude that the problem structure remains fixed for the higher-dimensional case (authentic filters with up to 20 layers). So the visualization of the problem not necessarily contradicts the assumption of Tikhonravov and Dobrowolski [25], who classified antireflection coating problems as members of convex objective function classes with linear restrictions. Based on fine properties of two-material coatings (details are omitted here) and the inverse Fourier transform synthesis method both authors created a new synthesis method for antireflection coatings:

1. With help of the *Q-function*[4] and an integral approximation by sums they transformed the highly multimodal antireflection problem of inhomogeneous layers into a quadratic problem with linear constraints.

2. In the next step of the synthesis method the unimodal substitute function is solved by a locally convergent optimization routine.

[3] As indicated by Dobrowolski and Kemp ([6], p. 2877) the borderline between refinement and synthesis methods is becoming blurred, because some of the refinement methods lost their disadvantage of converging to the closest minimum.

[4] The Q-function normally is used in the inverse Fourier transform synthesis method and is a function that describes the spectral properties of inhomogeneous coatings with high accuracy.

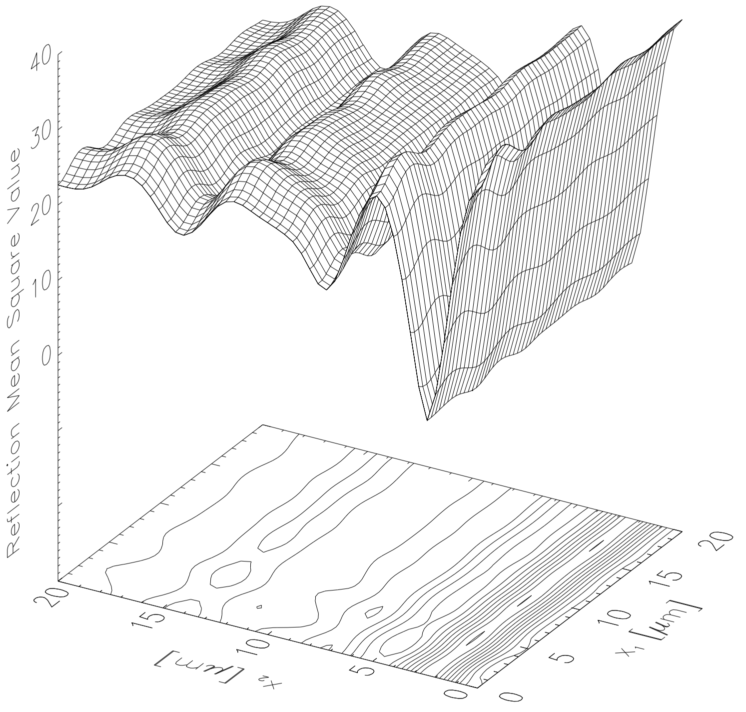

Figure 2: Topology of the RMS merit function in case of a fixed two-layer filter structure with $\eta_1 = 2.2$ (close to air) and $\eta_2 = 4.2$ (close to substrate). Optical thicknesses are varied in the range $0 - 20$ μm.

3. Afterwards they substitute the resulting inhomogeneous layer by the equivalent two-layer system.

4. Finally this system is again refined.

Inspired by the *maximum principle* of the optimality of a coating design (i.e., choosing the largest possible refractive index ratio) Tikhonravov and Dobrowolski [25] believe that "the inhomogeneous layer solutions obtained by the quadratic programming method" are "close to the two-materials solution." They also point out that the maximum principle is responsible for the fact that with a larger number of elementary thicknesses the inhomogeneous solution will even more resemble a homogeneous two-material multilayer.

Tikhonravov et al. [26] used the same idea to suggest that "the design of antireflection coatings belongs to a mathematical class of convex problems" which, as stated in a well-known mathematical theorem, have only one global optimum. Under this assumption the authors created an estimate of the lowest achievable mean reflectance for a normal incidence antireflection in dependence on the overall optical thickness. Figure 3 presents besides the estimation of Tikhonravov et al. an older estimation of Willey [28] which is up to 30μm more optimistic and over 30μm total optical thickness a little bit more pessimistic than the estimation of Tikhonravov et al.

