# Control of Micromixers, Jets, and Turbine Cooling using Evolution Strategies

Sibylle D. Müller[1] and Petros Koumoutsakos[1,2]

[1] Institute of Computational Sciences, Swiss Federal Institute of Technology (ETH) Zürich, Switzerland
[2] NASA Ames Research Center, Moffett Field, CA, USA

**Abstract.** We present a class of evolution strategies and we discuss three engineering applications in the fields of mixing control and turbomachinery: (i) flow in micromixers, (ii) jet flow, and (iii) turbine cooling. Evolution strategies are chosen as optimization method as they are capable of handling noisy and multimodal functions, inherent to these applications, in an automated fashion.

## 1 Introduction

Evolution strategies (ES) are parameter optimization techniques based on evolution principles such as mutation, selection, and adaptation. In the present article, we discuss ES as an optimization tool capable of handling diverse requirements for engineering applications such as the optimization of mixing (i) in micromixers and (ii) in jet flow, and (iii) the optimization of a cooling configuration in turbine blades. Numerical simulations are used to solve the governing equations of the flow or heat transfer problem. The control of mixing has important engineering applications ranging from micromixers for drug delivery to mixing in exhausts of aerodynamic configuration for noise reduction or to blade cooling in turbomachinery. The optimization of such processes has been subject of several theoretical and experimental works.

ES's offer several advantages for the optimization of such problems as these processes often involve noise, multimodality, or discontinuities of the objective function, which is confirmed by the results of our simulations. Moreover, often gradient information is not available which makes the use of gradient-free optimization techniques necessary. The inherently parallel character of ES's can be exploited in order to compensate for the high computational cost of the simulations.

The paper is organized as follows: In Section 2, we describe evolutionary algorithms in general and the evolution strategies implemented in this work. In Section 3, the micromixing device is described and the optimization results are presented. Jet mixing is presented in Section 4 while Section 5 describes the turbine blade cooling optimization, followed by a summary of our observations in Section 6.

## 2 Evolutionary Algorithms

Evolutionary algorithms are stochastic opimization algorithms which address the following problem: Minimize an objective function which is a mapping from a parameter vector $\boldsymbol{x} \in \boldsymbol{\mathcal{R}}^n$ to $\boldsymbol{\mathcal{R}}$.

They are based on biologically inspired principles such as reproduction, mutation, isolation, recombination, and selection applied on individuals in a population. The members of a population are capable of evolving over time, by adapting to their environment.

In contrast to many deterministic algorithms, evolutionary algorithms require only the value of the objective function for a given point in the parameter space, but do not need any sensitivity information. Certain evolutionary algorithms attempt to obtain this sensitivity information by exploring the search space and formulating prior information. They are comparatively easy to implement and work efficiently for global search. On the other hand, stochastic algorithms do not possess highly efficient convergence properties like e. g. the conjugate gradient method, but this disadvantage is compensated by the parallel character of evolutionary algorithms.

### 2.1 Evolution Strategies

Evolution strategies operate on a population with a number of individuals where each individual is represented by a real-valued vector.

The (1+1)-ES depicts the simplest evolutionary model with a population consisting of two individuals, one parent and one offspring. Here, each individual is identified by a parameter vector $\boldsymbol{x}$ with $n$ components representing its genotype. As basic elements of biological evolution, principles of mutation and selection are applied in this scheme.

In an iterative process, the strategy generates an offspring represented by the vector $\boldsymbol{x}_N^{(g)}$ by adding a normally distributed random vector $\boldsymbol{z} \sim \mathcal{N}(0, {\sigma_i}^2)$ to the parent vector $\boldsymbol{x}_E^{(g)}$:

$$\boldsymbol{x}_N^{(g)} = \boldsymbol{x}_E^{(g)} + \boldsymbol{z} \tag{1}$$

In the following selection step, for both the parent and the offspring the capacity of survival is determined by calculating the values of the objective function. For minimization, the one with the smaller function value is then chosen to be the parent vector of the next generation.

The (1+1)-Evolution Strategy can be written as the following algorithm:

1. Generation of an offspring:

$$\boldsymbol{x}_N^{(g)} = \boldsymbol{x}_E^{(g)} + \delta^{(g)} \cdot \boldsymbol{z} \tag{2}$$

2. Selection:

$$\boldsymbol{x}_E^{(g+1)} = \begin{cases} \boldsymbol{x}_N^{(g)} & , \quad \text{if} \quad f(\boldsymbol{x}_N^{(g)}) \leq f(\boldsymbol{x}_E^{(g)}) \\ \boldsymbol{x}_E^{(g)} & , \quad \text{otherwise} \end{cases} \tag{3}$$

with

$\boldsymbol{x}$ parameter vector to be optimized, $\boldsymbol{x} \in \boldsymbol{\mathcal{R}}^n$,
$g$ generation counter,
E Index of the vector representing the parent (*E*lder),
N Index of the vector representing the offspring(*N*ewer),
$\delta$ Step size determining the length of the mutation step, $\delta \in \boldsymbol{\mathcal{R}}_{>0}$,
$\boldsymbol{z}$ Realization of a $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ distributed random vector, $\boldsymbol{z} \in \boldsymbol{\mathcal{R}}^n$.

3. Increase the generation counter, choose an appropriate value for $\delta^{(g)}$, and go to Step 1 if termination criteria do not hold.

The (1+1)-evolution strategy was developed by Rechenberg and Schwefel in 1963, initially for optimization of experimental fluid dynamics problems [Rechenberg (1973)].

Generalizations of the (1+1) evolution scheme include plus and comma strategies, $(\mu+\lambda)$ and $(\mu, \lambda)$ evolution strategies where $\mu$ means the total number of parents and $\lambda$ the number of offspring.

In "plus" selection schemes $(\mu+\lambda)$, both the parent and offspring population are considered for selection. In a $(\mu+\lambda)$ ES, one or more individuals of the parent generation may survive for as long of a period as no better offspring are produced. In "comma" selection schemes $(\mu, \lambda)$, $\mu$ out of $\lambda$ individuals are selected to create the offspring of the next generation [Schwefel (1995)].

