Evolution Strategies for the Optimization of Microdevices

Sibylle D. Müller, Ivo F. Sbalzarini, Jens H. Walther, Petros D. Koumoutsakos
Institute of Computational Sciences
Swiss Federal Institute of Technology
CH - 8092 Zürich, Switzerland
muellers,ivos,walther,petros@inf.ethz.ch

Abstract- Single- and multicriteria evolution strategies are implemented to optimize micro-fluidic devices, namely the shape of a microchannel used for bioanalysis and the mixing rate in a micromixer used for medical applications. First, multimembered evolution strategies employing mutative step size adaptation are combined with the Strength Pareto Approach. In order to support targetting, an extension of the Strength Pareto Evolutionary Algorithm is proposed. Applied on the optimization of the microchannel, these algorithms suggest a novel design with improved properties over traditional designs. A comparison with a gradient method is presented. Second, an evolution strategy with derandomized self-adaptation of the mutation distribution is used to optimize the micromixer. The results agree well with dynamical systems theory.

1 Introduction

We apply evolution strategies to optimization problems for microdevices used in medical applications. In particular, the shape of a microchannel and the mixing rate in a micromixer are optimized. Optimizing microdevices presents a challenge due to the lack of knowledge to guide the design and due to the difficulty to formulate gradient algorithms for the governing equations. In conjunction with efficient computational methods for the simulation of such devices, evolutionary algorithms are highly suitable methods for this application due to their portability and inherent parallelization making possible the reduction of time to market in such designs.

Besides classical single objective evolution strategies, we also present multiobjective algorithms that find a set of optimal trade-off fronts, the so-called Pareto-optimal set, for problems with multiple, conflicting goals. An existing powerful multicriteria algorithm is embedded in an evolution strategy and extended for the purpose of targetting.

Section 2 of this paper introduces the evolutionary algorithms implemented in this work. Section 3 shows their application on the design of a microchannel while the application on a micromixer is presented in Section 4. Conclusions are drawn in Section 5.

2 Evolutionary Optimization Methods used in this Work

We use the following evolutionary optimization strategies:

- The (1+1) evolution strategy
- The Strength Pareto Approach for multiobjective optimization [Zitzler & Thiele (1999)] combined with an evolution strategy using mutative step size adaptation (see subsection 2.1)
- The Strength Pareto Approach with Targetting (see subsection 2.2)
- The derandomized $(\mu/\mu_I, \lambda)$ evolution strategy with covariance matrix adaptation (CMA-ES) with intermediate recombination [Hansen & Ostermeier (1996)], [Hansen & Ostermeier (1997)]

2.1 The Strength Pareto Approach combined with an Evolution Strategy

For this work, the *Strength Pareto Approach* for multiobjective optimization has been used because comparative studies have shown that, among all major multiobjective EAs, the *Strength Pareto Evolutionary Algorithm (SPEA)* is clearly superior ([Zitzler & Thiele (1999), Zitzler, Deb & Thiele (2000)]). It is based on Pareto-optimality and dominance. The algorithm as proposed by [Zitzler & Thiele (1999)] was implemented in a restartable, fully parallel code.

Following the notation in [Zitzler & Thiele (1999)], step 7 requires an algorithm for the selection of individuals into the mating pool and Step 8 includes some method for dynamical adaptation of step sizes (i.e. mutation variances). For this paper, selection was done using the following binary tournament procedure:

1. Select at random two individuals out of the population P.
2. Copy the one with the better fitness value to the mating pool.
3. If the mating pool is full, then stop, else go to Step 1

Adaptation of the step sizes was realized using the self-adaptive mutative technique. Each element of the population P and of the non-dominated individuals in the external population P' is assigned an individual step size for each dimension. The step sizes of all members of the mating pool are then either increased by 50%, cut to half, or kept the same, each at a probability of 1/3.

2.2 Strength Pareto Approach with Targetting

Compared to other methods like for example the Energy Minimization Evolutionary Algorithm (EMEA) (cf. [Jonathan, Zebulum, Pacheco & Vellasco (2000)]), the SPEA has two major advantages: it finds the whole Pareto-front and not just a single point on it and it converges faster. The latter is a universal advantage whereas the former is not. There are applications where a target value can be specified. In this case, one wants to find the point on the Pareto-front which is closest to the user-specified target (in objective space). This eliminates the need to analyse all the points found by SPEA in order to make a decision. EMEA offers such a possibility but it converges slower than SPEA. Moreover, EMEA is not able to find more than one point per run. Thus, we extend SPEA with a targetting facility that can be switched on and off depending on whether one is looking for a single solution or for the whole front, respectively. We added this capability to SPEA by the following changes to the algorithm:

1. Between Step 6 and Step 7 the fitnesses of all individuals in P and P' are scaled by the distance D of the individual from the target (in objective space) to some power q:

$$f_i = f_i \cdot D_i^q$$

   This ensures that enough non-dominated members close to the target will be found, so that the one with minimal distance will appear at higher probability. The parameter q determines the sharpness of the concentration around the target.

