UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Ingegneria Industriale

Scuola di Dottorato di Ricerca in Ingegneria Industriale
Indirizzo: Ingegneria dell'Energia
Ciclo: XXVI

DEVELOPMENT OF NUMERICAL PROCEDURES FOR TURBOMACHINERY OPTIMIZATION

Direttore della Scuola: Ch.mo Prof. Paolo Colombo

Coordinatore d'indirizzo: Ch.ma Prof.ssa Luisa Rossetto

Supervisore: Ch.mo Prof. Ernesto Benini

Dottorando: Giovanni Venturelli

Mater

Sommario

“Sviluppo di procedure numeriche per l’ottimizzazione di turbomacchine” raccoglie la ricerca svolta dall’autore nel periodo di Dottorato che va dal 2010 al 2013. Il lavoro è nato con una duplice finalità: da una parte sviluppare un algoritmo per l’ottimizzazione multi obiettivo; dall’altra, accoppiare il motore di ottimizzazione con strumenti di analisi basati sulla fluidodinamica computazionale (CFD) per studiare casi di interesse nell’ambito del “high speed turbomachinery”.

Gli algoritmi evolutivi hanno dimostrato alta affidabilità e robustezza nel raggiungimento del “Fronte di Pareto” (i.e., è la soluzione di un problema multi obiettivo), richiedendo però un numero di valutazioni delle funzioni obiettivo molto elevato, talvolta impraticabile dal punto di vista industriale. Infatti, quando la CFD è impiegata per valutare le funzioni obiettivo del sistema in esame, il costo computazionale può diventare il vero collo di bottiglia dell’intero processo. Una possibile soluzione viene fornita dai modelli surrogati, o metamodelli, cioè tecniche matematiche il cui scopo è quello di approssimare le funzioni obiettivo permettendo, di fatto, di diminuire le chiamate dirette alla CFD e di conseguenza anche il tempo totale del processo di ottimizzazione. Il vero dilemma è come affiancare gli algoritmi evoluti a uno o a più modelli surrogati, al fine di migliorare le prestazioni del motore di ottimizzazione. A oggi il problema non ha una soluzione univoca.

La tesi è costituita da cinque capitoli. Il primo capitolo vuol essere di introduzione sia ai modelli surrogati visti nell'ottica dell’ottimizzazione, sia alle strategie di ottimizzazione che sono state applicate per migliorare i compressori transonici e le schiere supersoniche di compressori, che rappresentano i casi di interesse studiati in questa Tesi. Il secondo capitolo è dedicato al motore di ottimizzazione sviluppato dall’autore, denominato GeDEA-II-K. Il GeDEA-II-K nasce dall’unione del preesistente algoritmo genetico GeDEA-II e di un modello surrogato basato sul Kriging. Le prestazioni del nuovo algoritmo sono state testate su problemi matematici a due e a tre obiettivi ben noti in letteratura. Nel terzo capitolo è stato approfondito in grande dettaglio la fisica alla base delle schiere supersoniche, cercando di comprendere il legame profondo tra la geometria della schiera e il campo di moto che si viene a creare. Nel quarto e nel quinto capitolo è

stato testato il loop automatico di ottimizzazione sviluppato dall'autore che comprende il motore di ottimizzazione, il tool di parametrizzazione della geometria, i modelli CFD, e tutti quegli elementi indispensabili per garantire robustezza ad una procedura automatica. Nello specifico è stata condotta l'ottimizzazione di una schiera supersonica e del compressore transonico NASA Rotor 67.

Abstract

This Doctoral Thesis deals with high speed turbomachinery optimization and all those tools employed in the optimization process, mainly the optimization algorithm, the parameterization framework and the automatic CFD-based optimization loop. Optimization itself is not just a mean to improve the performance of a generic system, but can be a powerful instigator that helps gaining insight on the physic phenomena behind the observed improvements.

As for the optimization engine, a novel surrogate-assisted (SA) genetic algorithm for multi-objective optimization problems, namely GeDEA-II-K, was developed. GeDEA-II-K is grounded on the cooperation between a genetic algorithm, namely GeDEA-II, and the Kriging methodology, with the aim at speeding up the optimization process by taking advantage of the surrogate model. The comparison over two- and three-objective test functions revealed the effectiveness of GeDEA-II-K approach.

In order to carry out high speed turbomachinery optimizations, an automatic CFD-based optimization loop built around GeDEA-II-K was constructed. The loop was realized for a UNIX/Linux cluster environment in order to exploit the computational resources of parallel computing. Among the tools, a dedicated parameterization framework for 2D airfoils and 3D blades has been designed based on the displacement filed approach.

The effectiveness of both the CFD-based automatic loop and the parameterization was verified on two real-life multi-objective optimization problems: the 2D shape optimization of a supersonic compressor cascade and the 3D shape optimization of the NASA Rotor 67. To better understand the outcomes of the optimization process, a wide section has been dedicated to supersonic flows and their behavior when forced to work throughout compressor cascades.

The results obtained surely have demonstrated the effectiveness of the optimization approach, and even more have given deep insight on the physic of supersonic flows in the high speed turbomachinery applications that were studied.

Contents

List of Figures

List of Tables

List of symbols

$\| \|$	Euclidean distance
$\| \|$	General distance
$\boldsymbol{a}$	Vector of regression parameters
a_i	Linear weight for the constraint violation function
$B_{i,j}()$	Spline basis function, i-th control point, j-th degree
C	Airfoil chord
c_{ax}	Airfoil axial chord
$c_{\theta_1}, c_{\theta_2}$	Tangential velocity components in the absolute frame of reference
C	SVR regularization constant
$EI\left(\boldsymbol{x}\right)$	Expected Improvement
f	Functional relationship between input variable and response; low-fidelity model response
$f_{\min}$	Minimum of the function
F	High-fidelity model response
$\boldsymbol{F(P)}$	Fitness functions vector
G, G_1, G_2	Constraint violation function
h	Specific enthalpy
$\boldsymbol{I}$	Identity matrix
k	Dimension of input space (design variable space)
L	SVR loss function
LH	Likelihood function
m	Order of polynomial
$\dot{m}$	Mass flow rate
$\dot{m}_{choke\_bl}$	Choking mass flow of the baseline configuration
$\dot{m}_{choke}$	Choking mass flow of a generic design
M	Mach number
M_{target}	Target Mach number
M_∞, M_1, M_2	Mach number at infinite, section 1, section2, respectively
np	Number of observations
N	Number of basis functions

p, p_l	Kriging smoothness parameter
p_1, p_2	Static Pressure at station 1 and 2, respectively
p_{01}, p_{02}	Total pressure at station 1 and 2, respectively
$\boldsymbol{p}()$	Generic spline model
$\boldsymbol{p}_{base}, \boldsymbol{p}_{\mathrm{mod}}, \boldsymbol{p}_{displ}$	Spline model of baseline geometry, modified geometry, displacement field, respectively
$\boldsymbol{P}$	Optimization design variables vector
$\boldsymbol{Q}$	Optimization penalty term
$\boldsymbol{r}$	Kriging vector of correlations
r_1, r_2	Transformation coefficients for the optimization penalty term
$\boldsymbol{R}$	Kriging Correlation matrix
s	Cascade spacing
t_i	Spline knot values
u	Spline natural parameter
w	Spline natural parameter
$w^{(i)}$	Weight of the *i*-th observation
$\boldsymbol{W}$	Weighting matrix
$\boldsymbol{x}$	Single observation or vector of observations ($npxk$)
$\boldsymbol{x}^{(i)}$	*i-th* observation
$\boldsymbol{x}_j^{(i)}$	*j*-th component of the *i*-th observation
$\boldsymbol{x}^*$	Unknown point for prediction
y	Observed response
$\hat{y}$	Predictor response
$\boldsymbol{y}$	Vector of observed responses ($npx1$)
$Y(\boldsymbol{x})$	Response of random function
Z_e, Z_c, Z_d	Expensive, Cheap and Gaussian process functions, respectively

Greek Symbols

$\beta_\infty, \beta_1, \beta_2$	Flow angle at infinite, section1, section2, respectively.
γ	Stagger angle
$\boldsymbol{\varepsilon}$	Measurement error; SVR accepted error
η	Coordinate along the direction orthogonal to the airfoil chord
η_{iso}	Isentropic efficiency
η_{pol}	Polytropic efficiency
θ, θ_l	Kriging correlation parameter
λ	Regularization parameter; size of the offspring population

μ	Size of the parents' population
$\boldsymbol{\mu}$	Vector of basis function centres
$\boldsymbol{\mu}^{(i)}$	*i-th* basis function centre
$\nu(M)$	Prandtl-Meyer function
ν_i	Kriging polynomial functions
ξ	Coordinate along the airfoil chord direction
π_c	Total pressure ratio; pressure ratio
ρ	Scaling factor
σ	Model parameter, bandwidth
$\boldsymbol{\Phi}$	Vandermonde matrix
$\boldsymbol{\Psi}$	Gram matrix
ω	Cascade loss coefficient

Abbreviations

ADEN	Adimensional Direct Evaluations Number
ASEMOO	Asynchronous Efficient Multi-Objective Optimization
AVDR	Axial Velocity Density Ratio
BPNN	Back Propagation Neural Networks
CFD	Computational Fluid Dynamics
CV	Cross-Validation
DOE	Design Of Experiments
EA	Evolutionary Algorithm
EHVI	Expected Hyper-Volume Improvement
E-O	Efficiency Optimized
FEM	Finite Element Method
GCV	Generalized Cross-Validation
HV	Hyper-Volume
LRP	Local Polynomial Regression
MLS	Moving Least-Squares
MLSM	Mach Loop Status Machine
MOEA	Multi-Objective Evolutionary Algorithm
MOOP	Multi-Objective Optimization Problem
PM	Pure Methods
RBF	Radial Basis Function
ROLS	Regularized Orthogonal Least-Square
RSM	Response Surface Model
SA	Surrogate-Assisted, Spalart-Almaras
SPX	Simplex Crossover
SVM	Support Vector Machine
SVR	Support Vector Regression
TPR-O	Total Pressure Ratio Optimized
UICL	Unique Incidence Control Loop
UISM	Unique Incidence Status Machine
WLS	Weighted Least-Squares

Chapter 1

Introduction

1.1 Why Metamodels?

Multi-Objective Evolutionary Algorithms (MOEAs) seem to be the most powerful tools apt to handle challenging real-world engineering optimization problems in which conflicting objective must be contextually optimized (e.g. Multi-Objective Optimization Problem (MOOP)). Owing to the inherent stochastic nature of MOEAs, a large number of individual evaluations are required to converge properly toward the *true* Pareto front Durillo *et al.* [**1**]. When computational expensive models are used for evaluating individual's performance (i.e. high-fidelity models as Computational Fluid Dynamics CFD and Finite Element Method FEM), the computational effort becomes one of the key issues of the entire optimization process. Although a way to limit computational resources is limiting the problem complexity by reducing the number of design variables, such strategy sounds more as a fold back on rather than an optimum solution. To try to enhance the search efficiency of MOEAs, in the last decade surrogate models, called also metamodels, or approximation models, or response surfaces, have been successfully employed.