Another principle of biological evolution, recombination, can also be used in evolution strategies. The I in the $(\mu_{\mathrm{I}}, \lambda)$ scheme stands for *i*ntermediate recombination and means that the $\mu$ parents are averaged to create one individual parent.

### 2.2 Mechanisms of Step Size Adaptation

In evolution strategies the choice of strategy parameters such as population size, mutation step size, or selection scheme is critical to the behavior and the convergence properties of the strategy.

A key issue is the adaptation of the mutation distribution during the optimization, e. g. [Schwefel (1995),**?**,**?**]. The adaptation of the step size is regarded to be a crucial factor in the convergence rate as well as the global search properties of an ES. In the following, three mechanisms for step size adaptation of global as well as individual step size will be described. Figure 1 shows lines of equal probability to place children where the center of the circles and ellipses are the location of the parent individual. The different types of step size adaptation include the following cases:

- The adaptation of one global step size yields circles representing lines of equal probability, Figure 1 (a).
- Individual adaptation of step sizes in each coordinate direction of the parameter space corresponds to mutation ellipses which are axis parallel, Figure 1 (b).

– A generalization of the latter adaptation mechanism are correlated mutations which can be generated by a linear transformation of a $\mathcal{N}(\mathbf{0}, \boldsymbol{I})$ - distributed random vector. This type of mutation distribution is represented by arbitrarily orientated mutation ellipses, Figure 1 (c).

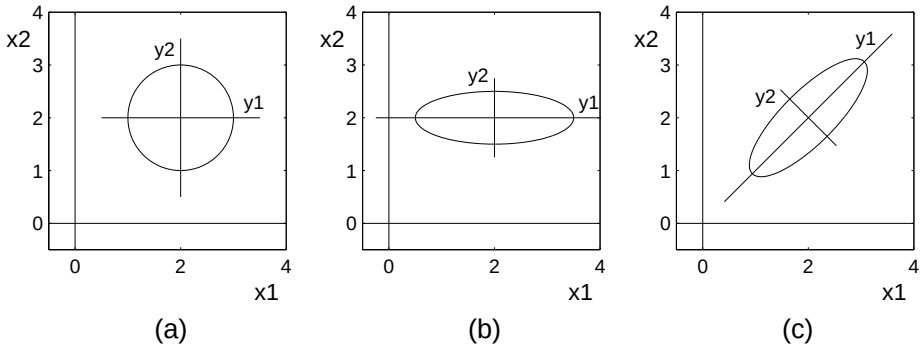


**Fig. 1.** Visualization of probability density distributions for the mutation step in 2 dimensions: (a) isotropic Gaussian probability density distribution of $\boldsymbol{z}$, (b) scaled probability density distribution, (c) scaled and rotated probability density distribution.

### 2.3 Covariance Matrix Adaptation

The Covariance Matrix Adaptation (CMA) approximates a mutation distribution which is properly adapted to the local topology of the fitness landscape. The process performs a principle component analysis of the weighted mutation steps. In contrast to other approaches, the CMA shows invariance properties to general linear transformations of the parameter space [Hansen & Ostermeier (1996)].

Although rapid convergence and global convergence behavior are two contradictory requirements [Schwefel (1995)], it is indicated in [Hansen (1998)] that the adaptation by the CMA-ES can even have positive effects on the global convergence behavior of the search process. Moreover, the CMA technique has been found to converge much faster than evolution strategies with only global and/or individual step size control for non-separable and badly scaled functions [Hansen (1998)].

Following the notation in [Hansen & Ostermeier (1997)] for the $(\mu/_{\mathrm{I}}\mu, \lambda)$-CMA-ES, the offspring of generation $g+1$ are computed by

$$\boldsymbol{x}_k^{(g+1)} = \langle \boldsymbol{x} \rangle_\mu^{(g)} + \delta^{(g)} \boldsymbol{B}^{(g)} \boldsymbol{D}^{(g)} \boldsymbol{z}_k \quad , \quad k = 1(1)\lambda \tag{4}$$

where

$$\langle \boldsymbol{x} \rangle_\mu^{(g)} = \frac{1}{\mu} \sum_{i \in I_{sel}^{(g)}} \boldsymbol{x}_i^{(g)} \tag{5}$$

represents the center of mass of the selected individuals of generation $g$. $I_{sel}^{(g)}$ is the index set of selected individuals at generation $g$, $|I_{sel}^{(g)}| = \mu$. The random vectors $\boldsymbol{z}$ are $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ distributed and serve to generate offspring for generation $g+1$, as can be seen from equation (4). They can be used to calculate

$$\langle \boldsymbol{z} \rangle_{\mu}^{(g+1)} = \sum_{i \in I_{sel}^{(g+1)}} \boldsymbol{z}_{\mu}^{(g+1)}. \tag{6}$$

The columns of $\boldsymbol{B}^{(g)}$ represent eigenvectors of the covariance matrix $\boldsymbol{C}^{(g)}$. $\boldsymbol{D}^{(g)}$ is a diagonal matrix whose elements are identical to the square roots of the eigenvalues of $\boldsymbol{C}^{(g)}$. Hence, the relation of $\boldsymbol{B}^{(g)}$ and $\boldsymbol{D}^{(g)}$ to $\boldsymbol{C}^{(g)}$ can be expressed by

$$\boldsymbol{C}^{(g)} = \boldsymbol{B}^{(g)} \boldsymbol{D}^{(g)} \left( \boldsymbol{B}^{(g)} \boldsymbol{D}^{(g)} \right)^{T}. \tag{7}$$

and

$$\boldsymbol{C}^{(g)} \boldsymbol{b}_i^{(g)} = (d_{ii}^{(g)})^2 \cdot \boldsymbol{b}_i^{(g)} \tag{8}$$

where $\boldsymbol{b}_i^{(g)}$ represents the $i$-th column of $\boldsymbol{B}^{(g)}$. Each covariance matrix corresponds to a hyperellipsoid which defines the surface of equal probability to place offspring. Here, the eigenvectors of the covariance matrix define the orientation of the hyperellipsoid and the eigenvalues define the squared ratio of the lengths of its axes.