2. Another external storage P_{best} is added that always contains the individual out of P' which is closest to the target. Therefore, between steps 4 and 5, the algorithm calculates the distances of all members of P' to the target and picks the one with minimal distance into P_{best}. At all times, P_{best} contains only one solution.

3. At the end of the algorithm, not only the Pareto-front is put out but also the solution stored in P_{best}. Note that due to clustering and removal in P', the solution in P_{best} is not necessarily contained in P'. Therefore, it is an optimal solution which otherwise would not have appeared in the output.

The algorithm has been implemented and tested for convex and nonconvex testfunctions. Figures 1 to 4 show some results for the nonconvex testfunction $\mathcal{T}_2$ as proposed in [Zitzler, Deb & Thiele (2000)]:

$$\begin{array}{ll} \text{Minimize} & \mathcal{T}_2(\mathbf{x}) = (f_1(x_1), f_2(\mathbf{x})) \\ \text{subject to} & f_2(\mathbf{x}) = g(x_2, \cdot \ldots, x_m) h(f_1(x_1), g(x_2, \ldots, x_m)) \\ \text{where} & \mathbf{x} = (x_1, \ldots, x_m) \\ & f_1(x_1) = x_1 \\ & g(x_2, \ldots, x_m) = 1 + 9 \cdot \sum_{i=2}^{m} x_i/(m-1) \\ & h(f_1, g) = 1 - (f_1/g)^2 \end{array} \quad (1)$$

where m is the dimension of the parameter space and $x_i \in [0,1]$. The exact Pareto-optimal front is given by $g(\mathbf{x}) = 1$. The parameters of the algorithm were set as summarized in Table 1.

Parameter	Value
Dimension of parameter space (m)	5
Size of population (λ)	50
Size of mating pool (μ)	30
Size of non-dominated set (N')	70
Number of generations	250
Target value for (f_1, f_2)	(0.5, 0.7)
Concentration parameter q	4

Table 1: Settings for targetting SPEA

The chosen target value is slightly off-front. Therefore, the targetting error will not become zero. Figure 1 shows the final population after 250 generations without targetting. The diamonds indicate members of the external non-dominated set (Pareto-optimal front) whereas members of the regular population are denoted by crosses. In Figure 2 the same run has been repeated with targetting. Figure 3 shows the targetting error as a function of the generation number. The dashed line indicates the theoretical minimum of the distance. After about 80 to 100 generations, the point on the front which is closest to the target has been found with good accuracy. Figure 4 shows the path of P_{best} towards the target. The jumps are due to the fact that the individual stored in P_{best} gets replaced as soon as another individual is closer to the target.

The actually achieved best objective value was $\mathbf{f}(P_{best}) = (0.5265, 0.7247)$; its Euclidean distance from the target is $3.6287 \cdot 10^{-2}$, which is equal to the theoretical minimal distance within the given computational accuracy.

3 Microchannel Flow Optimization

Both single- and multiobjective EAs have been applied to a fluidic microchannel design problem. Bio-analytical applications require long thin channels for DNA sequencing by means of electrophoresis. In order to pack a channel of several meters in length onto a small square plate, curved geometries are required. However, curved channels introduce dispersion and therefore limit the separation efficiency of the system. The question is now how to shape the contour of the channel in order to minimize dispersion.