Looking at the big picture, irrespective of the appellation given to metamodels, surrogates offer an efficient way with which information coming out from the expensive model can be collected, manipulated and exploited to improve the search efficiency. The ability of doing so, however, is not just an inherent peculiarity of metamodel itself, but greatly depends on how metamodels are integrated within the optimization framework. Metamodels infer knowledge from the *true information*; however, metamodels themselves do not create new information.

1.2 Metamodels in Practice

The final aim of metamodel techniques is to construct a reliable predictor of the deterministic function we would like to emulate, bearing in mind that willing accurate prediction all over the function landscape would involve much more resources than willing it only in the region of the optimum.

Despite the huge amount of different metamodels offered on the market, it is possible to make a few classifications. First, there are *global* and *local* metamodels techniques (Sacks *et al.* [**2**]) depending on the validity of the approximation on the design space: local surrogates are valid in a neighborhood of a point, whereas global methods provide information throughout all the design space, or a large portion of it. There is also a third category so-called *midrange* approximation technique, which provide local approximation exploiting global qualities (Toporov [**3**]). Pay attention to the terminology *local* and *global* because it is also used to describe how fitted data influence predictor‘s approximation; this ambiguity will be clarified in §1.2.2.

Moreover, we distinguish other two main categories: *parametric* and *nonparametric* techniques. A parametric technique is the one which obeys the rules of the conventional statistical regression analysis Hill and Lewicki [**4**], and is often used within the realm of physical experiments as it typically smoothes out the random errors which inevitably affect the tests. It basically consists in predefining a form of a response surface, usually of the low-order polynomial type because of their intrinsic “physical” meaning, the unknown coefficients of which are determined using a generalized least-square regression criterion to fit the response surface predicted values to the observed data.

On the other hand, in a nonparametric approach the response function is not assumed to belong to a specified parametric class of functions: on the contrary, it is only supposed to obey to a few and rather general smoothness conditions. The very attractive feature of this approach is that data to be fitted is not forced into a prescribed mathematical structure in order for the unknown model parameters to be determined, but it is left free to build the statistical model on its own without being trapped into a predefined, constrained formulation. In other words, the response function is identified only on the basis of the assigned data, and its determination becomes actually the final goal of the model identification. In this sense it is called “nonparametric”, i.e. not because it is parameterless but because the goal of the regression is now to estimate the regression function f directly rather than the parameters.

1.2.1 Parametric Models

1.2.1.1 Response Surface Model (RSM)

RSM consists in a collection of statistical and mathematical techniques for parametric model building, aimed at developing a reliable model that exhibits the highest correlation with observations, while keeping the number of explanatory variables to a minimum (Box [**5**]). In general, the underlying assumption is that data coming out from the function to be estimated has the following form:

$$y = f(\boldsymbol{x}) + \varepsilon \qquad \text{Eq. 1.1}$$

in which the response y has got two contributions: the first, $f(\boldsymbol{x})$, accounts for the systematic component of the functional relationship between the response y and the independent variables $\boldsymbol{x}$, the second, ε, is the measurement random error that is assumed to be normally, identically and independently distributed, with zero mean and constant variance. In RSM approach $f(\boldsymbol{x})$ is modeled as polynomial of order m and, in the case $\boldsymbol{x}$ has dimension $k = 1$, it can written as:

$$\hat{y}(\boldsymbol{x}) = \sum_{i=0}^{m} \boldsymbol{a}_i \boldsymbol{x}^i \qquad \text{Eq. 1.2}$$

where the predictor $\hat{y}$ is linear in the parameters $\boldsymbol{a}_i$ (unknown vector of regression coefficients), which can be estimated through a least squares solution of $\boldsymbol{\Phi a} = \boldsymbol{y}$, where $\boldsymbol{\Phi}$ is the $(npxk)$ Vandermonde matrix and $\boldsymbol{y}$ is the $(npx1)$ vector of observed responses. The maximum likelihood estimate of $\boldsymbol{a}$ is thus

$$\boldsymbol{a} = \left(\boldsymbol{\Phi}^T \boldsymbol{\Phi}\right)^{-1} \boldsymbol{\Phi}^T \boldsymbol{y} \qquad \text{Eq. 1.3}$$

The extension of RMS to multivariate space is straightforward.

Since the polynomial approximation of order m is similar to a Taylor series expansion truncated after m+1 terms (Box [**5**]), it is clear that the higher is m the better is the approximation accuracy. However, a high m order also implies a more flexible polynomial that can be trapped in noisy data with the danger of over fitting. As a result, a second-order polynomial is, *de facto*, the most popular solutions in applications.

1.2.2 Nonparametric Techniques

While polynomial models can be regarded for as "global" models, in which both the observations near to (in the Euclidean distance sense) and far from a location $\boldsymbol{x}$ in the input parameters' domain equally influence the predicted response over $\boldsymbol{x}$, nonparametric approaches have a somewhat "local" character Giunta *et al.* [**6**]. Specifically, the closer the available observations to $\boldsymbol{x}$, the higher their weight in the determination of the predicted response $\hat{y}(\boldsymbol{x})$. This seems particularly attractive when the unknown response function is highly multimodal.

1.2.2.1 Moving Least-Squares

Moving Least-Squares technique (MLS), often referred to as Local Polynomial Regression (LRP) (Toporov *et al.* [**7**] and Cleveland [**8**]), was developed to overcome well-known drawbacks of traditional multivariate polynomials, such as excessive smoothing that makes them not flexible enough to achieve an adequate fit, as well as their attitude to exalt individual observations influence on remote parts of the fitted hypersurface.

The MLS derives from and improves the Weighted Least-Squares (WLS) approach. In WLS, observations may not be equally important in estimating the polynomial coefficients. To this end, each observation is given a weighting $w^{(i)} \geq 0$. With $w^{(i)} = 0$ the observation is neglected in the fitting. The coefficients of WLS model are

$$\boldsymbol{a} = \left(\boldsymbol{\Phi}^T \boldsymbol{W} \boldsymbol{\Phi}\right)^{-1} \boldsymbol{\Phi}^T \boldsymbol{W} \boldsymbol{y} \qquad \text{Eq. 1.4}$$

, where the weighting matrix $\boldsymbol{W}$ is

$$\boldsymbol{W} = \begin{pmatrix} w^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w^{(np)} \end{pmatrix}. \qquad \text{Eq. 1.5}$$

The enhancement of MLS lies on the "local" nature of the weighting matrix $\boldsymbol{W}$, since now it depends on the location of the point to be predicted and each observed data point. The weighting is controlled by a kernel function which decays with increasing distance $\left|\boldsymbol{x}^{(i)} - \boldsymbol{x}\right|$. An example of kernel function may be the Gaussian function:

$$w^{(i)} = \exp\left(-\frac{\sum_{j=1}^{k} \left(x_j^{(i)} - x_j \right)^2}{\sigma^2} \right) \qquad \text{Eq. 1.6}$$

Besides the kernel function, there are two more parameters which greatly affect the performance of the estimator. First of all, the order *m* of the local polynomials drives the *bias-variance trade-off*. In order to decrease the bias one can increase the value of *m*, though this might in turn increase the variability since more local parameters are used and vice versa.

The most critical parameter for the estimation result is the bandwidth, σ, since it controls the size of the local neighbourhood of the response function. The choice of σ is a trade-off between variance and bias. By choosing a large bandwidth the local estimate is influenced by many observations and thus the variance is small. On the other hand, the influence of remote observations might increase the squared bias. Many different techniques have been proposed in the literature for bandwidth selection, for instance, the conditional Mean Square Error (MSE) minimization in Fan and Gijbels [**9**] and the so-called "plug in method" based on Cross-Validation (CV) proposed by Ruppert *et al.* [**10**].

1.2.2.2 Radial basis functions (RBF):

As it is known, a generic radial basis function can be expressed in the form:

$$\psi\left(\| \boldsymbol{x} - \boldsymbol{\mu} \| \right) \qquad \text{Eq. 1.7}$$

, where $\boldsymbol{x}$ is an n-dimensional input vector, $\boldsymbol{\mu}$ is called centre, ||.|| denotes the Euclidean distance, and ψ is a univariate function, that is often referred to as the "profile function" or "basis function". Typically, a fitting model is set up as a linear combination of *N* radial basis functions having *N* distinct centers:

$$\hat{y}(\boldsymbol{x}) = \sum_{i=1}^{N} w_i \psi\left(\left\| \boldsymbol{x} - \boldsymbol{\mu}^{(i)} \right\| \right) = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\Psi} \qquad \text{Eq. 1.8}$$

, where w_j is the weight associated with the *j*-th radial basis function centred at $\boldsymbol{\mu}^{(j)}$. While Eq. 1.8 is linear in terms of the basis function weights w_j, the predictor $\hat{y}(\boldsymbol{x})$ can express highly non-linear responses. This is equivalent to

build a linear neural network having a number of inputs corresponding to the number of input vectors, primitive nodes whose transfer function is given by Eq. (1.6), and a single output (which corresponds to the values to be fitted) (Broomhead and Lowe [**11**], Hassoun [**12**]).

A variety of radial basis functions are used in practice [**13**] and lead to different techniques:

$$\psi(\boldsymbol{r}) = \|\boldsymbol{r}\| \qquad \text{(linear)}$$

$$\psi(\boldsymbol{r}) = \|\boldsymbol{r}\|^3 \qquad \text{(cubic)}$$

$$\psi(\boldsymbol{r}) = \|\boldsymbol{r}\|^2 \log\left(\|\boldsymbol{r}\|\right) \qquad \text{(thin plate spline)} \qquad \text{Eq. 1.9}$$

$$\psi(\boldsymbol{r}) = \sqrt{\|\boldsymbol{r}\|^2 + \sigma^2} \qquad \text{(multiquadratic)}$$

$$\psi(\boldsymbol{r}) = \exp\left(-\frac{\sum_{j=1}^{k} \theta_l \left|\boldsymbol{r}_j\right|^{p_l}}{2\sigma^2}\right) - \sqrt{\|\boldsymbol{r}\|^2 + \sigma^2} \qquad \text{(Kriging)}$$

, here σ is a positive constant. The basis function shown for Kriging is only one possibility, but is a popular choice that appeared in an influential article by Sacks *et al.* [**2**].

First, RBFs need the centres to be specified. At present, no general rule exists for selecting them [**14**], even though some criteria have been developed [**15**], such as the ROLS procedure (Regularized Orthogonal Least Squares), where the centres are chosen one at a time using a forward selection procedure from a candidate set consisting of all the data points or a subset thereof. However, a simple solution would be to choose $N = np$, thus leading to a square system of equations with a unique solution. The closure of the problem is achieved by imposing the centres of the basis directly on the data points. Thus will lead to the matrix equation:

$$\boldsymbol{\Psi w} = \boldsymbol{y} \qquad \text{Eq. 1.10}$$

,where $\boldsymbol{w}$ is the unknown parameters vector, $\boldsymbol{y}$ the observed response vector, $\boldsymbol{\Psi}$ denotes the so-called Gram matrix and it is defined as

$$\Psi_{i,j} = \psi\left(\left\|\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(j)}\right\|\right), \quad i, j = 1, ..., np \qquad \text{Eq. 1.11}$$

The importance of the Gram matrix in the computation of $\boldsymbol{w}$ will soon be clear. As $\boldsymbol{w}$ is the solution of Eq. 1.10

$$\boldsymbol{w} = \boldsymbol{\Psi}^{-1}\boldsymbol{y} \qquad \text{Eq. 1.12}$$

, the computational issues of $\boldsymbol{\Psi}^{-1}$ strictly depends on the mathematical properties of the Gram matrix $\boldsymbol{\Psi}$, and this is where the choice of basis function can have an important effect. It can be shown that, under certain assumptions, Gaussian and inverse multi-quadratic basis functions always lead to a symmetric positive definite Gram matrix (Vapnik [**16**]), ensuring safe computation of $\boldsymbol{w}$.