The results from Willey ([27],[28]) were deduced from a few basic principles and some empirical data, whereas those of Tikhonravov et al. [26] are mainly based on theoretical considerations concerning [25], such that the authors of [26] believe to have found a better approximation of the lowest average reflectance.

The results (see section 5) of our runs concerning the two PESs give a strong hint that the real estimate must be a mixture made out of both estimates. This believe is underpinned by the fact that Tikhonravov et al. [26] give no real analytical proof for their results. Instead of

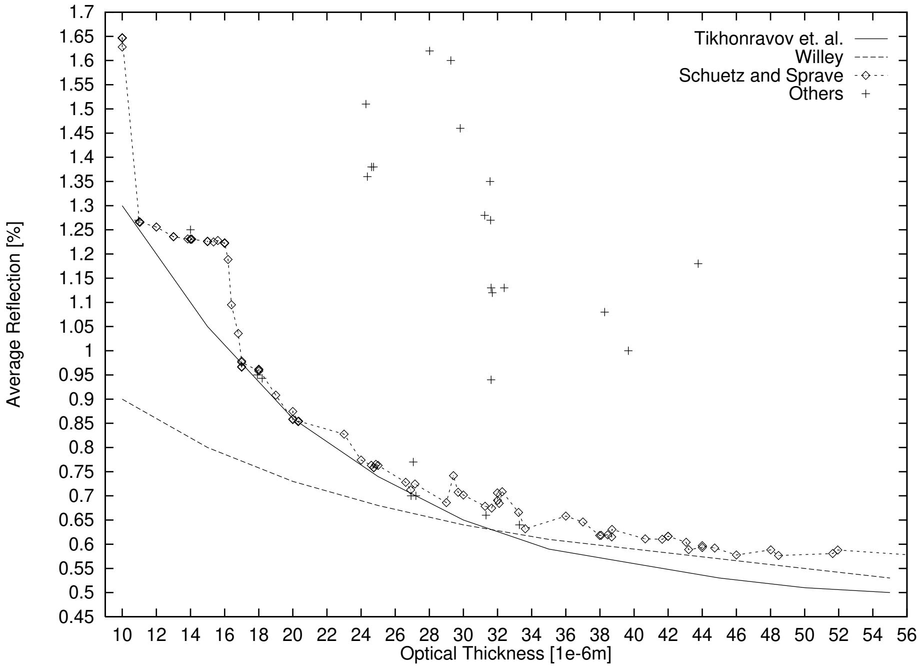

Figure 3: Several approximations of the best achievable mean reflectance to a given optical thickness.

doing so they refer to papers following in future.

4 The Algorithms

4.1 Mixed-Integer Representation

All algorithms used in this work are based on evolution strategies. Thus, the regular data representation is a real-valued vector. To integrate the discrete part of the problem the representation had to be extended. Two different representations have been tested, guided by either one of the following ideas:

- The natural way to add discrete variables to an ES is to use integer vectors and define mutation and recombination analogous to their real-valued counterparts
- Because GAs already deal with discrete variables, a binary representation is used for discrete parameters. Then the common mutation and crossover operators can be applied

Following the first idea, the mutation of discrete variables is guided by the principle of maximum entropy (in the case of MOC design a uniform distribution over the set of indices is the distribution with maximum entropy). The decision when to mutate discrete parameters is driven by an adaptable mutation rate from the interval $(0, 1)$ (for details see [3]). In [20] it was shown that self-learning of the mutation rate can not be easily accomplished.

The advantages and drawbacks of the GA operators used in the second idea are widely discussed. Nevertheless, due the relatively low cardinality of the integer part of the MOCs' representation (typically four or less different refractive indices) it is unlikely that problems like Hamming cliffs take any effect.