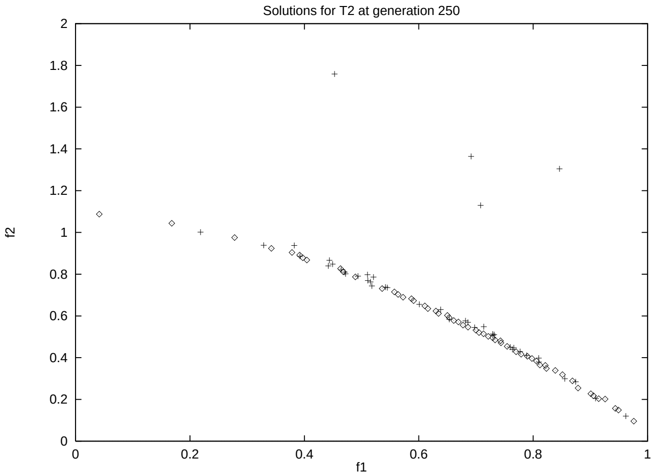

Figure 1: Final population without targetting

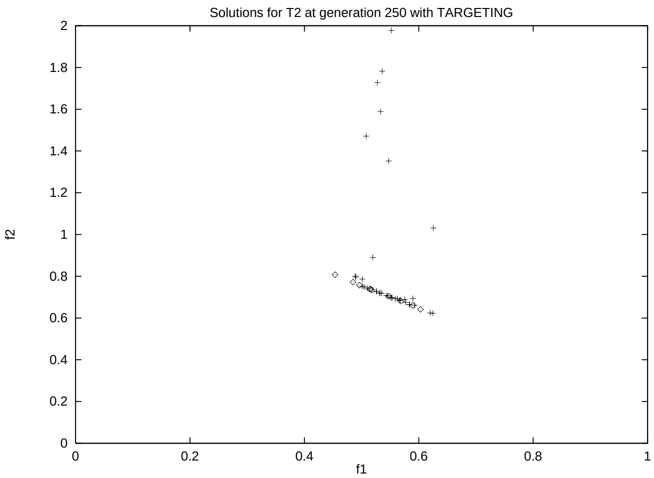

Figure 2: Final population with targetting

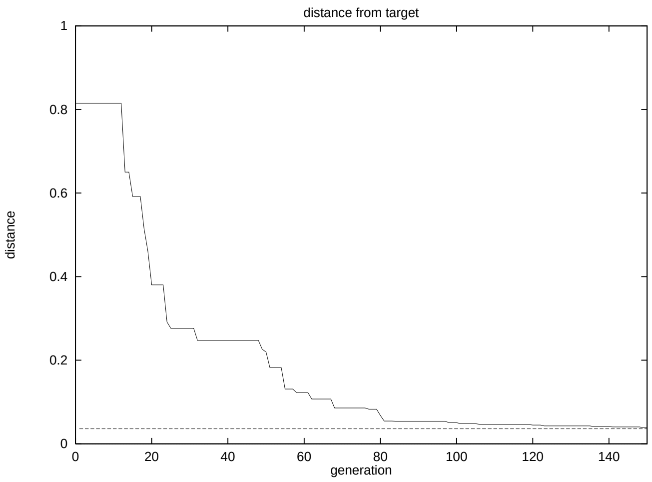

Figure 3: Distance between P_{best} and target

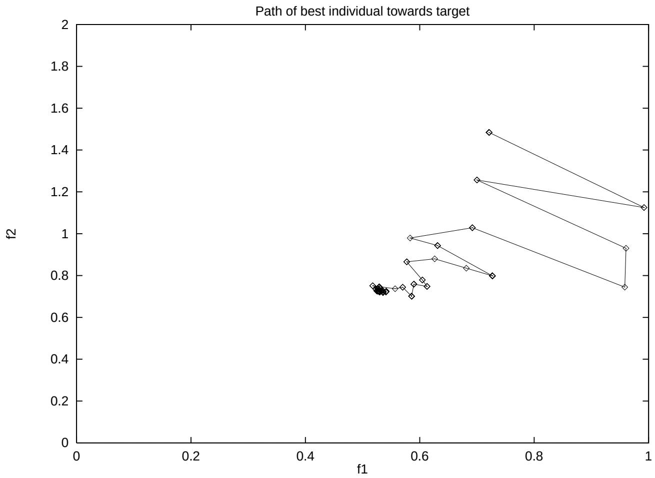

Figure 4: Path of P_{best} towards the target

A detailed description of the problem as well as an optimization solution using gradient methods can be found in [Mohammadi, Molho & Santiago (2000)].

3.1 Single Objective Optimization

The goal of this optimization run was to minimize the final skewness of the flow inside the channel, i. e. it was required, that the iso-values of the advected species a be normal to the flow field U by time T when they exit the channel. The objective function defined by [Mohammadi, Molho & Santiago (2000)] is therefore:

$$J = \int_{\Omega} \left(\nabla a(x,T) \times U(x)\right)^2 \; dx \tag{2}$$

with Ω being the cross section of the channel exit. The shape of the 90 degrees turn is described by 11 parameters. Therefore, the parameter search space is of dimension 11. The objective space is scalar since it is a single objective problem.

The objective function is computed by solving the governing equations of the motion of the species in an electric fields. Both a (1+1)-ES and a $(3/3_I, 12)$-CMA-ES were applied to the problem and their convergence was compared. The results were statistically averaged from 5 runs with different initial conditions, i.e. starting points.