Instead of using selection information of a single generation step, one could use a path that the population takes over a number of generations. Such a path is called an evolution path, $\boldsymbol{s}^{(g+1)}$, and is calculated by

$$\boldsymbol{s}^{(g+1)} = (1-c) \cdot \boldsymbol{s}^{(g)} + c_u \cdot \underbrace{\frac{\sqrt{\mu}}{\delta^{(g)}} \left( \langle \boldsymbol{x} \rangle_{\mu}^{(g+1)} - \langle \boldsymbol{x} \rangle_{\mu}^{(g)} \right)}_{= \sqrt{\mu}\, \boldsymbol{B}^{(g)} \boldsymbol{D}^{(g)} \langle \boldsymbol{z} \rangle_{\mu}^{(g+1)}}. \tag{9}$$

With the evolution path, the covariance matrix of generation $g+1$ is build by

$$\boldsymbol{C}^{(g+1)} = (1 - c_{cov}) \cdot \boldsymbol{C}^{(g)} + c_{cov} \cdot \boldsymbol{s}^{(g+1)} \left( \boldsymbol{s}^{(g+1)} \right)^{T}. \tag{10}$$

For the adaptation of the global step size, the evolution path $\boldsymbol{s}_{\delta}^{(g+1)}$, which is not scaled by $\boldsymbol{D}$, is calculated by

$$\boldsymbol{s}_{\delta}^{(g+1)} = (1 - c_{\delta}) \cdot \boldsymbol{s}_{\delta}^{(g)} + \\ + c_{\delta u} \cdot \underbrace{\boldsymbol{B}^{(g)} \left( \boldsymbol{D}^{(g)} \right)^{-1} \left( \boldsymbol{B}^{(g)} \right)^{-1} \frac{\sqrt{\mu}}{\delta^{(g)}} \left( \langle \boldsymbol{x} \rangle_{\mu}^{(g+1)} - \langle \boldsymbol{x} \rangle_{\mu}^{(g)} \right)}_{= \sqrt{\mu}\, \boldsymbol{B}^{(g)} \langle \boldsymbol{z} \rangle_{\mu}^{(g+1)}} \tag{11}$$

and its length is used to compute the overall step size for generation $g+1$

$$\delta^{(g+1)} = \delta^{(g)} \cdot \exp\left(\frac{\|\boldsymbol{s}_\delta^{(g+1)}\| - \hat{\chi}_n}{D\,\hat{\chi}_n}\right). \tag{12}$$

The underbraced terms in equations (9) and (11) emphasize that the evolution path for updating the covariance matrix does not depend on the global step size and that the global step size is not scaled by the lengths of the axes of the mutation ellipsoid.

Recommended strategy parameters are [Hansen & Ostermeier (1997)]:

$$\begin{array}{lll} c &= \frac{1}{\sqrt{n}}, & c_{cov} = \frac{2}{n^2+n}, \quad c_\delta = \frac{1}{\sqrt{n}}, \\ c_u = \sqrt{c \cdot (2-c)}, & c_{\delta u} &= \sqrt{c_\delta \cdot (2-c_\delta)}, \; D = \sqrt{n}. \end{array} \tag{13}$$

Initial values are $\boldsymbol{s}^{(0)} = \mathbf{0}$ and $\boldsymbol{s}_\delta^{(0)} = \mathbf{0}$ and the covariance matrix $\boldsymbol{C}^{(0)}$ is the identity matrix, [Hansen & Ostermeier (1997)]. If $c_{cov} = 0$, then the mutation distribution is isotropic, and the strategy is called cumulative step size adaptation (CSA) evolution strategy.

## 3 Micromixer Optimization

### 3.1 Micromixing Design

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 2 and the flow configuration is illustrated in Figure 3.

As seen from Figure 3, 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

$$U(y, x=0) = U_m \left(1 - \left(\frac{y}{h}\right)^2\right), |y| <= h \tag{14}$$

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

$$V_i = \hat{V}_i \sin(2\pi f_i t + \phi_i), i = 1(1)3 \tag{15}$$

is sinusoidal in time where $f_i$ is the oscillation frequency, $\phi_i$ the phase shift relative to the first set of secondary channels, and $\hat{V}_i$ the velocity amplitude where $\hat{V}_i$ according to [Volpert, Mezic, Meinhart & Dahleh (2000)].

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 $Re = 5$, with $U_m = 1\,\mathrm{m/s}$ and $h = 1\,\mathrm{m}$. In the following, we will skip the SI-units.