4.2 Mutation Rate Pooling

An important property of EAs is the robustness against non-optimum parameter configurations. Adding a fixed mutation rate for integer variables would result in a sensitive external parameter. The algorithms used for this paper therfore use a simple but nevertheless promising technique: only a range of mutation rates has to be specified, and each subpopulation gets its own unique rate from this range. By means of this approach, it is possible to provide an almost optimal mutation rate at each iteration. Moreover, because each good solution undergoes different mutations rates, this approach seems to be a compromise between exploration (high mutation rates) and exploitation (low mutation rates) of the search space.

4.3 Variable Dimensions

In [20], variable dimension has been introduced into ES to cope with the design of optimum MOCs. Since this did not work very successfully, the approach described in [3] has been integrated in the algorithms presented this paper: the calculation of the dimension is hidden from the ES which operates on a fixed-sized real vector and an integer vector of the same length. In this way, an upper bound for the number of layers is given. On the other hand, the objective function merges adjacent layers of the same refractive index, so the resulting filter design can be of any dimension below the maximum.

4.4 Handling of Constraints

One of the major advantages of EAs compared to traditional optimizing methods is that constraints can be handled easily by treating infeasible points as *lethal* mutations. This is the most general approach, because sometimes we simply know just if a point is feasible or not. In cases where we can obtain more information about the degree of infeasibility of points, some information gets lost by simply ignoring these points. To profit from this knowledge, in [10] a generalized penalty function was introduced. If it is possible to measure the distance to the feasible region, a comparison operator can be constructed which is capable of comparing two arbitrary points in the search space. A measure that is available in most cases is the number of violated constraints. Or, if constraints are given to separately restrict the range of each parameter, the Euclidean distance to the feasible region can be computed. In order to find optimal MOC designs under the restriction of a maximum optical thickness, the degree of infeasibility is obviously the difference between the calculated thickness and the given maximum.

4.5 Models of Parallelism

Because serial EAs can be easily extended to migration model EAs, the program used in [3] was ported first to obtain preliminary observations whether the qualilty of the results can be improved. The subpopulations are placed in a ring topology, and, whenever a certain number of generations, the *isolation time* has passed, some individuals are exchanged between contiguous populations.

Motivated by the results from [24], a neighborhood model EA has also been implemented using a logical torus as the population structure. The neighborhood of an individual consists of all individuals in a certain distance in the maximum norm.

Slices of the torus are mapped to processors so that the balance between communication and computation can be biased as well as in the migration model. Moreover, an isolation mechanism has been introduced, resulting in a hybrid approach between the neighborhood model (no isolation, small neighborhood size) and the migration model (with isolation, huge neighborhood size). In contrast to [16], the immigrants remain unchanged in a population until new ones arrive to give them a chance to establish in the new environment.

All local selection operators are taken from [24]. The localized variants of the ES selection schemes, namely the $(\mu + \lambda)$ selection and the (μ, λ) selection, depend on generating a surplus of offspring at each position. The parents of each newly generated individual are randomly selected from the neighborhood. The best child replaces the individual at the given position in case of comma selection. When using the plus selection scheme, an individual is only replaced if one the offspring generated from its neighborhood is better with respect to the fitness function.

The mating selection operator simply selects the two best individuals from the neigborhood and generates one offspring which replaces the individual at the current position in the next generation.

5 Experimental Results

5.1 Selection of the Final PEA

First experiments[5] were made without restricting the optical thickness. It is by no means surprising that the solutions tended to huge optical thicknesses. But at the same time, the merit function values dropped below all values offered in the literature.

While all algorithms behaved qualitatively similar the neighborhood model with mating selection showed by far the best performance. According to [24], this is a strong hint that the MOCs optimization problem is relatively smooth in higher dimensions, perhaps even uni-modal when a certain optical thickness is given.

To determine the effect of mutation rate pooling some experiments were repeated with fixed mutation rates (one run for each rate in the pool) and otherwise unchanged parameters. The pooling variant always found a solution at least as good as the best found with a fixed rate.

The PEA showed a very robust behaviour against variation of population size, neighborhood size and isolation time. We finally chose a subgrid of 20×25 individuals for each of the 15 subpopulations, a neighborhood size of 7×7, and an isolation time of 30 generations.