Since the CMA-ES has a population size of 12, it performs 12 function evaluations per generation. Figure 5 shows the convergence normalized to the same number of function calls. Figures 6 and 7 show the corresponding solutions after 20 and 180 generations of the best 1+1 run out of the ensemble (the lines are iso-potential lines of the electric field). After 20 generations the contour of the channel gets a clearly visible dent in it. After 80 evaluations of the objective function, the algorithm has found a double-bump shape to be even better and after 180 calls to the solver, no further significant improvement is observed. The value of the objective function has dropped to about 10^{-6} for the best run out of the ensemble.

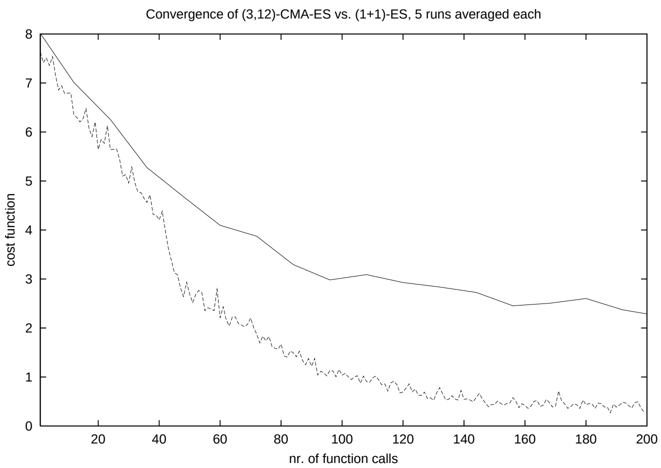

Figure 5: Convergence of $(3/3_I, 12)$-CMA-ES [solid line] and (1+1)-ES [dashed line] vs. number of evaluations of the objective function.

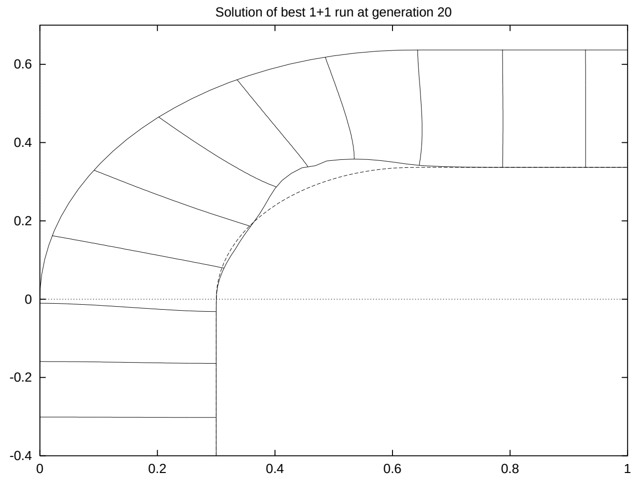

Figure 6: Solution at generation 20 using ES

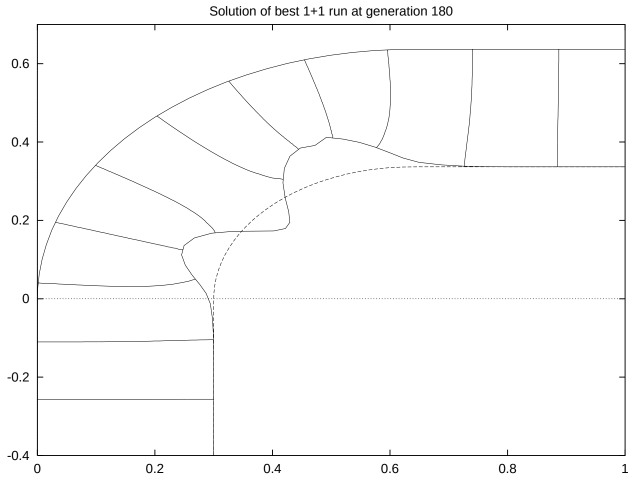

Figure 7: Solution at generation 180 using ES

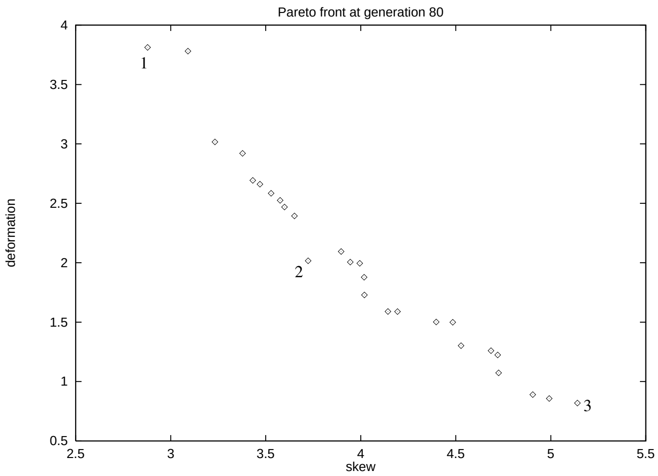

Figure 8: Pareto-front of non-dominated solutions after 80 generations using ES

This means that dispersion is almost zero and the channel has good separation properties.