Beyond determining $\boldsymbol{w}$, some of the radial basis functions have an associated width parameter σ, which is related to the spread of the function around its centre. A heuristic approach is given in [**14**], where the width is the average over the centres of the distance of each centre to its nearest neighbor. However, this holds true for Gaussians RBFs, and it is only a rough guide that provides a starting value. Some algorithms exist for the width selection [**17**], including Generalized Cross-Validation GCV, but basically all of them proceed from a tentative value and test several widths values equally spaced between specified initial upper and lower bounds; then the width value minimizing log10(GCV) is selected.

Clarification must be made on the difference influence of $\boldsymbol{w}$ and basis function parameters on the predictor performance. A correct choice of the former will make sure that the approximation can reproduce the training data, while the correct estimation of the latter will enable us to minimize the estimated generalization error of the model.

1.2.2.3 RBF models of noisy data

When data is corrupted by noise, the interpolation scheme proposed in RBF ($N = np$) leads to deterioration of the predictor performance, because the model structure is not able to distinguish between the underlying response and the noise, thus observing overfitting behavior.

Adding more flexibility to the model could be an easy and practical way to deal with such problem. The insertion of a regularization parameter lambda within the Gram matrix implies that the approximation model will no longer pass through the training points and $\boldsymbol{w}$ will be the least-square solution of

$$\boldsymbol{w} = \left(\boldsymbol{\Psi} + \lambda \boldsymbol{I}\right)^{-1} \boldsymbol{y} \qquad \text{Eq. 1.13}$$

, where $\boldsymbol{I}$ is an (*npxnp*) identity matrix. The best choice for λ would be the variance of the noise in the response $\boldsymbol{y}$. Unfortunately, the noise variance is usually not known, thus lambda is added to the parameters to be estimated.

Another way to minimize undesirable overfitting is to construct a regression model reducing the number of bases functions $N < np$. To this end, Support Vector Regression (SVR) can give a very elegant and effective bases function selection method. An alternative is to use forward selection (Orr 1995 [**14**]) in which the basis function, which most reduces some error metric, are chosen and added one at a time in empty subsets of the design space; the process continues until there is no significant decrease in the error metric.

1.2.2.4 Ordinary Kriging

Kriging is a particularly famous method of creating metamodels as it can effectively represent a wide variety of responses while providing useful error estimate of the predictor. Derivation of Kriging equations can be carried out in different ways. The reader interested in the standard derivation may consult the article of Sacks *et al.* [**2**], while a "gentle introduction to Kriging" has been proposed by Jones [**18**]. Hereafter the main equations are reported trying to highlight the essential statistical concepts of the method.

In Kriging, the function to fit is seen as a random function $Y(\boldsymbol{x})$ whose realization, or response, is characterized by constant mean μ and variance σ^2. As a result, two contributions flow into the realization of $Y(\boldsymbol{x})$ at a given point $\boldsymbol{x}$:

$$Y(\boldsymbol{x}) = \mu + \varepsilon(\boldsymbol{x}) \qquad \text{Eq. 1.14}$$

. where $\varepsilon(\boldsymbol{x})$ is the deviation or error of the response due to the random attribute of $Y(\boldsymbol{x})$ and it is related somehow to σ^2. If a new point $\boldsymbol{x}^{(2)}$ is sampled near to $\boldsymbol{x}^{(1)}$, the realization of $Y(\boldsymbol{x}^{(2)})$ should be close to the previous response, because it is likely that the function under exam has some smoothness properties: as $\boldsymbol{x}^{(2)}$ approaches $\boldsymbol{x}^{(1)}$, the response $y^{(1)}$ tends to $y^{(2)}$. In other words, between $y^{(1)}$ and $y^{(2)}$ there is a certain degree of correlation that increases as the distance $\left|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(2)}\right|$ decreases; such correlation can be modeled statistically assuming that:

$$\boldsymbol{R} = Corr\left[Y(\boldsymbol{x}^{(i)}), Y(\boldsymbol{x}^{(j)})\right] = \exp\left(-\sum_{l=1}^{k} 10^{\theta_l} \left\| x_l^{(i)} - x_l^{(j)} \right\|^{p_l}\right) \quad \text{Eq. 1.15}$$

, where θ_l and p_l represent the hyperparameters of the l-th variable. The θ hyperparameter determines the rate at which the correlation decreases, and p determines the degree of smoothness in each coordinate direction. The hyperparameters are chosen via a maximization of the concentrated likelihood function (Jones [**18**]):

$$LH = -\frac{np}{2}\log\left(\hat{\sigma}^2\right) - \frac{1}{2}\log\left(|\boldsymbol{R}|\right) \quad \text{Eq. 1.16}$$

, where the optimal variance and mean are

$$\hat{\sigma}^2 = \frac{1}{n}\left(\boldsymbol{y} - \boldsymbol{I}\hat{\mu}\right)^T \boldsymbol{R}^{-1}\left(\boldsymbol{y} - \boldsymbol{I}\hat{\mu}\right) \quad \text{Eq. 1.17}$$

and

$$\hat{\mu} = \frac{\boldsymbol{I}^T\boldsymbol{R}^{-1}\boldsymbol{y}}{\boldsymbol{I}^T\boldsymbol{R}^{-1}\boldsymbol{I}}. \quad \text{Eq. 1.18}$$

Since the likelihood function gives a measure of how consistent a model is with respect to observed outcomes (i.e. training points), choosing the parameters to maximize LH intuitively means that we want our model parameters to be optimally tuned in such a way that model's behavior can well explain the data we have seen. This tuning is the main reason Kriging often outperforms other basis-function methods in terms of prediction accuracy.

Given a known set of hyperparameters, the Kriging prediction is the result of the maximization of the *augmented likelihood* function. The procedure is very similar to the LH maximization with the difference that the unknown is the new response (Jones [**18**]). The Kriging predictor can be written as:

$$y(\boldsymbol{x}^*) = \hat{\mu} + \boldsymbol{r}^T\boldsymbol{R}^T\left(\boldsymbol{y} - \boldsymbol{I}\hat{\mu}\right) \quad \text{Eq. 1.19}$$

, where here $\boldsymbol{r}$ denotes a vector of correlations between the unknown point $\boldsymbol{x}^*$ and the previous sample point $\boldsymbol{x}^{(i)}$.

One of the key features of Kriging is the provision of an estimated error in its prediction. At training points, the mean-squared error is null, while in all the rest of the variables space it can be calculated by:

$$s^2 = \hat{\sigma}^2 \left[1 - \boldsymbol{r}^T \boldsymbol{R}^{-1} \boldsymbol{r} + \frac{\left(1 - \boldsymbol{r}^T \boldsymbol{R}^{-1} \boldsymbol{r}\right)^2}{\boldsymbol{I}^T \boldsymbol{R}^{-1} \boldsymbol{I}} \right] \quad \text{Eq. 1.20}$$

1.2.2.5 Universal Kriging

A more general formulation of the model can be found in the Universal Kriging formulation (Cressie [**19**]), in which the mean term becomes some function of $\boldsymbol{x}$:

$$\hat{\mu}(\boldsymbol{x}) = \sum_{i=0}^{m} \mu_i \nu_i(\boldsymbol{x}) \quad \text{Eq. 1.21}$$

, where the ν_i's are some known functions and the μ_i's are unknown parameters. Usually $\hat{\mu}(\boldsymbol{x})$ takes the form of a low-order polynomial regression. The idea is that $\hat{\mu}(\boldsymbol{x})$ captures known trends in the data and bases functions added to this will fine-tune the model, thus giving better accuracy than ordinary Kriging where a constant $\hat{\mu}$ is used. However, if the underlying trend is not known a priori, the introduction of a low order polynomial could even deteriorate model prediction.

1.2.2.6 Bling Kriging

Blind Kriging (Joseph *et. al* [**20**]) gives an effective answer to the problem of polynomial selection emerged in universal Kriging. The approach proposed by Joseph *et. al* [**20**] aimed to identify the ν_i's through a Bayesian forward selection technique (Joseph [**21**]) and uses candidate variables of linear effects, quadratic effects, and two-factor interactions. Although the results reported by Joseph demonstrated the effective of such procedure, the computation cost of the bling Kriging process is much more computationally expensive and this may outweigh increased accuracy.

1.2.2.7 Co-Kriging

In the presence of multi fidelity data, the Co-Kriging technique (Cressie [**19**]) allows to enhance the accuracy of a surrogate of the expensive function by

coupling a greater quantity of cheap data with a small amount of expensive data. This can be done with a correction process that models the differences between the cheap and expensive function:

$$Z_e(\boldsymbol{x}) = \rho \, \mathrm{Z}_c(\boldsymbol{x}) + \mathrm{Z}_d(\boldsymbol{x}) \qquad \text{Eq. 1.22}$$

. where the expensive function Z_e is approximated by multiplying the cheap code Z_c by a scaling factor ρ plus a Gaussian process Z_d that accounts for the difference between Z_e and Z_c. Following maximization of likelihood function, predictor and error estimate of the predictor are calculated. The reader is referred to the work of Forrester *et al.* [**22**] for a detailed discussion on the subject.

1.2.2.8 Kriging with noisy data

In the same way as for an RBF prediction, a Kriging model may regress the data by adding a regularization constant to the diagonal of the correlation matrix (Forrester *et. al* [**23**]). Although the predictor is calculated similarly to ordinary Kriging, the predictor error estimate must be carefully evaluated in order to not include in such error both the model and the noise errors. Accounting only for the error model reflects the deterministic nature of the computer experiments and, consequently, gives to the error estimate a property of utmost importance for the optimization process, this is, to have zero error on sampling sites.

1.2.2.9 Support vector machines

The theory of Support Vector Machines (SVM) is mainly inspired from statistical learning theory of Vapnik [**16**]. Suitable SVM for interpolation and regression have been developed recently which are called Support Vector Regression (SVR) (Smola [**24**]). The main idea is to find a function with at most ε deviation from y and as the same time minimizes the model complexity (see Figure 1.1). Thus, the construction of the model reduced to the minimization of the following regularized ε-insensitive loss function:

$$L = \|\boldsymbol{w}\|^2 + C\frac{1}{np}\sum_{i=1}^{np}\max\left\{\left|y^{(i)} - \hat{y}(\boldsymbol{x}^{(i)})\right|, \varepsilon\right\} \qquad \text{Eq. 1.23}$$

, where ε is the accepted error, C is a regularization constant and $\hat{y}$ is the function to be estimated:

$$\hat{y}(\boldsymbol{x}) = \mu + \sum_{j=1}^{n} \boldsymbol{w}^{(j)} \psi(\boldsymbol{x}, \boldsymbol{x}^{(j)}) \qquad \text{Eq. 1.24}$$

The reader is referred to the work of Forrester and Keane [**25**] for a detailed derivation of SVR equations.