5.2 Sampling of Designs

While EAs need a relatively large amount of computation time, there is no need for manual calculation of starting designs and careful use of control parameters, which favours this kind of algorithms for the given task, i.e., finding optimal designs of arbitrary optical thicknesses without any need for manual adjustments. About 100 designs in the range from 10μm to 50μm were calculated in completely unsupervised runs.

The first goal was obviously reached: for each design

[5] All experiments were run on a ParsyTec Xplorer system with 16 nodes, each equipped with a Motorola PowerPC-601 processor. The largest runs took about $2 - 3$ days.

which has been published by others who applied refinement methods to promising starting designs, the PEA yielded a solution which is at least very close (e.g. at 10, 14, 27 and 34μm), and the PEA performed much better beyond 34μm.

Figure 3 shows our experimental curve for the best merit values achievable with respect to a given optical thickness, together with the curves of Tikhonravov et al. [26] and Willey [28]. Single points indicate results from Aguilera et al. [1] and other researchers [13, 14, 12, 15, 7, 25, 19].

The most interesting point is the fact that the empirical estimation approximates up to 30μm the curve of Tikhonravov et al., and beyond 30μm the estimate of Willey is better adapted. This leads us to the hypothesis that a mixture, constructed by a combination of parts of both curves, presents the most authentic estimate of the best achievable mean reflectance to a given optical thickness.

Another interesting observation illustrates the "Bermuda Triangle", i.e. an area of great difference between the empirical and the analytical estimate. No result from runs of the other researchers nor from our own give an adequate approximation of the Tikhonravov estimate in the optical thickness range $11 - 17\mu$m.

In order to get an idea how mutation rate pooling works, in figure 4 the merit value of the best individual in each generation has been plotted together with the mutation rate that was used to generate this individual. The range $[0.001, 0.015]$ from which the rates where chosen arose from experiments. Finally, it should be mentioned that in the case of stagnation an increase of the mutation rates was observed, such that the search process got the possibility to leave a local structure optimum. This phenomenon is not visible in graphic 4 because of its scaling which was chosen in order to present the whole fitness curve and not only a part of it.

6 Conclusions

Parallel evolutionary algorithms have been successfully applied to a practical optimization problem which otherwise requires expert knowledge and time to get solutions of high quality. Moreover, in the field of MOCs our results question the two estimated relations of optical thicknesses and achievable filter qualities given by Tikhonravov and Willey, i.e. represent a new insight into the problem structure given by the problem of designing antireflection coatings for germanium infrared optics.

In the range of 11 to 17μm optical thickness, which we called the "Bermuda Triangle", none of the estimated curves could be approached. There are two possible explanations:

1. There is a very attracting non-global local optimum that makes it nearly impossible to find the global one.
2. The assumption that the relationship of optical thicknesses and achievable filter qualities results in a convex curve is wrong.

The latter would mean that it is just a matter of computing power to generate a very accurate estimation by PEAs.

In the field of EAs a fundamental result of our work is the hypothesis that in some complex applications mutation rate pooling seems to work better than learning mutation rates. One reason for this might be the fact that on-line learning of strategy parameters, which can be seen as an dynamic optimization problem, complicates the minimization problem under consideration. In our opinion the choice of the range in which the pool is placed is application dependent, i.e. the range has to be established experimentally. Despite the choice of the mutation rate range an on-line adaptation of this range could be interesting for further research. Since the use of one mutation probability leads to acceptable results one can conclude that there must be a trade-off between learning and pooling.

Local selection appears to be a promising technique to balance EAs' facilities of exploration and exploitation of the search space. Although mating selection emphasizes fast progress to local optima, it performs best on the optimization of MOCs which perhaps means that the difficulty of this problem is not multi-modality, but more likely narrow high-dimensional valleys.

In this paper, we mostly described what worked, but not why and how. These are of course open questions and subject of further work. Which conditions must be given to make self-adaption work? How can one decide which selection type is appropriate for a given class of problems? Both empirical and theoretical work are neccessary to get closer to the answers.