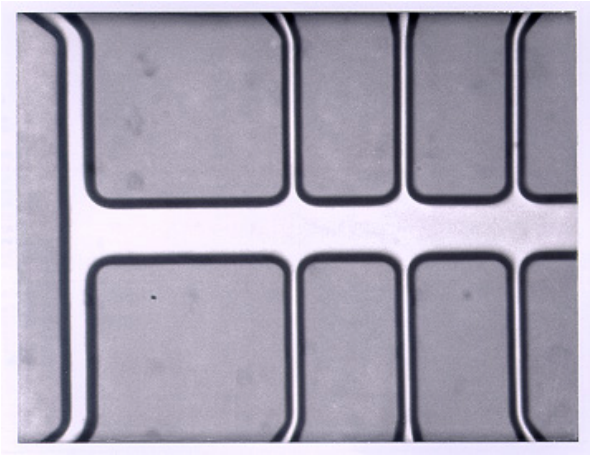

**Fig. 2.** Micrograph of the mixing chip

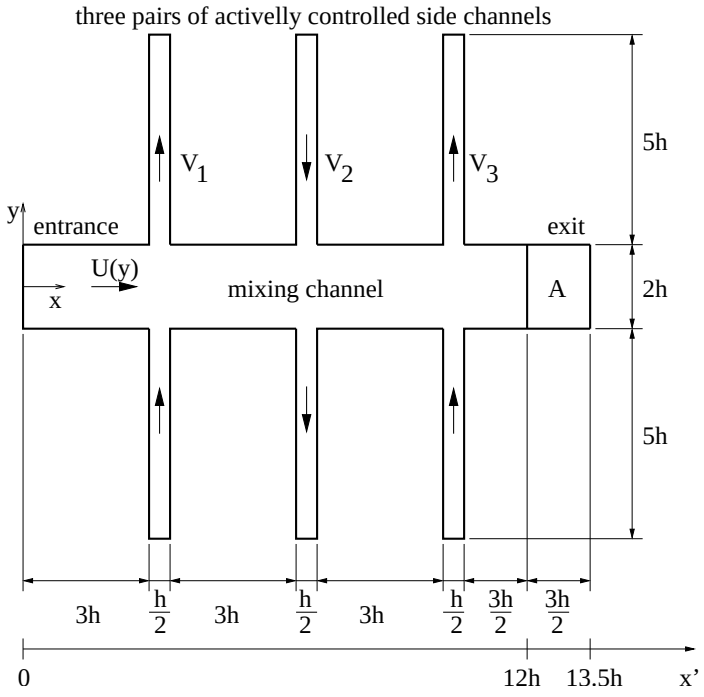


**Fig. 3.** Schematic of the flow configuration

### 3.2 A Theoretical Approach to Determine the Optimal Frequency

Dynamical system theory provides the means to compute the optimum frequency for a mixer with only one channel. In our assumption, optimum mixing occurs when a fluid particle at the lower wall of the main channel is moved to the upper wall and back to the lower wall during one actuation period, thus $l = 4h$ for the distance the particle has traveled during one period. $v$ denotes the velocity of the fluid in the side channel. For the simulation, we set $U_m = 1$ and $h = 1$.

In a first approximation, we set v to the maximum velocity amplitude in the side channel, $v_1 = 2U_m$. This yields a frequency of

$$f = \frac{v_1}{l} = \frac{2U_m}{4h} = \frac{1}{2} \tag{16}$$

In a second approximation, we set v to the average velocity amplitude in the side channel, $v_2 = \frac{2}{3}v_1$, as found for two-dimensional parabolic velocity profiles. Note, however, that a parabolic profile is not used for the simulations. This yields a frequency of

$$f = \frac{v_2}{l} = \frac{\frac{2}{3} \cdot 2U_m}{4h} = \frac{1}{3} \tag{17}$$

In a third approximation, we set v to the average velocity in the side channel,

$$v_3 = v_2 \frac{1}{T/2} \int_0^{T/2} \sin\left(\frac{2\pi t}{T}\right) \mathrm{d}t = \frac{2}{\pi} v_2 \tag{18}$$

This yields a frequency of

$$f = \frac{v_3}{l} = \frac{2v_2}{\pi l} = \frac{2 \cdot \frac{2}{3} \cdot 2U_m}{\pi \cdot 4h} = \frac{2}{3\pi} \approx 0.21 \tag{19}$$

For a mixer with 3 side channels, we need to study the ratio between the 3 frequencies. One can learn that there is a difference if frequency ratios are 1, rational, or irrational. We distinguish the parameter vector, $\boldsymbol{x} = (f_1, f_2, f_3)$, with identical frequencies $\boldsymbol{x} = (1/2, 1/2, 1/2)$, "rational" frequencies $\boldsymbol{x} = (1/2, 1/3, 1/4)$, and "irrational" frequencies $\boldsymbol{x} = (1/(2\sqrt{5}), 1/(2\sqrt{2}), 1/2)$ as proposed in [Volpert, Mezic, Meinhart & Dahleh (2000)].

### 3.3 Optimization Parameters and Objective Function

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 $\Phi_1 = \Phi_2 = 0$, $\Phi_3 = \pi$ as stated 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 $\boldsymbol{x} = (f_1, f_2, f_3)$.

The Navier-Stokes and convection-diffusion equations are discretized using a second order finite volume technique and solved on a Cartesian using a standard computational fluid dynamics package [Star-CD (1997)]. The main flow channel contains $20 \times 135$ grid points in the span- and streamwise direction, respectively, while the secondary channels are discretized using $5 \times 50$ grid points each. The initial conditions involve a uniform concentration of zero and unity for the grid points below and above the symmetry line ($y = 0$), respectively. The mixing rate is estimated from the local variance of the concentration field,

$$m = \overline{c^2} - \overline{c}^2, \tag{20}$$

and $\overline{c}$ denotes the mean (spatial) concentration in the region $A$ (Fig. 3) which extends between $x = 12h$ and $x = 13.5h$. If there is no mixing, $m = 0.25$, and if there is perfect mixing, $m = 0$.

The objective function is the mixing rate $m$, averaged between time $t = 90$ and time $t = 135$, a time regime in which the flow has reached steady-state. Time $t = 135$ corresponds to 10 flow through times which is defined as

$$T_f = \frac{L}{U_m} = \frac{13.5h}{U_m} = 13.5. \tag{21}$$

The CPU time for one function evaluation, i. e. one flow simulation until time $t = 135$, takes about 3 CPU hours on a Sun Sparc Ultra-2 processor.

The chosen optimization method is an evolution strategy with covariance matrix adaptation with intermediate recombination of the parents, called $(\mu_I, \lambda)$-CMA-ES. The population consists of $\mu = 2$ parents and $\lambda \in [6, 8, 10]$ children. Using the explicit message passing (MPI), the optimization is performed in parallel on a cluster of $\lambda/2$ Sun workstations.

### 3.4 Optimization Results

For actuation of one, two, and three frequencies, the initial and the final parameters and mixing rates are given in Table 1.