3.2 Multiobjective Optimization

Afterwards, the total deformation of the channel contour is introduced as a second objective to be minimized simultaneously in order to minimize manufacturing costs. Thus, the second objective reads:

$$K = \sum_i p_i^2 \quad i = 1, \ldots, 11 \tag{3}$$

where p_i are the shape parameters of the channel as introduced by [Mohammadi, Molho & Santiago (2000)]. The first objective remains unchanged. The algorithm used for this optimization is a SPEA with a population size of 20, a maximum size of the external non-dominated set of 30, and a mating pool of size 10.

Figure 8 shows the Pareto-optimal trade-off front after 80 generations of the algorithm and figures 9 and 10 show the corresponding solutions, i.e. optimized shapes of the channel. One is now free to choose whether to go for minimal skewness at the expense of a higher deformation (cf. Figure 9), choose some intermediate result or minimize deformation in order to minimize manufacturing costs and still get the lowest skewness possible with the given amount of deformation (cf. Figure 10).

3.3 Comparison with Gradient based Methods

Figures 11 and 12 show two classes of optimized shapes obtained by [Mohammadi, Molho & Santiago (2000)] using gradient methods. It is interesting that the gradient technique offers two different designs, namely the single-dented (Fig. 11) and the double-dented shapes (Fig. 12) which we found with the evolution strategy also. Therefore, we obtain qualitatively similar results from both methods. Using the gradient method, the skew is reduced by one order of magni-

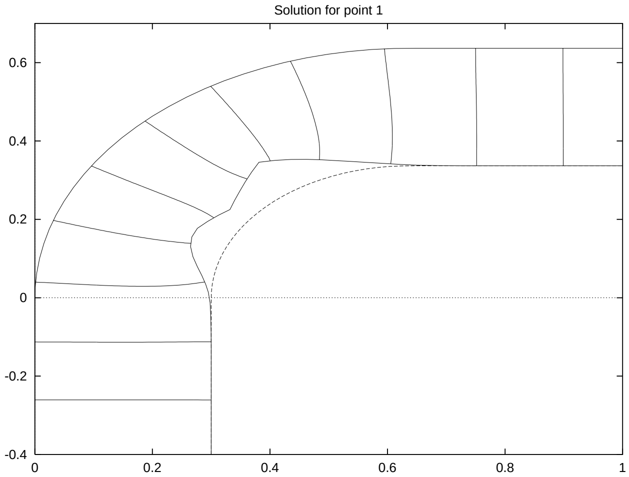

Figure 9: Solution at point 1 using ES

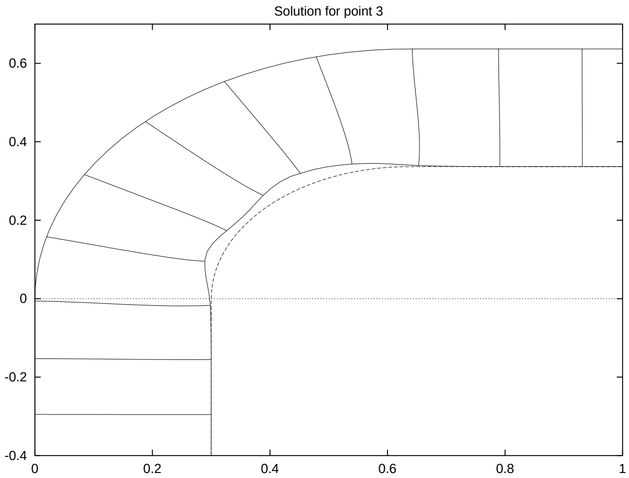

Figure 10: Solution at point 3 using ES

tude [Mohammadi, Molho & Santiago (2000)] which is comparable to the numbers obtained by evolutionary optimization. While trial and error procedures were used in the gradient methods to obtain various solutions, ES provides us with a number of solutions and a Pareto front in a fully automated fashion. Unlike the gradient based methods which require an explicit formulation of the optimization problem in hand, the evolution strategy provided a straightforward optimization procedure. Moreover, the small cost of the computations implies that ES are a reliable method leading to greater flexibility and shorter “time-to-solution”.