Acknowledgments

The authors gratefully acknowledge support by the BMBF, grant 01 JR 509 A (Joachim Sprave) and the VW foundation, grant I/69 254 (Martin Schütz).

References

[1] J. A. Aguilera, J. Aguilera, P. Baumeister, A. Bloom, D. Coursen, J. A. Dobrowolski, F. T. Goldstein, D. E. Gustafson, and R. A. Kemp. Antireflection coatings for germanium IR optics: A comparison of numerical design solutions. *Applied Optics*, 27(14):2832–2840, July 1988.

[2] Th. Bäck and F. Hoffmeister. Basic aspects of evolution strategies. *Statistics and Computing*, 4(2):51–63, 1994.

[3] Th. Bäck and M. Schütz. Evolution strategies for mixed-integer optimization of optical multilayer sys-

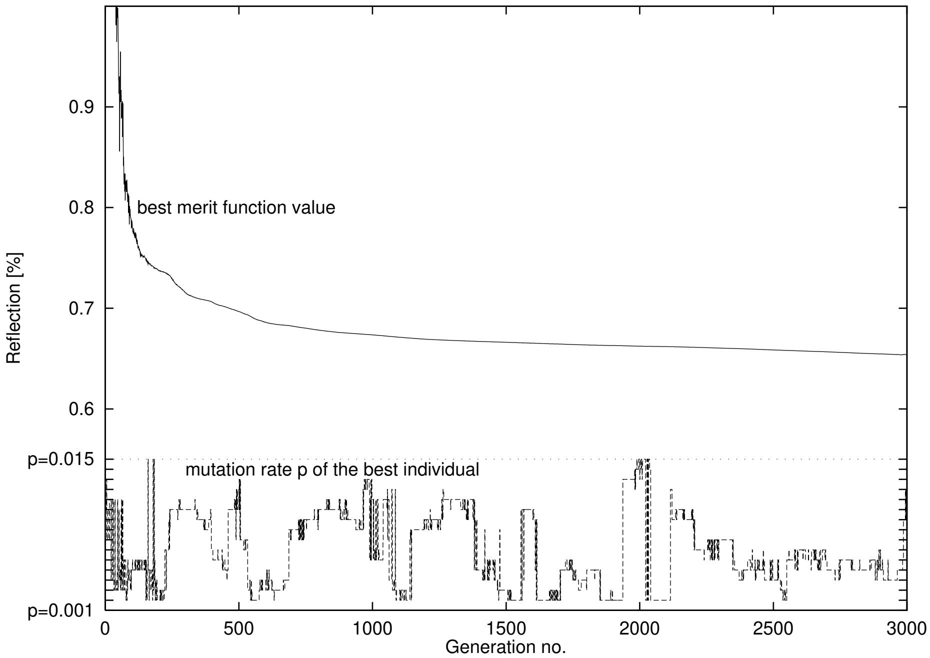

Figure 4: Example for the effect of pooled mutation rates

tems. In J. R. McDonnell, R. G. Reynolds, and D. B. Fogel, editors, *Evolutionary Programming IV: Proceedings of the Fourth Annual Conference on Evolutionary Programming*, pages 33–51. MIT Press, Cambridge, MA, 1995.

[4] Th. Bäck and H.-P. Schwefel. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1–23, 1993.

[5] J. A. Dobrowolski. Design of optical multilayer coatings at NRCC. *SPIE*, 652,(Thin Film Technologies II):48–56, 1986.

[6] J. A. Dobrowolski and R. A. Kemp. Refinement of optical multilayer systems with different optimization procedures. *Applied Optics*, 29(19):2876–2893, 1990.

[7] J. A. Dobrowolski and R. A. Kemp. Interface design method for two-material optical multilayer coatings. *Applied Optics*, 31:6747–6756, 1992.

[8] Sh. A. Furman and A. V. Tikhonravov. *Basics of Optics of Multilayer Systems.* Basics of. Editions Frontières, Gif-sur-Yvette Cedex – France, 1992.

[9] M. Gorges-Schleuter. *Genetic Algorithms and Population Structures — A Massively Parallel Algorithm.* Dissertation, Universität Dortmund, 1990.