Besides the three above-mentioned optimizations, we could not afford to run other optimizations with a different direct search technique and compare it with the ES due to the large computational cost. However, we compare the optimization results with dynamical system theory. For 3 actuated frequencies, we obtain the following mixing rates: The parameter vector $\boldsymbol{x} = (1/2, 1/2, 1/2)$ with identical frequencies yields the mixing rate $m = 0.1596$ while our initial "rational" frequency set $\boldsymbol{x} = (1/2, 1/3, 1/4)$ gives us $m = 0.0345$. "Irrational" frequencies as proposed in [Volpert, Mezic, Meinhart & Dahleh (2000)] $\boldsymbol{x} = (1/(2\sqrt{5}), 1/(2\sqrt{2}), 1/2)$ 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 1. This result means a considerable improvement compared with numbers from a well-established theory. Figure 4 shows two snapshots of the flow in the micromixer at time $t = 45$ for the identical and optimal frequencies, respectively.

| Number of actuated frequencies | | 3 | 2 | 1 |
|---|---|---|---|---|
| Initial frequencies | $f_1$ | 0.25 | 0.5 | 0.5 |
| | $f_2$ | 0.33 | 0.5 | 0 |
| | $f_3$ | 0.5 | 0 | 0 |
| Initial mixing rate | $m$ | 0.0345 | 0.17 | 0.18 |
| Best frequencies | $f_1$ | 0.14 | 0.14 | 0.15 |
| | $f_2$ | 0.32 | 0.32 | 0 |
| | $f_3$ | 0.50 | 0 | 0 |
| Best mixing rate | $m$ | 0.0213 | 0.037 | 0.06 |
| Number of function evaluations | | 460 | 420 | 64 |

**Table 1.** Optimization results

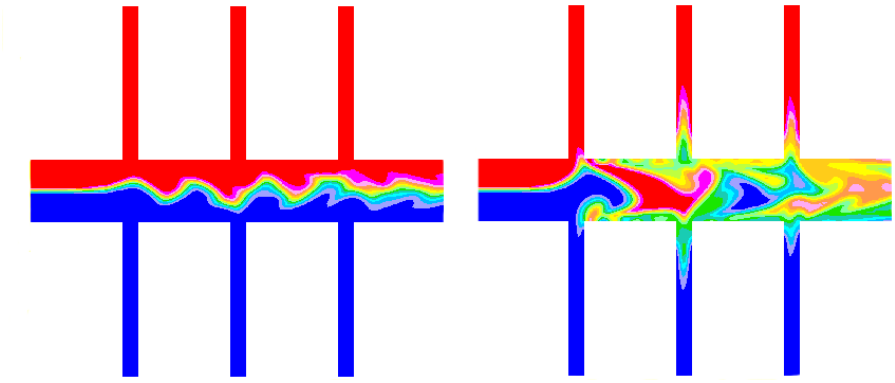

**Fig. 4.** Flow actuated by the frequencies $\boldsymbol{x} = (1/2, 1/2, 1/2)$ (left) and by the optimal frequencies $\boldsymbol{x} = (0.14, 0.32, 0.50)$ (right)

The difficulty for the optimization of two frequencies is the unsteadiness in the region where $f_1 \sim f_2$. We found that when the frequencies are similar, the dynamical behavior of the systems changes such that shedding of fluid packages occurs and averaging the mixing rate in the last third of the simulation time stops being a characterisic measure of the flow behavior. Actually, we did not learn this behavior before optimization. A direct search method without a capability of handling noise would not have been able to overcome the above difficulty and would have converged to the unsteady region. The recombination (= averaging) feature of the CMA-ES managed to escape this region. For the two-frequency actuation, the mixing rate as a function of the actuation parameters is reconstructed from the computed points during the optimization, see Fig. 5.

The optimal frequency of the first side channel appears to be independent of the number of actuated frequencies. In the three optimizations, the first parameter attains values of about $f_1 = 0.14 - 0.15$. The same phenomenon occurs

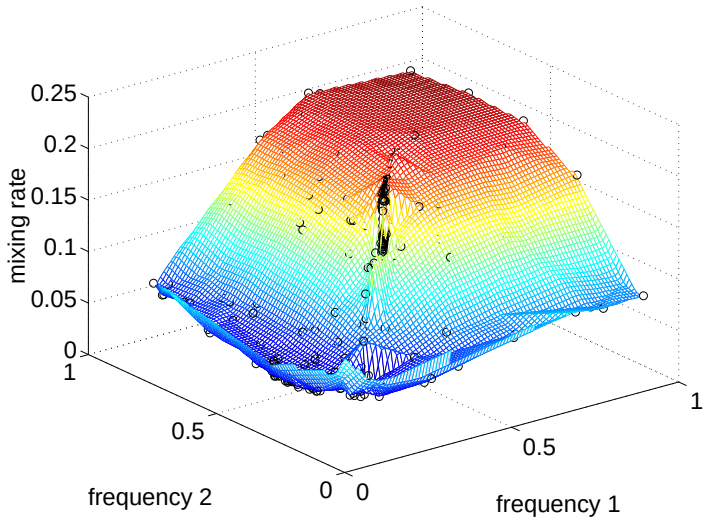


**Fig. 5.** Surface of the mixing rate as a function of the two actuation frequencies

for the second frequency for the optimization of two and three frequencies where $f_2 = 0.32$.

## 4 Jet Mixing

The control of jet flows has applications in various fields such as combustion, aerodynamic noise, and jet propulsion. The mixing rate of a jet can be altered by applying a suitable excitation at the jet orifice. We examine a prototypical configuration for jet mixing using a vortex filament method to simulate temporally varying jets. The vortex filaments are subject to azimuthal perturbations. Our objective is to maximize the length of all vortex filaments, i.e. to find the perturbation parameters (amplitudes and phase shifts) that cause the greatest perturbation. Optimization parameters are the phase shifts $\beta_j$, $(j = 2, ..., 5)$ and the amplitudes $\epsilon_j$, $(j = 1, ..., 5)$ of the azimuthal perturbation. The phase shift $\beta_1$ is set to zero. The amplitude $\epsilon_j = 0.2$ given by [Martin & Meiburg (1992)] based on linear stability arguments corresponds to a perturbation of 4%. For the optimization, we implement a (1+1)-ES running for 300 generations with initial values given by [Martin (2001)]. With the phases constrained between 0 and 1, two optimizations are performed. The first optimization, for which the amplitudes are constrained between 0 and 1, yields an increase of 42 % for the filament length. In the second optimization with lower and upper amplitudes of 0 and 0.25, respectively, the filament length is increased by only 1 %. This indicates that the upper limit of the amplitudes has a high impact on the optimization result and therefore should be chosen carefully. Figures 6 and 7 show

the filament structure of the upper half jet from an axial view. The structure resulting from the first optimization seen in Figure 3 has longer filaments than the one from the initial state seen in Figure 2.