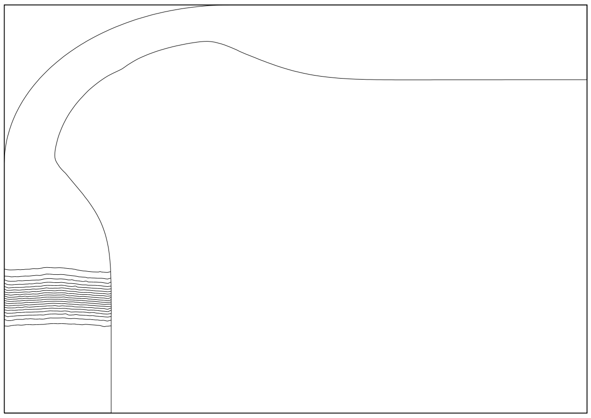

Figure 11: First optimized shape using a gradient method [Mohammadi, Molho & Santiago (2000)]

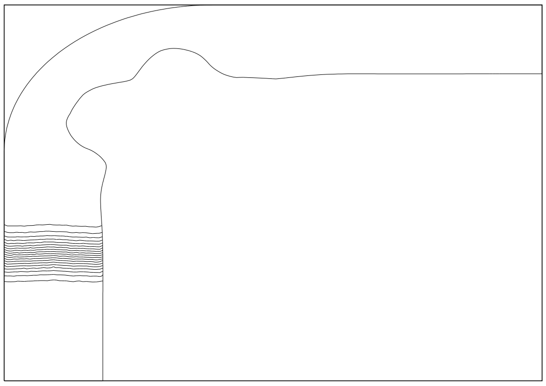

Figure 12: Second optimized shape using a gradient method [Mohammadi, Molho & Santiago (2000)]

4 Micromixer Optimization

An evolution strategy with covariance matrix adaptation (CMA-ES) is applied to optimize a micromixer. This strategy has been found to preserve invariance properties against transformations of optimization parameters and to perform better than other evolutionary algorithms for non-separable and badly scaled functions [Hansen & Ostermeier (1997)].

Moreover, since we are dealing with a highly dynamical system that introduces noise to the objective function, we take advantage of the recombination feature in the proposed method.

The proposed mixer is actively controlled to enhance mixing in a straight channel. Flow in the main channel is manipulated by controlling time-dependent flow from six secondary channels. From these secondary channels, time-dependent cross-flow momentum is imparted on the main channel flow which alters the trajectories of flow-tracing particles. A micrograph of the mixing chip is shown in Figure 13 and the flow configuration is illustrated in Figure 14.

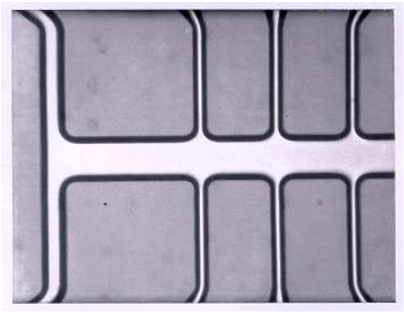

Figure 13: Micrograph of the mixing chip

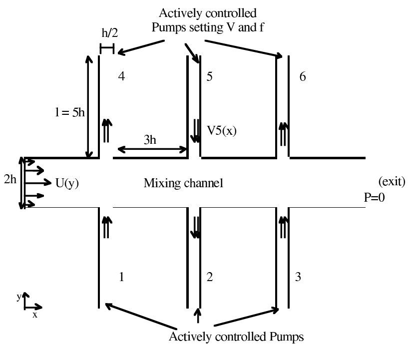

Figure 14: Schematic of the flow configuration

As seen from Figure 14, the main channel is $2h$ in height and $13.5h$ in length where h is a varying length scale. With a distance of $3h$ between secondary channels, they are $h/2$ in width and $5h$ in length. The inlet velocity $U(y)$ of the main channel

$$U(y, x=0) = U_m \left(1 - \left(\frac{y}{h}\right)^2\right), |y| <= h \tag{4}$$

is parabolic, and the inlet velocity of each secondary channel set i

$$V_i = \hat{V}_i \sin(2\pi f_i t + \phi_i), i = 1(1)3 \tag{5}$$

is sinusoidal in time where f_i is the oscillation frequency, ϕ_i the phase shift relative to the first set of secondary channels, and $\hat{V}_i$ the velocity amplitude.

In order to study the performance of the mixer, we compute the mixing rate of the flow by numerical simulations of the governing Navier-Stokes equations at a Reynolds number $Re = 5$.

The aim of the optimization is to obtain the parameter vector which leads to the most pronounced mixing rate in the micromixer. Actuation parameters are the frequency, the amplitude, and the phase shift for each pair of the secondary channels, yielding a total number of 9 actuation parameters. Within this work, we set amplitudes and phases to constant values, namely $\hat{V}_1 = \hat{V}_2 = \hat{V}_3 = 2U_m = 2$, and $\phi_1 = \phi_3 = 0$, $\phi_2 = \pi$ as recommended in [Volpert, Mezic, Meinhart & Dahleh (2000)]. As optimization parameters remain the three frequencies f_1, f_2, f_3 which vary within the limits $[0, 1]$, summarized in the parameter vector $\mathbf{x} = (f_1, f_2, f_3)$.