[10] F. Hoffmeister and J. Sprave. Problem-Independent Handling of Constraints by Use of Metric Penalty Functions. Submitted to EP'96, 1996.

[11] Z. Michalewicz. Genetic algorithms, nonlinear optimization, and constraints. In L. Eshelman, editor, *Genetic Algorithms: Proceedings of the 6th International Conference*, pages 151–158. Morgan Kaufmann Publishers, San Francisco, CA, 1995.

[12] J. A. Dobrowolski P. G. Verly and R. R. Willey. Fourier-transform method for the design of wideband antireflection coatings. *Applied Optics*, 31:3836–3846, 1992.

[13] A. Premoli and M. L. Rastello. Stochastic synthesis of multilayers. *SPIE - Optical Thin Films and Applications*, 1270:18–27, 1990.

[14] A. Premoli and M. L. Rastello. Minimax refining of optical multilayer systems. *Applied Optics*, 32:1597–1605, 1992.

[15] M. L. Rastello and A. Premoli. Continuation method for synthesising antireflection coatings. *Applied Optics*, 31:6741–6764, 1992.

[16] G. Rudolph. Globale Optimierung mit parallelen Evolutionsstrategien. Diplomarbeit, Universität Dortmund, Fachbereich Informatik, 1990.

[17] G. Rudolph. Global optimization by means of distributed evolution strategies. In H.-P. Schwefel and

R. Männer, editors, *Parallel Problem Solving from Nature — Proceedings 1st Workshop PPSN I*, volume 496 of *Lecture Notes in Computer Science*, pages 209–213. Springer, Berlin, 1991.

[18] G. Rudolph. Parallel approaches to stochastic global optimization. In W. Joosen and E. Milgrom, editors, *Parallel Computing: From Theory to Sound Practice, Proceedings of the European Workshop on Parallel Computing*, pages 256–267. IOS Press, Amsterdam, 1992.

[19] N. K. Sahoo and R. V. S. R. Apparao. Modified complex method for constrained design and optimization of optical multilayer thin-film devices. *Applied Physics A*, A 59:317–326, 1994.

[20] M. Schütz. Eine Evolutionsstrategie für gemischt-ganzzahlige Optimierungsprobleme mit variabler Dimension. Diploma thesis, University of Dortmund, September 1994.

[21] H.-P. Schwefel. *Evolution and Optimum Seeking.* Sixth-Generation Computer Technology Series. Wiley, New York, 1995.

[22] P. Spiessens and B. Manderick. A genetic algorithm for massively parallel computers. In R. Eckmiller, G. Hartmann, and G. Hauske, editors, *Parallel Processing in Neural Systems and Computers*, pages 31–36. North–Holland, 1990.

[23] J. Sprave. *Parallelisierung Genetischer Algorithmen zur Suche und Optimierung.* Diploma thesis, University of Dortmund, December 1990.

[24] J. Sprave. Linear neighborhood evolution strategy. In A. V. Sebald and L. J. Fogel, editors, *Proceedings of the 3nd Annual Conference on Evolutionary Programming*, pages 42–51. World Scientific, Singapore, 1994.

[25] A. V. Thikonravov and J. A. Dobrowolski. Quasi-optimal synthesis for antireflection coatings: A new method. *Applied Optics*, 32(22):4265–4275, 1993.

[26] A. V. Tikhonravov, M. K. Trubetskov, J. A. Dobrowolski, and B. T. Sullivan. Optimum Solutions to Single-Band Normal Incidence Antireflection Coating Problems. In *Optical Interference Coatings, Vol. 17, OSA Technical Digest Series*, pages 49–51, Washington DC, 1995. Optical Society of America.

[27] R. R. Willey. Broadband Antireflection Coating Design Performance Estimation. In *34th Annual Technical Conference Proceedings*, pages 205–209. Society of Vacuum Coaters, 1991.

[28] R. R. Willey. Predicting achievable design performance of broadband antireflection coatings. *Applied Optics*, 32:5447–5451, 1993.