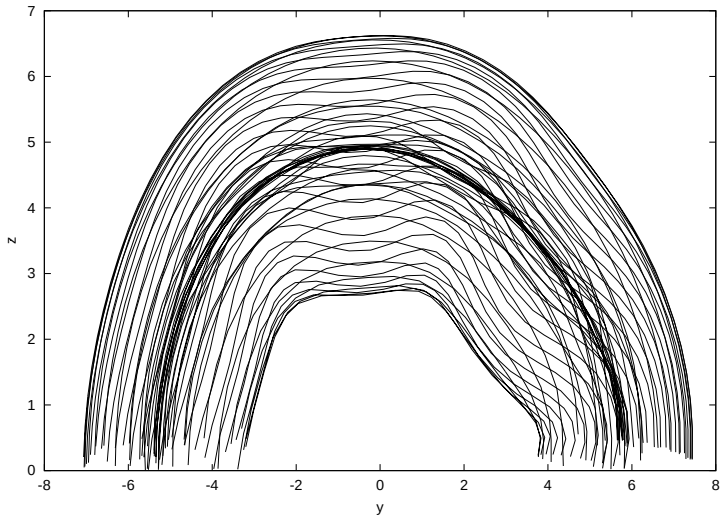


**Fig. 6.** Filament structure of the upper half of the jet, excitation with initial values

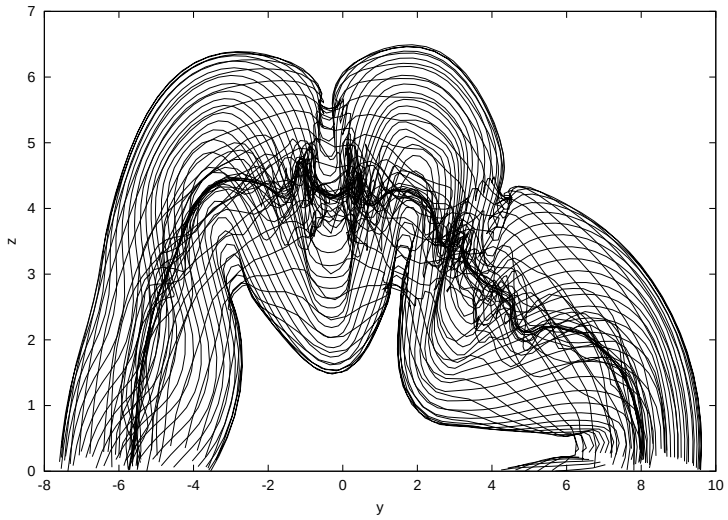


**Fig. 7.** Filament structure of the upper half of the jet, excitation with optimum values (first optimization)

## 5 Turbine Blade Film Cooling

An evolutionary algorithm is implemented in a realistic automated design cycle of turbine blade film cooling. This design cycle involves the use of empirical formulas for the flow calculation and the use of a commercial software package for the simulation of the heat transfer problem. The overall process consists of an engineering multiobjective optimization problem with constraints. In this work we consider the solution of this problem in the context of an automated optimization cycle using evolution strategies. In film cooling, the coolant leaves the blade through rows of holes, forming a protective film on the outer surface of the blade, separating the hot gas from the metal. The design of the cooling configuration aims at both minimizing the coolant supply and achieving optimum cooling, while obeying certain engineering constraints. For the simulation of the cooling configuration, the mass flow of the coolant and the surface temperature of the blade are computed, based on given external aerodynamic and thermal conditions for a two-dimensional blade geometry.

### 5.1 Description of a Blade Cooling Configuration

A typical vane cooling configuration is shown in Figure 5.4. It consists of three components: (i) an insert for impingement cooling; (ii) the trailing edge geometry for convection cooling; (iii) coolant air leaving the inside of the blade through holes and keeping attached as a film, thereby cooling the outer surface. The coolant flow is described by three different cooling mechanisms based on empirical correlations which compute the heat transfer coefficient and the temperature in the boundary layer [McGreehan & Schotsch (1987)], [Goldstein (1971)], [Seller (1963)]. The geometry as well as the external aerodynamic and thermal conditions are considered two-dimensional and the shape of the blade is fixed.

### 5.2 Optimization Algorithm and Parameters

Two optimization techniques [Hansen & Ostermeier (1997)] are chosen: The first method is an evolution strategy with covariance matrix adaptation (CMA-ES) while the second method is an evolution strategy with cumulative step size adaptation and isotropic mutation distribution (CSA-ES), see Section 2. Both methods work with a population size of $\mu = 3$ and $\lambda = 12$ and with intermediate recombination.

The optimization parameters are the number of film cooling rows ($R_f$), the position of film cooling rows ($s_{pos}(1 \text{ to } R_f)$), the injection angles of the holes, measured from the hole axis to the surface ($\alpha(1 \text{ to } R_f)$), the number of holes per row ($N_r(1 \text{ to } R_f)$), the number of impingement holes ($N_j$), and the loss parameter that is a function of the pressure loss in the trailing edge ($A_{loss}$).

The only integer parameter in the blade cooling optimization is the number of rows. Note that optimizing integer variables with an evolution strategy only makes sense if the number of possibilities to adjust the variables is large enough. Due to manufacturing requirements, the number of rows should not exceed about

10. With respect to this small number, this parameter is not included in the optimization process but rather changed manually for different optimization runs. First tries with $R_f = 4$ showed that the constraints for maximum and minimum temperature cannot be met. To homogenize the temperature distribution, we increased the number of rows to $R_f = 7$. This number was found to be sufficient for achieving the optimization goals and is used in this study.