The objective function to be minimized is the mixing rate m which is computed by

$$m = \frac{\sum_{i=1}^{N}(c(i))^2}{N} - \left(\frac{\sum_{i=1}^{N}(c(i))}{N}\right)^2 \tag{6}$$

where c is the scalar concentration and N the number of vertices in the outlet for which the concentration is measured.

For the optimization, we evaluate the mixing rate until time $t = 135$ in steps of 0.05. After trying different sets of frequencies, we decided on averaging the mixing rate between time $t = 90$ and time $t = 135$, a time regime in which the flow reaches steady state for the tested frequencies. Time $t = 135$ corresponds to 10 flow through times. The CPU time for one simulation of the flow until time $t = 135$ takes about 3 CPU hours on a Sun Sparc Ultra-2 processor.

We implement a derandomized evolution strategy with covariance matrix adaptation and with recombination of all parents. The population consists of $\mu = 2$ parents and $\lambda = 10$ children. Using MPI, the optimization is run in parallel on 5 processors of a Sun workstation cluster.

The results of the optimization of the three frequencies are documented in Table 2.

Number of actuated frequencies		3
Initial frequencies	f_1	0.25
	f_2	0.3333
	f_3	0.5
Initial mixing rate	m	0.0345
Best frequencies	f_1	0.1388
	f_2	0.3165
	f_3	0.4956
Best mixing rate	m	0.0213
Number of function evaluations		460

Table 2: Initial and optimized frequencies

To obtain the required accuracy of the parameters, the function had to be evaluated 460 times. Due to the huge computational cost of the optimization, we could not afford to run another optimization with a different direct search technique and compare it with the ES. However, we compare the optimization result with dynamical system theory. The optimum frequency for a mixer with only one side channel can be determined analytically by considering the movement of a fluid particle in the main channel which yields a value of $f = 1/2$ (non-dimensional units). From studying a mixer with three secondary channels, one can learn that there is a difference if frequency ratios are 1, rational, or irrational. The computation of the mixing rate for the parameter vector $\mathbf{x} = (1/2, 1/2, 1/2)$ with identical frequencies yields $m = 0.1596$ while our initial ”rational” frequency set $\mathbf{x} = (1/2, 1/3, 1/4)$ gives us $m = 0.0345$. ”Irrational” frequencies as proposed in [Volpert, Mezic, Meinhart & Dahleh (2000)] $\mathbf{x} = (1/(2\sqrt{5}), 1/(2\sqrt{2}), 1/2)$ yield $m = 0.0285$ which is the best mixing found theoretically. As one can see, the evolutionary optimization yields a much better mixing rate of $m = 0.0213$ with the frequencies reported in Table 2. This result means a considerable improvement compared with numbers from a well-established theory.

Figures 15 and 16 show a snapshot of the flow in the micromixer at time $t = 45$ for the identical and optimal frequencies, respectively.

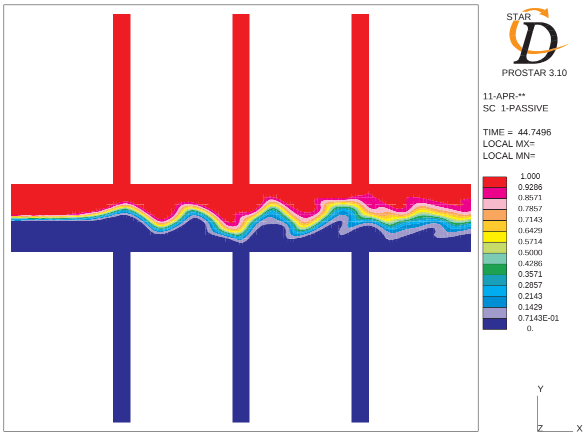

Figure 15: Flow actuated by $\mathbf{x} = (1/2, 1/2, 1/2)$

5 Conclusions

Single- and multiobjective evolutionary algorithms have been implemented and assessed. The SPEA has successfully been extended to support targetting in objective space, an important feature when faced with a problem where some problem-specific knowledge is available. It has been shown that these algorithms are easy to apply to microdevice related problems. For the microchannel, the solutions are comparable to those found by gradient based methods while remaining portable and providing a multitude of efficient new designs in an automated fashion. For the micromixer, the evolution strategy suggests a solution even better than theoretical results.