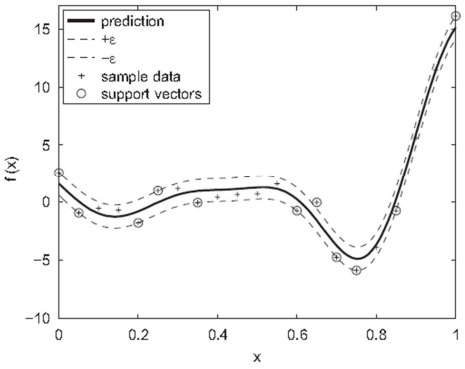


Figure 1.1: A SVR prediction using Gaussian kernel (from Forrester and Keane [**25**]).

1.2.3 Curse Of Dimensionality

When trying to map a high-dimensional input variables' space into an output space, a series of difficulties usually arise: specifically, multivariate data is difficult to work with because of the relevant amount of observations that are necessary to get good estimates. Furthermore, adding more features to the explanatory variables' space increases their interdependency relationships and can also cause an augmented noise, which may adversely affect prediction reliability. This is usually referred to as the "curse of dimensionality" (Bellmann [**26**]).

More generally, the curse of dimensionality is the expression of all phenomena that arise with high-dimensional data, and that have most often undesirable consequences on the behavior and performance of data fitting algorithms. Specifically, a nonparametric model building approach may show dramatically deteriorated prediction performance, unless it is fit with a proper number of independent observations [**27**]. Nevertheless, approximation deterioration seems to be inevitable and affects any meta-models as dimensionality of the problem on hand increases.

1.3 Metamodels At The Service Of Optimization

The use of metamodels in optimization framework has widely been adopted in the last few decades. There are no doubts that the introduction of approximation techniques have taken multiple advantages to the field and the huge amount of works out there give testimony of such. Here, the hard question is if there is an optimum way to integrate and exploit such techniques within the optimization framework. We would like to summarize the main ideas from which all the other works branch out, being aware that it could be considered an inexcusable reduction but necessary.

1.3.1 *Off-Line* And *In-Line* Approach

Generally, the creation of surrogates involves a training dataset over which the model is build or trained. We refer to *off-line* approach when processes of both training points sampling and model creation are performed before optimization starts. On the other hand, *in-line* approach denotes the situation where further training points are somehow selected and then added to dataset throughout optimization.

Several *off-line* data sampling methods have been proposed in the fields of design of experiment [**28**]- [**29**] as the famous Design Of Experiments (DOE), the Orthogonal arrays, Central Composite Design, and D-optimality. The final aim of these methods is to produce an "even-spread cloud" of sampling locations that pursues different optimal distribution concepts, which should encourage surrogates prediction capability. An *off-line* approach may be useful when the computational budget allows only few optimization cycles, or when the surrogate accuracy is so high that the optimizer can just be run over the metamodel.

When search space dimensions increase, a more efficient way of dealing with the poor approximation of surrogates is the *in-line* sampling strategy. Among the others, bagging [**30**] and boosting [**31**] are two statistical learning methods that have been developed to globally improve the quality of approximation using bootstrap techniques. Global quality is desirable but expensive, and is not strictly necessary for global optimization. In fact, we would like to improve surrogates only in those regions of interest that lead to the optimum, avoiding a detailed exploration of all the rest of the search space but, contextually, analyzing its main topology to be sure to evade local minima. Given a fixed computational budget, the solution to such problem is a trade-off between *exploration* and *exploitation* of the design space.

There are promising sequential sampling strategies that balance exploration and exploitation based on surrogate uncertainty prediction. For instance, Kriging techniques make use of statistical information to implement sampling rules, the so-called *infill criteria* [**18**]- [**32**]- [**33**]. The selection of new samples location may be driven by the Predicted Error (PE), the Probability of Improvement (PI), and the Expected Improvement (EI), or a variation on the theme. These concepts can be extended in a multi-optimization framework.

1.3.2 Multifidelity

The main hypothesis underneath the multifidelity approach is that high-fidelity models as CFD are more time consuming, require a larger computing effort, and are more accurate than a low-fidelity model, for instance an Euler inviscid solver or even a one-dimensional algebraic equation. On the other hand, low-fidelity models are cheaper but less accurate. How to mitigate the drawbacks related with computational cost of high-fidelity models with the use of low-fidelity models is the main concern of multifidelity approaches.

Many strategies have been developed over the years in order to link, or better to correct, somehow the low-fidelity model with the high-fidelity one. The main idea is that the link between the two can be analyzed and mathematically described by taken into account a low number of points evaluated for both high-fidelity and low-fidelity model according to a suitable design of experiment. Of course, this link can be refined in an adaptive way during optimization by evaluating new points on both models.

The ways metamodels could fulfill in such scope were generalized in the work of Toporov [**3**], in which metamodels are considered as tuned low-fidelity models based on the interaction of high- and low-fidelity models response:

$$\hat{y}(f(\boldsymbol{x}),\boldsymbol{a}) \approx F(\boldsymbol{x}) \qquad \text{Eq. 1.25}$$

, where $F(\boldsymbol{x})$ is the high-fidelity model response, $f(\boldsymbol{x})$ is the low-fidelity model response and $\boldsymbol{a}$ is a vector of tuning parameters used for minimizing the discrepancy between the high-fidelity and the low-fidelity responses at sampling points. Three different tuning approaches were proposed:

1. Linear and multiplicative with two tuning parameters
2. Correction function approach
3. Use of low-fidelity model inputs as tuning parameters

An alternative to Toporov's approach is the so-called *space mapping* that aims to establish a mapping of one model's parameter space on the other model's space such that the low-fidelity model with the mapped parameter accurately reflects the behavior of the high-fidelity model [**34**]. In other words, a distortion, or tuning, is performed on the input variable of the low-fidelity model rather than on its response; for instance, distortion could pursue the matching between the low- and –high fidelity model optimum [**35**].

1.3.3 Ensembles Of Surrogates

Another idea to enhance the prediction capability of a metamodel is to take into account various metamodels at the same time. In fact, since we do not now *a priori* which metamodel performs better, the use of multiple surrogates can reduce the risk associated with poorly fitted models [**36**]. This *a priori* uncertainty is inherent on how metamodels are built, which depends primarily on the combination of three components [**37**]: the statistical model and its assumption; the basis functions or kernel functions, depending on the specific surrogate; the minimization function or metrics selected to asses model parameters. The mix of these three ingredients could in principle lead to infinite recipes.

Bearing in mind how easy it is to produce different surrogates, being a variation of the same statistical model or based on different techniques, the number of metamodels to be created is more related to the computational budget and software capabilities rather than an "optimum magic number". The tough question is how to exploit information coming out from such multiple metamodels. In the literature there have been explored two main alternatives both based on criteria apt to evaluate the *goodness* of every surrogate at hand. Examples of such criteria are Cross-Validation, the Akaike Information Criterion and the Bayesian information criterion [**38**]. Regardless peculiarities of these criteria, they make it possible to rank surrogates. At this point, the first option is to promote as the best predictor the surrogate with the higher rank and rely on its response, discharging all the others; this is referred to as the selecting based approach. The other option is to create a "master" surrogate through proper weighting selection in the linear combination of the models, in which the weights reflect the goodness of surrogates. This second approach would like to mitigate the errors in prediction that could affect single metamodels.

Selection or combination? It seems that there is no ultimate answer as discussed by Yang [**39**]. He pointed out that selection can be better when the

errors in prediction are small, and combination performs better when the errors are large. Moreover, Viana *et al.* [**40**] highlighted that potential gains from using combination diminish drastically in high dimensional spaces and that in such circumstances criterion like cross-validation loses its effectiveness and makes the gain very difficult in practice.

Another aspect to be stressed is that from the optimization point of view it is not really important the prediction capability itself, but it is paramount the capability of reproducing the landscape of the fitness function. Samad itself in the optimization of the NASA Rotor 37 (Samada *et al.* [**41**]) had declared that the most accurate surrogate did not always lead to the best design, demonstrating that using combination of surrogates can improve the robustness of the optimization at a minimal computational cost.

1.3.4 Comparison Of Surrogates

Comparative studies have led to a no clear conclusion, despite the fact that nonparametric models seems superior to parametric ones (among others, Jin *et al.* [**42**] and Benini and Ponza [**43**]). Surrogates performance depends on both the nature of the problem and the DOE. Nevertheless, there are two interesting works related with turbomachinery, which indeed is the application field of the present work, that try to compare EA coupled with surrogates. First, Kean [**44**] addressed the optimization of a gas-turbine compressor blade section subject to damage in service and uncertainty in manufacture. Different optimization algorithms were compared, starting from the direct search with NSGA-II, then implementing a Kriging-based surrogate-assisted NSGA-II, concluding with Kriging EI formulation, in which statistical information coming out from Kriging model is used to drive the search towards the Pareto-optimal front. Although the second approach consumes only 30% of computational budget compared to NSGA-II, it was observed that further improvements on the Pareto front become difficult to find as the Kriging models are not as good at uncovering novel new behavior as the direct NSGA-II search, because they are always based on points from previous updates. On the other hand, expected improvement method carries out a more careful exploration of the design space, although slightly more slowly than surrogate-assisted NSGA-II run.

Second, Peter and Marcelet [**45**] compared different types of surrogate models (i.e., least square polynomials, radial basis function, multi-layer perceptron, simple Kriging) on a turbine cascade optimization problem, it turns

out that among all the models that were tested the Kriging models and the radial basis function network appear to give the best results. It is worth nothing that the cascade was parameterized with only two design variables, thus leading to a very low dimension design space. In order to formulate a precise ranking of surrogates, multiple evaluation criteria were considered, i.e., the ability to approximate the exact function and to find the global and local maxima. Starting from a DOE, performance was measured upon the number of exact CFD evaluations required for a certain surrogate in order to reach a predetermined threshold of the evaluation criteria.

1.4 Application Study: High-Speed Blade Optimization

It is well known that effectiveness of optimization algorithms strictly depends on the problem at hand; even though their performance may be evaluated and compared over synthetic functions, the harsh reality impose a real-world application as test bench to consecrate the promising algorithm to *superior realm*. For this reason, attention is here devoted to a specific optimization problem in the turbomachinery field, that is *blade shape optimization* and, in particular, high-speed or transonic blades optimization, which indeed is attractive for both the physic inside transonic blades and the complexity of the optimization landscape. Of course, the focus is on metamodels and on their contribution to the improvement of the optimization process efficiency.

To carry out any sort of shape optimization, there are two essential ingredients:

- geometric parameterization
- optimization strategy

A brief review on high-speed blade optimization works is hereafter presented with the aim to highlight these two aspects rather than the physic explanation of improvement achieved optimizing the blade. Although most of the woks are focused on NASA Rotor 37 and Rotor 67 (see Figure 5.1), attention will be first devoted to Rotor 37 (see Table 1.1) and others interesting minor works, while there is a specific section about the NASA Rotor 67 in chapter §0.