All the other parameters are real valued. For each film row $j = 1, \ldots, R_f$, the position $s_{pos}$, the angle $\alpha$, and the number of holes per row $N_r$ can be chosen. The other two parameters are the number of impingement holes $N_j$ and the loss parameter $A_{loss}$. There are $3 \cdot R_f + 2$ parameters altogether; in our case 23 parameters.

The limits of the optimization parameters are set to $0 <= s_{pos} <= 1$, $30^o <= \alpha <= 90^o$, $10 <= N_r <= 100$, $500 <= N_j <= 1000$, $1 <= A_{loss} <= 3$. The parameter range of $s_{pos}$ covers the whole surface starting from $s_{pos} = 0$ at the trailing edge, continuing on the suction side to the stagnation point and on the pressure side back to the trailing edge. For the angle $\alpha$, a special constraint is imposed by engineering considerations. In the leading edge of the blade, defined by values of $s_{pos}$ between 0.49 and 0.53, holes with angles other than $90^o$ are difficult to manufacture due to the high bend. Therefore, holes in the leading edge region should always be set to $90^o$. This is a constraint that can be neglected for the following reason: In all of the optimization runs, the objective function values are not sensitive to changes of the hole angles. This indicates that the influence of the injection angle is not sufficiently incorporated in the model. Instead of the constraint, we compare the outputs of the converged solution with the outputs computed with angles of $90^o$ for rows of holes in the leading edge. For all of our optimization runs, the difference between both outputs was below 1%.

### 5.3 The Objective Function

The main goal of the optimization is to minimize the mass flow of the coolant, $\dot{m}_c$. At the same time, constraints have to be met. These constraints are expressed by one equality constraint $\varepsilon_{mean} = 0.5$ where $\varepsilon_{mean}$ is the normalized mean surface temperature, and also by the two inequality constraints $\varepsilon_{max} <= 0.6$ and $\varepsilon_{min} >= 0.4$ where $\varepsilon_{max}$ is a function of the minimum temperature at the outer surface, and where $\varepsilon_{min}$ is a function of the maximum temperature. These constraints are incorporated in the objective function $f$ using a distance method with Euclidian metric, $f = \left(\sum_{i=1}^{4} f_i^2\right)^{0.5}$ where $f_1 = \dot{m}_c$, $f_2 = 2|\varepsilon_{mean} - 0.5|$, and

$$f_3 = \begin{cases} 10000 \cdot (\varepsilon_{max} - 0.5)^4 & \text{if } \varepsilon_{max} >= 0.5, \\ 0 & \text{if } \varepsilon_{max} < 0.5. \end{cases}$$

$$f_4 = \begin{cases} 10000 \cdot (\varepsilon_{min} - 0.5)^4 & \text{if } \varepsilon_{min} <= 0.5, \\ 0 & \text{if } \varepsilon_{min} > 0.5. \end{cases}$$

The coefficients and the exponent can be considered weights and are chosen such that the function $f_2$ is in the same order of magnitude as $f_3$ and $f_4$ for the initial parameters.

The computation of the optimization objectives entails two steps: The mass flow of the coolant as well as the air temperatures and heat transfer coefficients in the boundary layer as a function of the geometry have to be computed first. In a second step, the steady heat equation is solved for the unknown temperature distribution using second order finite volume techniques subject to the aforementioned boundary conditions. The blade geometry is discretized using 660 x 20 mesh points and the solution is obtained applying a commercial package, STAR-CD [Star-CD (1997)]. One function evaluation is done in about 15 CPU seconds on a Sun Ultra 60/Creator 3D, 300 MHz.

### 5.4 Results

With the initial parameter configuration $s_{pos} = (0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8)$, $\alpha = 50^o$, $N_r = 50$, $N_j = 800$, $A_{loss} = 2$, we obtain the objectives: $\dot{m}_c = 0.367$, $\varepsilon_{mean} = 0.522$, $\varepsilon_{max} = 0.616$, $\varepsilon_{min} = 0.360$, $f = 4.22$.

A reduction of the initial objective function value by 60 % is achieved with the CMA-ES already after about 200 generations (= 2400 function evaluations). After about 1800 generations (= 21600 function evaluations), convergence is reached.

Optimum parameters $s_{pos} = (0.490, 0.500, 0.502, 0.508, 0.510, 0.706, 0.786)$, $\alpha = (65^o, 90^o, 89^o, 30^o, 30^o, 73^o, 30^o)$, $N_r = (100, 10, 13, 95, 100, 35, 41)$, $N_j = 501$, $A_{loss} = 2.97$ yield the outputs: $\dot{m}_c = 0.257$, $\varepsilon_{mean} = 0.502$, $\varepsilon_{max} = 0.595$, $\varepsilon_{min} = 0.400$, $f = 0.229$. Here, the reduction of the cost function equals 95 %.

The results of this study are summarized as follows:

- The original row configuration (rows equidistantly placed) was drastically modified. The final configuration suggests to place five rows in the leading edge and two rows on the pressure side, as seen in Figure 5.4.
  The number of holes per row attains small values for the rows on the pressure side whereas two rows in the leading edge are equal to the upper limit, $N_r = 100$. This result makes sense when considering that the hot gas is approaching the leading edge of the turbine, thus requiring efficient cooling in this area.
- The mass flow can be reduced significantly and the mean temperature distribution comes closer to the desired one. At the same time, the maximum and minimum temperatures are not exceeded, as seen from the constraints $\varepsilon_{max} < 0.6$ and $\varepsilon_{min} > 0.4$ that are met.
- As one can see from Figure 9 where the improvement of the multi-objective function is shown, convergence is reached after about 1800 generations (= 21600 function evaluations in a population with 12 offspring) for the evolution strategy with covariance matrix adaptation. Convergence is defined to be reached when the difference of the cost function for the best individual is less than $10^{-3}$ (in absolute values) over 10 generations. On the