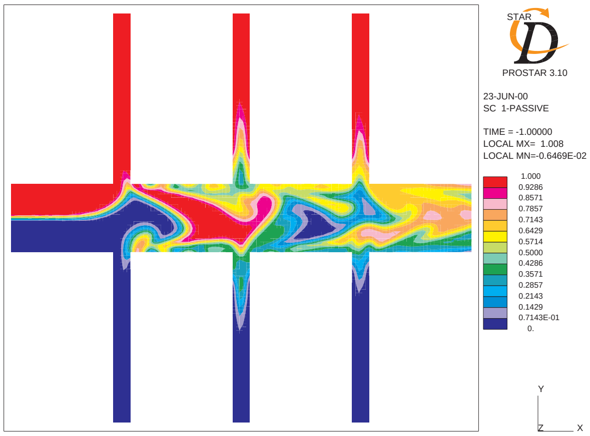

Figure 16: Flow actuated by optimal frequencies

Bibliography

[Zitzler & Thiele (1999)] Zitzler, E. & Thiele, L. Nov. 1999 Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach *IEEE Transactions on Evolutionary Computation*, **3(4)**, 257-271.

[Zitzler, Deb & Thiele (2000)] Zitzler, E., Deb, K. & Thiele L. 2000 Comparison of Multiobjective Evolutionary Algorithms: Empirical Results *Evolutionary Computation*, **8(2)**, 173-195.

[Schaffer (1984)] Schaffer, J. D. 1984 Multiple Objective Optimization with Vector Evaluated Genetic Algorithms, Unpublished Ph.D. thesis, Vanderbilt University, Nashville, Tennessee.

[Schaffer (1985)] Schaffer, J. D. 1985 Multiple objective optimization with vector evaluated genetic Algorithms, *Proceedings of an International Conference on Genetic Algorithms and Their Applications*, sponsored by Texas Instruments and the U.S. Navy Center for Applied Research in Artificial Intelligence (NCARAI), 93-100.

[Fonseca & Fleming (1995)] Fonseca, C. M. & Fleming, P. J. 1995 Am overview of evolutionary algorithms in multiobjective optimization, *Evolutionary Computation*, **3(1)**, 1-16.

[Horn (1997)] Horn J. 1997 Multicriteria decision making, In Bäck, T., Fogel, D. B. and Michalewicz, Z., editors, *Handbook of Evolutionary Computation*, **97/1**, F1.9, IOP Publishing Ltd. and Oxford University Press.

[Van Veldhuizen & Lamont (1998)] VAN VELDHUIZEN, D. A. & LAMONT, G. B. 1998 Multiobjective evolutionary algorithm research: A history and analysis, Technical Report TR-98-03, Department of Electrical and Computer Engineering, Graduate School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, Ohio.

[Coello (1999)] COELLO, C. A. C. 1999 A comprehensive survey of evolutionary-based multiobjective optimization, *Knowledge and Information Systems*, **1(3)**, 269-308.

[Morse (1980)] MORSE, J. N. 1980 Reducing the size of the non-dominated set: Pruning by clustering, *Computers and Operations Research*, **7(1-2)**, 55-66.

[Jonathan, Zebulum, Pacheco & Vellasco (2000)] JONATHAN, M., ZEBULUM, R. S., PACHECO, M. A. C., VELLASCO, M. B. R. 2000 Multiobjective Optimization Techniques: A Study Of The Energy Minimization Method And Its Application To The Synthesis Of Ota Amplifiers, *Proceedings of the second NASA/DoD Workshop on Evolvable Hardware*, July 13-15, 2000, Palo Alto, CA, 133-140.

[Mohammadi, Molho & Santiago (2000)] MOHAMMADI, B., MOLHO, J. I. & SANTIAGO J. A. 2000 Design of Minimal Dispersion Fluidic Channels in a CAD-Free Framework, *Center for Turbulence Research Annual Research Briefs 2000*.

[Hansen & Ostermeier (1996)] HANSEN, N. & OSTERMEIER, A. 1996 Adapting Arbitrary Normal Mutation Distributions in Evolution Strategies: The Covariance Matrix Adaptation, *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation (ICEC '96)*, 312-317.

[Hansen & Ostermeier (1997)] HANSEN, N. & OSTERMEIER, A. 1997 Convergence Properties of Evolution Strategies with the Derandomized Covariance Matrix Adaptation: The $(\mu/\mu_I, \lambda)$-CMA-ES, *Proceedings of the 5th European Conference on Intelligent Techniques and Soft Computing (EUFIT '97)*, 650-654.

[Volpert, Mezic, Meinhart & Dahleh (2000)] VOLPERT, M., MEZIC, I., MEINHART, C.D. & DAHLEH, M. 2000 Modeling and Analysis of Mixing in an Actively Controlled Micromixer, *University of Santa Barbara, Internal Report, 2000*.