	Authors	Parameterization		Opt. Strategy		
		Locus	Design variables	Method	Objectives	CFD runs
No Surrogates	Benini [**46**]	Lean+Foils	23	GA+CFD	η_{iso} π_{c0} π_{c0}	2000
	Burguburu and Pape [**47**]	Foils suction side	9	Gradient+CFD	η_{iso}	41
Off-line	Jang *et al* [**48**].	Sweep	2	RSM+CFD	η_{iso}	7
	Ahn and Kim [**49**]	Lean	3	RSM+CFD	η_{iso}	15
	Chen *et al.* [**50**]	Foils	3	RSM+CFD	η_{iso}	-
In-line	Wang *et al.* [**51**]	Sweep+Lean	4	GA+BPNN+CFD	η_{iso} π_{c0}	232

Table 1.1: Major works related with NASA Rotor 37 shape optimization.

1.4.1 Geometric Parameterization Of The Blade

Generally, a blade can be reconstructed interpolating a few spanwise 2D sections located along the 3D stacking line. As a result, blade reshaping may involve the 3D stacking line (referred to as *sweep* and *lean* deformation, see §5.2 for a detailed discussion), 2D spanwise profiles, or both of them. The parameterization may have different levels of complexity according to (i) how in depth it can reshape the geometry and (ii) how many design variables are associated to the parameterization framework. Usually, control points (CP) based curves as Bézier or B-splines or cubic splines are adopted because CP are straightforward related with the optimization design variables, and also for their ability to produce smoothed and contextually complex shapes.

Let's consider the works on the NASA Rotor 37. In Jang *et al.* [**48**] the stacking line was parameterized with two design variables allowing only sweep deformation. Ahn and Kim [**49**] considered only lean deformation handled by three design variables. For the same rotor, parameterization complexity was increased including both lean and sweep but leaving the profiles shape unchanged (Wang *et al.* (2011) [**51**]), requiring four design variables in the whole.

On the other hand, Burguburu and Pape [**47**] adopted a pure 2D profile deformation; modified suction sides were derived applying displacements to the reference blade by translating each point along the blade local normal taking into account only the suction side of the foils. In Chen *et. al* [**50**]the profiles

deformation was focused on design parameters (or geometric parameters) with strong physical meaning, i.e., position and value of maximum thickness, stagger angle, leading and trailing edge angles, etc. The modifiable spanwise sections were reduced to 3 (i.e., hub mid-span and tip) and only 3 design variables per section were activated (i.e., stagger angle, maximum camber location and maximum thickness location), being the rest fixed to the baseline value.

A whole parameterization of Rotor 37 was proposed by Benini [**46**], in which three profiles along span were selected (i.e., hub, midspan, and tip profiles), each of which was represented by camber and thickness distributions. These were defined by fourth-order Bezier polynomials. 14 parameters for the camber lines plus 9 parameters for the thickness, this is, 23 parameters in total were used.

1.4.2 Optimization Strategies

First, we make a distinction between direct and inverse methods based on the nature of the input variables of the optimization problem. In the direct methods, inputs are mainly geometric variables that affect the shape of the blade; the algorithm searches for new geometries for which objectives are optimized. Contrarily, in the inverse method, conventional design quantities are the distribution of pressure, or generally, of any fluid dynamic property on the surface of the blade; the aim of the process is to find those distributions that optimize the objectives. Of course, the way in which the blade geometry is computed on the basis of the specification of quantities distribution is the core of inverse methodologies, which however are beyond the scope of this brief survey and will not be investigated further; exception is made for Rotor 67.

1.4.2.1 Avoiding surrogates

Most of the approaches encountered in the literature belong to the direct method category. Some of these did not implement metamodels. For instance, in Benini [**46**] the Rotor 37 multi-objective optimization was performed by means of GeDEA genetic algorithm [**52**]. Isentropic efficiency and total pressure ratio were to be maximized at a given mass flow rate. A population of 20 individuals evolved for a total of 100 generations, against 23 design variables. An improvement of 1.5% in the adiabatic efficiency was achieved without modifying the total pressure ratio (particular point on Pareto-optimal front). In Burguburu and Pape [**47**] the Rotor 37 was optimized by means of a gradient method coupled with CFD solver. Isentropic efficiency at a given design point was maximized,

keeping mass flow and total pressure ratio fixed. Only three iterations with the gradient method were performed and 41 CFD calls were required, against 9 design variables. The increment in isentropic efficiency was around 1.2% compared to the reference blade.

1.4.2.2 Off-line strategies

The simplest strategies are the one based on the *off-line* approaches (§1.3.1). In Jang *et al.* [**48**]a polynomial response surfaces was built with seven points evaluated via CFD, against two design variables; then, the maximum isentropic efficiency was searched on the metamodel. It was found that the optimum shape was a backward sweep deformation, which increased the isentropic efficiency by 1.25%. In Ahn and Kim [**49**] the isentropic efficiency was maximized adopting a response surface optimization approach. The 15 sample points (against 3 design variables) prescribed by full factorial design were selected using D-optimal design strategy. Then, a polynomial response surface was constructed and a search algorithm was run over the surrogate to find the global optimum. The optimization of the rotor blade produced an efficiency enhancement of 0.7%. In Chen *et al.* [**50**] isentropic efficiency was maximized by means of a gradient algorithm applied to response surface technique. The isentropic efficiency was increased by 1.73% compared to the baseline Rotor 37.

1.4.2.3 In-line strategies

More elaborated strategies belong to the *in-line* approaches. In Wang *et al.* [**51**] a multi-objective optimization framework using NSGA-II and back propagation neural network was applied to redesign the NASA Rotor 37. In this framework, a modified crowding distance was proposed to enhance the robustness of NSGA-II and a course-to-fine approaching strategy was implemented to refine the approximation model, keeping to a minimum the expensive CFD evaluations. Efficiency increased about 1.1% and total pressure ratio increased about 1%, while the chocking mass flow only decreased 0.04%.

Other works are not strictly related with high-speed turbomachinery, but they make use of surrogates to speed up the optimization process. Karakasis *et al.* [**53**] have used a hierarchical evolutionary algorithm based on multi-fidelity models, whose accuracy and computational cost increase from the lowest to the highest level. The role of the lower levels is to explore the design space with the minimum computational effort and guide the higher ones to scrutinize particular regions by modeling additional flow features, which cannot be described by the

lower-level tool. This approach was applied to a controlled diffusion compressor cascade optimization aiming to minimize the pressure losses while preserving the flow turning. The hierarchical approach comprises two levels, a viscous-inviscid flow solver and Navier-Stokes equation solver, both approximated with radial-basis function networks. Compared to conventional EA approach, it was demonstrated that 8-fold time benefit which is mainly obtained from the approximation of the expensive flow analysis by using meta-model approach. In Okus and Akmandor [**54**] a novel multilevel genetic algorithm was presented, which takes advantage of the successful solutions adopted in dynamic environments, this is, "multiploid" GAs, in which the genetic operators are rewritten in such a way that the multi-fidelity information can be treated and exploited to enhance the search to converge toward the Pareto-optimal front. A 3D blade turbine, which was parameterize by 37 design variables, was optimized in order to maximize isentropic efficiency and torque. Compared to a simple MOGA (Multi Objective Genetic Algorithm) approach, this method reduced the computational cost by a factor of 4, while producing a superior Pareto-optimal frontier with respect to MOGA. In Keskin *et al.* [**55**] the multi-objective NSGA-II genetic algorithm assisted by adaptive Kriging-based response surfaces was applied to the optimization of a 3D compressor blade in order to minimize loss production at a required flow turning, considering both design and off-design performance. Surrogate's accuracy was automatic controlled throughout the entire optimization process by updating the CFD training points when strictly needed, in order for the genetic algorithm to find the Pareto-optimal front on a reliable model response. In Giannakoglou *et al.* [**56**], a multi-layers network, which can be trained on both known responses and response gradients, operates as approximate evaluation tool during the evolutionary search. This novel implementation was used to design 3D blade of both turbine and compressor, the former being parameterized with 33 Bézier control points while the latter with 20 points. Furthermore, the 3D Euler and adjoint equations were solved to provide the objective function gradient, which indeed speeds up the optimization algorithm. Compared to traditional EA, the novel approach drastically reduced the number of exact evaluations by a factor of 5. In fact, only 200 evaluations were needed to reach the target pressure distribution, which was the optimization goal. Contrarily, a traditional EA after 1000 evaluations was not capable of reproduce comparable results.

1.5 Thesis Objectives

This thesis deals with turbomachinery optimization and all those tools employed in the optimization process, mainly the optimization algorithm, the parameterization framework, and the automatic CFD-based optimization loop. The first and most speculative objective is the enhancement of the optimization algorithm by means of surrogate models. The state-of-the-art genetic algorithm GeDEA-II, which was developed at Università di Padova, needs to be equipped with "metamodels' technologies" to improve its convergency efficiency and repeatability, while preserving its robustness.

The second objective is related with the design and development of the automatic CFD-based optimization loop built around the improved GeDEA-II algorithm. This comprises a robust strategy to handle all the optimization tools in a Linux cluster environment in order to exploit the computational resources of parallel computing. Among the tools, the most important one is the parameterization tool able to reshape both 2D supersonic foils and 3D transonic compressor blades. Once the analysis tools are ready, the optimization of high-speed turbomachinery may start. The third and last objective of this thesis is two-fold: prove the effectiveness of the optimization approach and gain insight on the physics phenomena of transonic and supersonic flows with the aim to explain the reason of the observed improvements.

1.6 Thesis Outline

The document is organized in five chapters. In Chapter 2 the surrogate-assisted multi-objective genetic algorithm GeDEA-II-K is presented. Based on the cooperation between the GeDEA-II genetic algorithm and the Kriging technique, GeDEA-II-K is tested over two- and three-objective synthetic test functions proving to be a promising tool in a multi-objective optimization context.

The effectiveness of a CFD-based automatic loop developed during this PhD course is verified on two real-life multi-objective optimization problems: the 2D shape optimization of a supersonic compressor cascade and the 3D shape optimization of the NASA Rotor 67. Due to the inherent misleading behavior of supersonic cascades compared with subsonic ones, Chapter 3 is entirely devoted to the physic of supersonic compressor cascades. The reader must be aware of the *physics constraints* of a supersonic flow throughout compressor cascades in order to gain a better understanding of the results coming out from the cascade shape

optimization. In chapter 4, the study and the optimization of a supersonic compressor cascade is presented. In particular, the validation of the CFD model, the parameterization technique and the *unique incidence* control loop are illustrated, in addition to the optimization results. In the last chapter, the optimization of the NASAS Rotor 67 transonic compressor is reported.

Although a lot of time was invested to design and perfect the automatic optimization framework, the loop is briefly reported in Appendix A, because the relevant peculiarities were not judged to be in such amount to require a separate chapter.