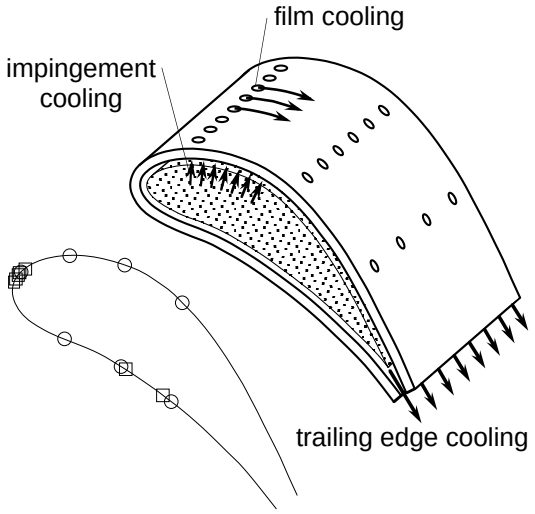


**Fig. 8.** The three cooling mechanisms and the initial (circles) and final (squares) row positions from the optimization with the CMA-ES .

other hand, the CSA-ES converges slower and cannot reach the same good value as the CMA-ES within the same number of iterations. The difference in convergence properties of these two types of evolution strategies agrees well with results from the optimization of several test functions [Hansen & Ostermeier (1997)].

– Although evolution strategies are often considered good candidates for global optimization, they may also converge to local optima. Therefore, to find a good optimum it is recommended to perform several optimization runs with different start configurations. Since the computations in our case are relatively expensive, we optimized with only two different initial settings of the strategy parameters. The converged function values in both optimization runs did not significantly $(1 - 2\%)$ differ from each other. The final model parameters were not the same but they represented a very similar configuration with five rows close to each other in the leading edge region and two rows of holes on the pressure side.

Interpreting the results, one should keep in mind that we use an empirical model that may not take into full account the properties of the flow and heat problem. Moreover, the results could be improved by replacing the combination of the several objectives into one single objective function by searching the Pareto-front of the multi-objective problem.

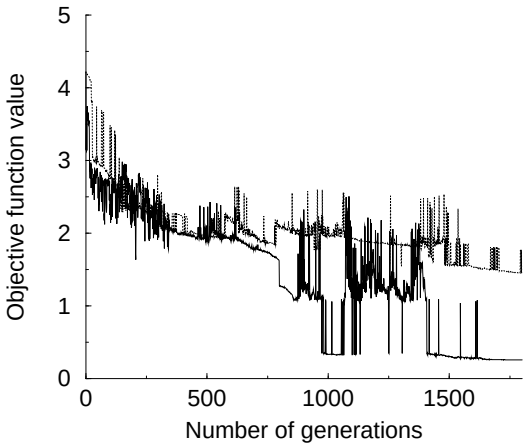


**Fig. 9.** Convergence of the objective function f (— CMA-ES, $\cdots$ CSA-ES)

## 6 Summary and Conclusions

Evolution strategies were applied to three engineering problems, (i) flow in micromixers, (ii) jet flow, and (iii) turbine blade cooling. In case of the micromixing example, the evolutionary optimization results outperformed results from a well-established theory. Moreover, the chosen evolution strategy was able to handle noise during the optimization process and did not get trapped one of the sub-optimal solutions. For the jet flow mixing, even a simple evolution strategy, the (1+1)-ES can yield a great improvement over solutions found by engineering intuition. Turbine blade cooling is an for the use of evolution strategies to optimize efficiently a complex engineering problem in an automated fashion. In summary, the results show that these optimization methods are highly suitable for optimization in applications that are characterized by noise, multimodality, and no availability of gradient information.

## References

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

[Rechenberg (1973)] Rechenberg, I., "Evolutionsstrategie : Optimierung technischer System nach Prinzipien der biologischen Evolution," Fromann-Holzboog, Stuttgart, 1973.

[Schwefel (1995)] Schwefel, H.-P., "Evolution and Optimum Seeking," John Wiley and Sons, New York, 1995.

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

[Ostermeier (1997)] Ostermeier, A., "Schrittweitenadaptation in der Evolutionsstrategie mit einem entstochastisierten Ansatz." Dissertation, Fachbereich 6 der Technische Universität Berlin, 1997.

[Hansen (1998)] Hansen, N., "Verallgemeinerte individuelle Schrittweitenregelung in der Evolutionsstrategie: Eine Untersuchung zur entstochastisierten, koordinatensystemunabhängigen Adaptation der Mutationsverteilung," Mensch & Buch Verlag, Berlin, 1998.

[Martin & Meiburg (1992)] J.E. Martin and E. Meiburg, "Numerical investigation of three-dimensionally evolving jets under helical perturbations," *Journal of Fluid Mechanics,* Vol. 243, pp. 457-487, 1992.

[Martin (2001)] J.E. Martin, *Personal communication,* 2001.

[Volpert, Mezic, Meinhart & Dahleh (2000)] Volpert, M., Mezic, I., Meinhart, C.D., Dahleh, M., "Modeling and Analysis of Mixing in an Actively Controlled Micromixer," *Unpublished report, University of Santa Barbara, CA*, 2000.

[McGreehan & Schotsch (1987)] McGreehan, W.F., Schotsch, M.J. "Flow characteristics of long orifices with rotation and corner radiusing," ASME paper 87-GT-162, 1987.

[Goldstein (1971)] Goldstein, R.J., "Film cooling," *Advances in Heat Transfer*, Vol. 7, pp. 321-379, 1971.

[Seller (1963)] Seller, J.P., "Gaseous film cooling with multiple ejection stations," *AIAA Journal*, Vol. 1, No. 9, pp. 2154-2156, 1963.

[Bestle (1993)] Bestle, D., "Analyse und Optimierung von Mehrkörpersystemen," Springer, Berlin, 1993.

[Star-CD (1997)] *Star-CD User Manual,* Computational Dynamics Ltd., London, 1997.