Chapter 2

GeDEA-II-K: A Kriging-Assisted Evolutionary Algorithm

In this chapter, a novel surrogate-assisted (SA) evolutionary algorithm for MOOPs developed during my PhD course is presented. The GeDEA-II-K is grounded on the cooperation between the GeDEA-II, which is a state-of-the-art "pure" genetic algorithm, and some Kriging statistical criteria featured in the ASEMOO, which is a surrogate-assisted algorithm based on the Kriging technique. How far can a SA algorithm go when very few direct evaluations are available? Comparison over two- and three-objective test functions have demonstrated that the GeDEA-II-K, exploiting synergistically the strengths of both parents the GeDEA-II and the ASEMOO, can achieve high performance in the approximation of the Pareto-optimal front mitigating the drawbacks of its "parents".

2.1 Introduction

Evolutionary Algorithms (EAs) play an important role in the framework of metaheuristics in dealing with multi-objective problems in real-world engineering optimization. Research in this field is primarily concentrated toward reducing the computational effort for obtaining multiple optima. At the same time, quality and variety of optimal solutions is of fundamental importance to engineers in order to give them a number of choices among which to select the most appropriate ones with a high level of confidence regarding their performance. The latter can be referred to as "convergence ability". A complete review and synthesis on metaheuristics can be found in [**57**].

Most of the times, computational effort and convergence ability are conflicting tasks: the lower the former, the lower the quality of the obtained

solutions. Generally speaking, in non-deterministic algorithms found in metaheuristics, both the computational effort and convergence ability depend on the absolute number of direct function evaluations. While "exact" solutions can be obtained using a direct objective function evaluation only, generation of "non-exact" or approximate solutions can be obtained using a response surface (often referred to as a "metamodel" or "surrogate"), which mimics the real objective function landscape being computationally cheaper to evaluate [**58**].

In the following, we shall refer to a "pure" method (PM) when dealing with an algorithm performing direct function evaluation only, and to a "surrogate-assisted" method (SA) when considering procedures that make use of approximated function landscapes somewhere in the optimization method.

The purpose of the present work is to show how cooperation between particular type of PM and SA methods, namely the GeDEA-II and the ASEMOO algorithms, have led to a novel algorithm: GeDEA-II-K. More in detail, two- and three-objective test functions are selected among up-to-date multidimensional problems in the literature that stress the search algorithms hampering convergence towards to the Pareto-optimal front. Performance of the algorithms is compared using a universally accepted measure of fitness function evaluation cost, The Adimensional Direct Evaluations Number (ADEN). Furthermore, two metrics of performance were used to analyze algorithms' results, i.e. the Hyper-volume (HV) and the so-called D-metric.

2.2 Brief Review Of GeDEA-II And ASEMOO Algorithms

2.2.1 GeDEA-II

The GeDEA-II algorithm is a multi-objective real-coded evolutionary algorithm (MOEA) developed at University of Padua. It mainly follows the basic steps of an Evolution Strategy implementing a Pareto-like evaluation method based on both fitness and distance among individuals [**52**].

A Simplex-Crossover (SPX) operator is implemented for individuals' recombination, while Tournament-Selection (TS) and Shrink-Mutation (SM) are employed to complete the genetic manipulation [**59**]. While keeping firmly in mind that the No free lunch theorem [**60**] applies to optimization algorithms, the GeDEA-II is able to outperform some state-of-the-art competitor algorithms on a number of state-of-the-art test problems. Following is a brief description of the

GeDEA-II operators, while for an extensive explanation on its architecture and performance the reader is referred to [**59**].

The main steps of the GeDEA-II algorithm are described as follows:

Step 1: An initial population of μ individuals is generated at random.

Step 2: A mating pool of 2λ individuals is formed, each individual having the same probability of being selected using TS.

Step 3: λ offspring are generated by SPX crossover. SM mutation is applied randomly with a probability *pmut*.

Step 4: The whole population of $\mu+\lambda$ individuals is checked to discover possible clones. These clones are removed and replaced with new randomly generated individuals.

Step 5: The objective function values of the $\mu+\lambda$ individuals are evaluated and the non-dominated sorting procedure by Goldberg is performed to assign the ranks to the solutions according to the objectives of the MOOP.

Step 6: The whole population of $\mu+\lambda$ individuals is processed to determine the value of the reciprocal distance-based genetic diversity measure for each individual.

Step 7: GeDEM [**52**], a special sort of as a genetic diversity preservation method, is applied according to the ranks scored in Step 5 and the values of the diversity measure assigned in Step 6. The non-dominated sorting procedure is used again to assign the ranks. GeDEM computes the actual ranks of the solutions maximizing (i) the ranks scored with respect to the objectives of the original MOOP, the non-dominated solutions having the highest rank, and (ii) the values assigned to each individual as a measure of its genetic diversity, calculated according to the chosen distance metric, i.e. the (normalized) Euclidean distance in the objective functions space.

Step 8: The best μ solutions among parents and offspring, according to the ranks assigned in Step 7 are selected for survival and the remaining λ are eliminated.

Step 9: If the maximum number of generations is reached then stop, else go to Step 2.

2.2.1.1 THE SPX Operator

Figure 2.1 depicts the pseudo-code of an SPX in a multi-objective context, extended to the most general case involving *M* objective functions. It is assumed that all of the objectives are to be minimized. At each generation *ignr*, the mean of

each objective function *mean* is calculated. Based on these values, the percentage variations *PV* are subsequently derived. Next, two selected parents are sorted according to these values and the child created. This choice guarantees that the objective function characterized by the greatest variation is selected every time, therefore ensuring the highest convergence rate to the Pareto Front (PF). For test problems involving more than two objective functions, the objective function considered to form the new child is chosen randomly in order to enhance the design space exploration of the crossover operator required in highly dimensional objective spaces.

Algorithm 1 Application of SPX in a multi objective context.

1: $Set\ M = number\ of\ objectives$
2: $Set\ \mu = number\ of\ parents$
3: $Set\ ignr = current\ generation$
4: **for** $i = 1 \rightarrow M$ **do**
5: $Mean(i) = 0$
6: **for** $j = 1 \rightarrow \mu$ **do**
7: $MEAN(i)_{(ignr)} = Mean(i)_{(ignr)} + \sqrt{oldfit(j)^2}$
8: **end for**
9: **end for**
10: **for** $i = 1 \rightarrow M$ **do**
11: $PV_i = \frac{MEAN(i)_{(ignr-1)} - MEAN(i)_{(ignr)}}{MEAN(i)_{(ignr-1)}}$
12: **end for**
13: $Set\ count = 1$
14: **while** $count \leq = \mu$ **do**
15: $Choose\ two\ parents,$ **p1** and **p2**, $(Selection\ operator)$
16: **for** $i = 1 \rightarrow M$ **do**
17: $oldfit(i,1) = i_{th}\ objective\ function\ of\ parent$ **p1**
18: $oldfit(i,2) = i_{th}\ objective\ function\ of\ parent$ **p2**
19: **end for**
20: $Set\ A = max_{i \in M}\ (PV_i)$
21: $Find\ index\ k \in M\ correspondent\ to\ A$
22: $Set\ OLDFIT = [oldfit(k,1), oldfit(k,2)]$
23: **if** $oldfit(k,1) \leq oldfit(k,2)$ **then**
24: $Set\ M = \left(\frac{1}{n}\right) \cdot$ **p1**
25: **Child** $= (1 + Refl) \cdot M - Refl \cdot$ **p2**
26: **else**
27: $Set\ M = \left(\frac{1}{n}\right) \cdot$ **p2**
28: **Child** $= (1 + Refl) \cdot M - Refl \cdot$ **p1**
29: **end if**
30: $Set\ offpsring(count, :) =$ **Child**
31: $count = count + 1$
32: **end while**

Figure 2.1: Pseudo-code for SPX operator.

2.2.1.2 The SM Operator

The mutation operator adds a random number taken from a Gaussian distribution with mean equal to the original value of each decision variable characterizing the entry parent vector. The shrinking schedule employed is:

$$Shrink_i = Shrink_{i-1} \cdot \left(1 - \frac{ignr}{ngnr}\right) \quad \text{Eq. 2.1}$$

where $Shrink_i$ is a vector representing the current mutation range allowed for that particular design variable, *ignr* represents the current generation and *ngnr* the total number of generations. The shape of the shrinking curve was decided after several experimental tests. Once the current variation range has been calculated, one decision variable of the mutated child is randomly selected and mutated according to the following formula:

$$Child_{mut} := Child_{cross} + \left[\left(\sqrt{Shrink_i} \cdot random\right) \cdot Shrink_i\right] \quad \text{Eq. 2.2}$$

where $Child_{mut}$ is the mutated decision variable, $Child_{cross}$is the decision variable generated by SPX and *random* is a random number taken from a normal distribution in the open interval]-1,1[.

2.2.2 ASEMOO

The ASEMOO (ASynchronous Efficient Multi Objective Optimization) algorithm is an optimization algorithm developed at the Warsaw University of Technology. Its concept is based on EGO algorithm by Jones *et al.* [**61**] and works of Jeong *et al.* [**62**]. The main optimization loop consists of the following steps:

1. A database of designs is initialized with a starting set of points selected with Latin Hypercube Sampling.
2. Objective functions are evaluated in all the points in the database.
3. Kriging model for all the objective functions is created.
4. Multi-objective optimization of a sampling criterion (EHVI) is performed. One point is selected and added to the database.
5. If more points are needed, go to algorithm goes to 2.

Kriging is a statistical model used for multi-dimensional approximation. A given objective function f is considered a realization of a random field $Y(\boldsymbol{x})$ with prescribed mean μ(x) and covariance $\boldsymbol{R}(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)})$ functions. Function f is then approximated with an unbiased, linear, least squares estimator of Y. In case

of ASEMOO algorithm, linear mean function is used and correlation function is given by equation:

$$\boldsymbol{R}(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}) = \exp\left(-\sum_{l=1}^{k} \left| \frac{x_l^{(i)} - x_l^{(j)}}{\theta_l} \right|^{p_l} \right) \quad \text{Eq. 2.3}$$

where p_l is 1 or 2, depending on the smoothness of the objective function. Shape parameters θ_l are chosen with Maximal Likelihood method and verified with generalized cross-validation (GVC). This approach for finding shape parameters is based on the Kriging implementation by Roustant *et al* [**63**]. One of the strong features of Kriging model is the ability to not only calculate the approximation, but also the variance of the estimator. This variance can be used as a good estimate of the error of the approximation. Jones *et al.* [**61**] combined these two information to create a sampling criterion which balances two, most important goals of the optimization loop:

1. exploration – improvement of the approximation
2. exploitation – finding the exact optimum

The sampling criterion is based on the assumption that the error of approximation has a normal distribution and we can calculate expected value of the improvement of the objective (EI):

$$EI\left(\boldsymbol{x}\right) = \boldsymbol{E}\left(max\left(f_{min} - Y\left(\boldsymbol{x}\right), 0\right) \middle| Y(\boldsymbol{x}^{(i)}) = f(\boldsymbol{x}^{(i)})\right). \quad \text{Eq. 2.4}$$

The $\boldsymbol{x}^{(i)}$ are the points where the objective function was already evaluated and f_{min} is the minimum objective function in the evaluated set. This function is high in two types of places: where the approximation of the objective has lower value and where the quality of approximation is low. Many approaches were considered for extending this criterion to multi-objective optimization problems, some of which were investigated by Łaniewski-Wołłk in [**64**]. The study found that Expected Hyper-volume function can achieve very good and consistent optimization results. The *Expected Hyper-Volume Improvement* (EHVI) function is constructed as:

$$EHVI\left(\boldsymbol{x}\right) = \boldsymbol{E}\left(HV\left(Y\left(\boldsymbol{x}^{(1)}\right),\ldots,Y\left(\boldsymbol{x}^{(n)}\right),Y(\boldsymbol{x})\right) - HV\left(Y\left(\boldsymbol{x}^{(1)}\right),\ldots,Y\left(\boldsymbol{x}^{(np)}\right)\right)\middle| Y\left(\boldsymbol{x}^{(i)}\right) = f(\boldsymbol{x}^{(i)})\right), \quad \text{Eq. 2.5}$$

where Y is a vector-valued objective function and $HV\left(\boldsymbol{y}^{(1)},\boldsymbol{y}^{(2)},\ldots,\boldsymbol{y}^{(np)}\right)$ is hyper-volume of the set dominated by the points $\boldsymbol{y}^{(1)},\boldsymbol{y}^{(2)},\ldots,\boldsymbol{y}^{(np)}$. In any hyper-volume calculation, there has to be a reference point selected. This point represents the maximum accepted values of the objective. The ASEMOO algorithm will find the part of the Pareto front that dominates this point. This feature of the algorithm can be used to refine a specific part of the Pareto front or to prevent deterioration of any objective with respect to the base (starting) design. The value of EHVI can be calculated analytically if we assume that the errors of the approximation of all the objectives are independent and have normal distribution. It is interesting to note that EHVI criterion reduces to Expected Improvement when applied to a one objective problem. Also like EI the EHVI criterion gives a good balance between exploration and exploitation. As it is based on the dominated-hyper-volume function it gives a good even, spread of points on the Pareto front.

Function EHVI is highly multi-modal and would be very hard to optimize with common algorithms. The final sampling criterion used in ASEMOO algorithm is a set of objectives, from which the first is -EHVI(**x**) and the rest are the approximates of all the objective functions. This makes the optimization problem well suited for common multi-objective genetic algorithms. The evaluation of the sampling criterion is very cheap compared to the evaluation of the objective, so the criterion in ASEMOO is optimized with NSGAII algorithm by Deb *et al.* [**65**] with a high number of generations and high population size.

At this stage, one point with the highest EHVI is selected from the Pareto front of the multi-objective optimization criterion. Objective functions are evaluated in this point, the data is added to the database and the optimization loop continues.

ASEMOO algorithm is also capable of generating new designs before the last one was evaluated, opening the possibility of asynchronous evaluation of the objective functions for different points in parallel. This feature is essential in industrial applications, where several evaluations can be run simultaneously and their running time (e.g., convergence of CFD solver) can substantially vary for

different designs. The same feature is used for designs for which the evaluation failed (e.g., divergent solution or failure of mesh generation).

2.3 GeDEA-II-K: Cooperation Between GeDEA-II And Kriging

How could we exploit the strengths of both GeDEA-II and ASEMOO avoiding destructive interactions between the two? GeDEA-II has proven to have outstanding exploration capability and diversity preservation [**59**] compared to EAs competitors, but still low convergence efficiency: it requires a quite high number of direct evaluations to cover the Pareto-optimal front. Do not blame it; it is inherent in its "genetic nature". On the other hand, ASEMOO has demonstrated to be an effective solution to *exploration-exploitation trade-off* with an elegant and sophisticated mathematical approach [**64**]. However, as it will be clear from the result section, if the Kriging does not get an appropriate insight on the fitness functions landscape the performance rapidly deteriorates. This is because the EHVI sampling criterion, which is adopted in ASEMOO to select new sampling points, relies on the effectiveness of the metamodel approximation. If the approximation is poor, it is high probable that new proposed sampling points will give little contribution to the advancement towards the Pareto-optimal front, and also to the Kriging model improvement. Of course, the balance exploitation-exploration inherent in EHVI will intervene sooner or later. However, in practice, a conspicuous number of direct evaluations, which depends on the complexity of the fitness function at hand, are needed to come out from a deceptive function landscape. In conclusion, a robust algorithm cannot rely on the metamodel only.

2.3.1 *Περί μαντικῆς*

The main idea behind GeDEA-II-K is to improve the GeDEA-II's reproduction operator with the integration of a Kriging filter. In particular, GeDEA-II-K doggedly takes advantage of GeDEA-II's exploration capability, which indeed is used to create a more densely populated offspring compared to the GeDEA-II one, while contextually adopts a Kriging based criterion to *filter* which individuals among the GeDEA-II's offspring may proceed to direct evaluation. Here the filter drives the search towards those regions promising from the Kriging point of view, without however having control on the regions themselves, since the offspring is suggested by GeDEA-II. The effect is twofold:

1. generation after generation, the Kriging database is augmented with offspring individuals that have a genotype inspired by GeDEA-II and a phenotype by Kriging: "*GeDEA-II in body and Kriging in soul!!*". Since the proposed offspring is potentially promising from the GeDEA-II point of view, the filtered offspring could lift a poor Kriging, even though the filter action is poor.
2. GeDEA-II reproduction operator (i.e., SPX+SM operators) receives *auxilium* from Kriging (i) by the inferred information on the fitness functions landscape and (ii) by retention of all the previous populations, which guarantees to do not select an offspring that is clone, or very close to an individual, of preceding populations.

Besides the cooperation GeDEA-II Kriging by means of the filter, it is reasonable to introduce in the offspring population individuals promoted by Kriging, for instance using the EHVI criterion. As suggested by Łaniewski-Wołłk [**64**], the best performance with the EHVI criterion is obtained with a *single-sampling strategy*, in which the selected point is evaluated and the Kriging model is updated before a new point is sampled. In the GeDEA-II-K context, only one point is selected based on EHVI.

2.3.2 The Framework

The GeDEA-II-K shares the same framework of the GeDEA-II (see §2.2.1). Starting from the first generation, the individuals evaluated so far are stored in the *strpop* and *strfit* variables. Step 3 and 4 of GeDEA-II outline are replaced with the following:

Step 3: 4λ offspring are generated by SPX crossover. SM mutation is applied randomly with a probability *pmut*.

Step 4a: Kriging models, one for each fitness function, are constructed over the database [*strpop*,*strfit*]. One point is selected according to the EHVI criterion.

Step 4b: the Kriging filter is applied to the whole population of $4\lambda+1$ individuals, after which only λ individuals are selected. Inside the Kriging filter routine an appropriate *clone-extermination* is accomplished.

2.3.3 The Kriging Filter

The Kriging filter (-K filter) is in charge to select λ individuals among $4\lambda+1$ and to assure a certain level of minimum distance in the design space among all individuals evaluated so far; clones are avoided since are just a particular case having null distance. The pseudo-code of the –K filter is reported in Table 2.1. It can be divided in three main chunks: the Kriging model creation, the minimum distance rejection, and the Pareto front selection.

First, one Kriging model for each objective is built upon the dataset storing all individuals so far evaluated, i.e. *strpop* and *strfit* variables; then, the optimum point according to the EHVI criterion *offEHVI* is selected. The Kriging models creation is the most expensive part of the filter algorithm.

Second, a minimum Euclidean distance in the genotype space is required. It is worth nothing that the distance among all individuals that flow into the Kriging database is of utmost importance due to the inversion of the Gram matrix, which becomes singular for distance getting close to zero. The *offpop* candidates too close to both the *strpop* and the *offEHVI* are deleted. Then, the *offpop* candidates too close to each other are rejected. This operation requires an iterative procedure, because deleting one element may change the minimum distance of other elements of the set that were paired with the deleted one. When the minimum distance law is broken, there are at least a couple of individuals that shares the same minimum distance (i.e., *popx*); only one individual will be going to the next phase. In order to decide which one should be preserved, the genotype diversity is rewarded: the individual that has the largest minimum distance with the set enclosing both the *strpop* and the *offEHVI* is retained; the others, one or more, are rejected.

In the third part, the fitness functions of the *offpop* set are predicted by means of the previous Kriging models and the non-dominated sorting procedure by Goldberg (1989) [**66**] is performed to assign the ranks to the solutions according to the predicted fitness functions. *Offpop* individuals are sorted by ranks and the first λ individuals are selected. If there are more than λ individuals on the Pareto front, the genotype diversity is rewarded.

Kriging Filter Algorithm

1: Define *Dmin_1(offpop)* := minimum genotype Euclidian distance between each individual in *offpop* and all the remaining individuals in *offpop*
2: Define *Dmin_2(offpop1,offpop2)*:= minimum genotype Euclidian distance between each individual in *offpop1* and individuals in *offpop2*
3: Set *strpop = population evaluated so far from beginning*
Set *d_toll = 1-e03*
4: Set *M = number of bjectives*
5: Set λ *= number of offspring required*
6: Set *strfit = population fitness evaluated so far from beginning*
7: **for** *i=1 to M*
8: **Build** *Kriging model on database (strpop, strfit(:,i))*
9: **end for**
10: Set *offEHVI = individual selected by EHVI criterion*
11: Set *dmin2=Dmin_2(offpop,[strpop; offEHVI])*
12: Find *index vector i for dmin2(i)<d_toll*
13: **Delete** *offpop(i,:)*
14: Set *dmin=Dmin_1(offpop)*
15: **while** *dmin<d_toll* **do**
16: find *index vector i for dmin(i)<d_toll*
17: Set *popx = offpop(i,:)*
18: Set *dmin2 = Dmin_2(popx ,strpop)*
19: Set *dmin2_sort = sort dmin2 in ascending order; index vector j:= dmin2_sort=dmin2(j)*
20: **Delete** *offpop(i(j[1:end-1]),:)*
21: Set *dmin=Dmin_1(offpop)*
22: **end while**
23: **for** *i=1 to M*
24: Set *offfit(:,i) = prediction Kriging model of offpop*
25: **end for**
26: Set *rankPF = Pareto Ranking of offfit*
27: Set *rankPF_sort = sort rankPF in ascending order; index vector j:= rankPF_sort=rankPF(j)*
28: Set *n_PF = number of individuals on the Pareto front*
29: **if** *size offfit* $> \lambda - 1$ *AND n_PF* $> \lambda - 1$
30: *Set dmin2 = Dmin_2(offpop ,strpop)*
31: *Set dmin2_sort = sort dmin2 in descending order; index vector k:= dmin2_sort=dmin2(k)*
32: *Set vector index isel = j(k(1:*$\lambda - 1$*))*
33: **else**
34: Set *vector index isel = j(1:*$\lambda - 1$*)*
35: **end if**
36: *Set selpop = [offpop(isel,:); offEHVI]*

Table 2.1: Kriging filter aglorithm.

2.4 Some Inference On Multidimensional Test Functions

2.4.1 Test Problem Suite

Two- and three-objective test functions were selected among those proposed by Deb (2000) [**67**] and Zitler *et al.* [**68**]. A brief review of the main characteristic featured by each test problem and its mathematical formulation are summarized hereafter. All these problems require the fitness f minimization and introduce several difficulties that stress the search algorithms hampering convergence towards to the Pareto-optimal front.

In the formula, *k* is the number of decision variables, *M* is the number of objective functions, and k_M is the number of variables of the functional $g(\boldsymbol{x}_M)$, $k_M = k - M + 1$. The decision variables $x_i \in [0,1]$ for $i = 1,\ldots,k$, and the subgroup of decision variables $\boldsymbol{x}_M : x_j \; for \; j = M-1, M, \ldots, k.$

2.4.1.1 Two-objective test functions

- ZDT1 has a convex Pareto-optimal front:

$$\begin{aligned} &f_1(x_1) = x_1 \\ &f_2(\boldsymbol{x}_M) = gh \\ &g(\boldsymbol{x}_M) = 1 + 9\sum_{x_i \in \boldsymbol{x}_M} \frac{x_i}{k-1} \\ &h(f_1, g) = 1 - (f_1 / g)^{0.5} \end{aligned} \qquad \text{Eq. 2.6}$$

- ZDT2 has a non-convex Pareto-optimal front:

$$\begin{aligned} &f_1(x_1) = x_1 \\ &f_2(\boldsymbol{x}_M) = gh \\ &g(\boldsymbol{x}_M) = 1 + 9\sum_{x_i \in \boldsymbol{x}_M} \frac{x_i}{k-1} \\ &h(f_1, g) = 1 - (f_1 / g)^{2} \end{aligned} \qquad \text{Eq. 2.7}$$

- ZDT3 features a disconnected Pareto-optimal front:

$$
\begin{aligned}
&f_1(x_1) = x_1 \\
&f_2(\boldsymbol{x}_M) = gh \\
&g(\boldsymbol{x}_M) = 1 + 9\sum_{x_i \in \boldsymbol{x}_M} \frac{x_i}{k-1} \\
&h(f_1, g) = 1 - (f_1 / g)^{0.5} - (f_1 / g)\sin(10\pi f_1)
\end{aligned}
\qquad \text{Eq. 2.8}
$$

- ZDT4 contains 21^9 local Pareto-optimal fronts:

$$
\begin{aligned}
&f_1(x_1) = x_1 \\
&f_2(\boldsymbol{x}_M) = gh \\
&g(\boldsymbol{x}_M) = 1 + 10(k-1)\sum_{x_i \in \boldsymbol{x}_M} (x_i^2 - 10\cos(4\pi x_i)) \\
&h(f_1, g) = 1 - (f_1 / g)^{0.5}
\end{aligned}
\qquad \text{Eq. 2.9}
$$

- ZDT6 has a non-uniformity of the search space:

$$
\begin{aligned}
&f_1(x_1) = x_1 \\
&f_2(\boldsymbol{x}_M) = gh \\
&g(\boldsymbol{x}_M) = 1 + 9\left(\sum_{x_i \in \boldsymbol{x}_M} \frac{x_i}{k-1}\right)^{0.25} \\
&h(f_1, g) = 1 - (f_1 / g)^2
\end{aligned}
\qquad \text{Eq. 2.10}
$$

2.4.1.2 Three-objective test functions

- DTLZ2 has a spherical Pareto-optimal front:

$$
\begin{aligned}
&f_1\left(x\right) = \left(1 + g\left(\boldsymbol{x}_M\right)\right)\cos\left(x_1\pi / 2\right)\cos\left(x_2\pi / 2\right) \\
&f_2\left(x\right) = \left(1 + g\left(\boldsymbol{x}_M\right)\right)\cos\left(x_1\pi / 2\right)\cos\left(x_2\pi / 2\right) \\
&f_3\left(x\right) = \left(1 + g\left(\boldsymbol{x}_M\right)\right)\sin\left(x_1\pi / 2\right) \\
&g\left(\boldsymbol{x}_M\right) = \sum_{x_i \in \boldsymbol{x}_M} \left(x_i - 0.5\right)^2
\end{aligned}
\qquad \text{Eq. 2.11}
$$

- DTLZ3 introduces many local Pareto-optimal fronts with respect to DTLZ2 by changing $g\left(\boldsymbol{x}_M\right)$ functional:

$$
\begin{aligned}
f_1\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(\frac{x_1\pi}{2}\right)\cos\left(\frac{x_2\pi}{2}\right)\\
f_2\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(\frac{x_1\pi}{2}\right)\cos\left(\frac{x_2\pi}{2}\right)\\
f_3\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\sin\left(\frac{x_1\pi}{2}\right)\\
g\left(\boldsymbol{x}_M\right) &= 100\left[k+\sum_{x_i\in x_M}\left(\left(x_i-0.5\right)^2-\cos\left(20\pi\left(x_i-0.5\right)\right)\right)\right]
\end{aligned}
\qquad \text{Eq. 2.12}
$$

- DTLZ4 implements a different meta-variable mapping $x_i \rightarrow x_i^{\alpha}$, which dense the set of solutions in specific region of the domain:

$$
\begin{aligned}
f_1\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(\frac{x_1^{\alpha}\pi}{2}\right)\cos\left(\frac{x_2^{\alpha}\pi}{2}\right)\\
f_2\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(\frac{x_1^{\alpha}\pi}{2}\right)\cos\left(\frac{x_2^{\alpha}\pi}{2}\right)\\
f_3\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\sin\left(\frac{x_1^{\alpha}\pi}{2}\right)\\
g\left(\boldsymbol{x}_M\right) &= \sum_{x_i\in\boldsymbol{x}_M}\left(x_i-0.5\right)^2
\end{aligned}
\qquad \text{Eq. 2.13}
$$

with $\alpha = 100$.

- DTLZ5 features the mapping θ_2 which transforms the Pareto-optimal front form a surface to a degenerated curve:

$$
\begin{aligned}
f_1\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(0.5x_1\pi\right)\cos\left(\theta_2\right)\\
f_2\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(0.5x_1\pi\right)\cos\left(\theta_2\right)\\
f_3\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\sin\left(0.5x_1\pi\right)\\
g\left(\boldsymbol{x}_M\right) &= \sum_{x_i\in\boldsymbol{x}_M}\left(x_i\right)^{0.1}\\
\theta_2 &= \frac{\pi}{4\left(1+g\right)}\left(1+2gx_2\right)
\end{aligned}
\qquad \text{Eq. 2.14}
$$

- DTLZ5 – mod has a more complex $g\left(\boldsymbol{x}_M\right)$ functional:

$$
\begin{aligned}
f_1\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(0.5x_1\pi\right)\cos\left(\theta_2\right)\\
f_2\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\cos\left(0.5x_1\pi\right)\cos\left(\theta_2\right)\\
f_3\left(x\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)\sin\left(0.5x_1\pi\right)\\
g\left(\boldsymbol{x}_M\right) &= \sum_{x_i\in\boldsymbol{x}_M}\left(x_i-0.5\right)^2\\
\theta_2 &= \frac{\pi}{4\left(1+g\right)}\left(1+2gx_2\right)
\end{aligned}
\qquad \text{Eq. 2.15}
$$

- DTLZ6 has 2^{M-1} disconnected local Pareto-optimal regions in the search space:

$$
\begin{aligned}
f_1\left(x_1\right) &= x_1\\
f_2\left(x_2\right) &= x_2\\
f_3\left(\boldsymbol{x}_M\right) &= \left(1+g\left(\boldsymbol{x}_M\right)\right)h\\
h &= M-\sum_{i=1}^{M-1}\frac{f_i}{1+g}\left(\sin\left(1+3\pi f_i\right)\right)\\
g &= 1+\frac{9}{k}\sum_{x_i\in x_M}\left(x_i\right)
\end{aligned}
\qquad \text{Eq. 2.16}
$$

2.4.2 Methodology

The comparison methodology adopted here follows the one proposed in [**68**]. The performance of ASEMOO, GeDEA-II and GeDEA-II-K was judged for three different dimensions of the test function design space, i.e. 6, 25 and 40 design variables. Each algorithm was run 30 times over each test instance with a limited number of direct evaluations. For this purpose, The *Adimensional Direct Evaluations Number* (ADEN) was taken into account as reference cost indicator. The ADEN is defined as ratio of the direct evaluations number to the design space dimension, and it is more suited than the number of generations when competitors are not all based on evolutionary process. ADEN was limited to ten for all the simulations.

2.4.3 Metric Of Performance

Two metrics of performance were used:

- **Hyper-volume (HV)** [**69**] is a high-quality unary metric based on the computation of the hyper-volume enclosed between a Pareto Approximation Set and an arbitrary reference point (anti-ideal solution). In these experiments, the reference point is problem dependent, namely, it is equal to the maximum fitness considering all the direct evaluations used by all competitors for a specific test problem, regardless design space dimensionality. The higher the HV, the better the coverage and diversity of solutions. The HV metric was normalized (HVnorm) with the maximum value of HV considering all simulations for a specific suite case.. HV metric is well suited to make comparison among different algorithms since its metric takes into account Pareto approximations, but the Pareto-optimal front. On the other hand, we do not know how close the bets algorithm is to the Pareto-optimal front. D-metric would like to answer to such issue.
- **D-metric**: Let P^* be a set of uniformly distributed points along the Pareto-optimal front. Let A be an approximation to the Pareto-optimal front.

$$D\left(A,P^*\right)=\frac{\sum_{v\in P^*}d(v,A)}{\left|P^*\right|} \qquad \text{Eq. 2.17}$$

where $d\left(v,A\right)$ is the minimum Euclidean distance between v and the points in A. A very low value of D-metric is representative of a close and well-spread of the Pareto approximation front. The P^* was approximated with 100 points for all the bi-objective problems and 1000 points for the three-objective.

The results are reported by means of box plots, which statistically infers the outcomes of the 30 runs. On each box, the central line represents the median, the edge of the box are the 25^{th} and 75^{th} percentiles, the whiskers extend to the most extreme data points not considered outliers. In order to analyze the convergence history, two snapshots for ADEN equal to 6 and 10 were taken.

2.4.4 Experimental Settings

As for GeDEA-II and GeDEA-II-K, the population size *nindi* depends on problem dimensionality *k*:

$$nindi = min\Big(round\big(1.5\mathrm{k}\big), 20\Big) \qquad \text{Eq. 2.18}$$

while for the other genetic parameters the reader is referred to [**59**]. Equation Eq. 2.18 is the result of different experiments on GeDEA-II aimed to squeeze out the best performance from the algorithm reducing the overall direct evaluations. As a result, the number of generations depends on the problem dimensionality; however, in the best case, which is for ADEN equal to 10 and 40 design variables, only 20 generations are performed. On the other hand, ASEMOO was used in synchronous mode, one sampling point at a time, with an initial sampling database of 2*k* points. The synchronous mode is the one that guarantees the best performance.

2.4.5 Results Of Comparison

The fundamental aspect of this investigation is the very low number of the direct evaluations offered to the three algorithms in order to converge towards the Pareto-optimal front. Compared to the test campaign followed in [**59**], in which GeDEA-II has demonstrated to be superior with respect to competitors, the number of direct evaluations are here reduced by 75 % on average. As a result, the Pareto approximation is intrinsically weak compared to other investigations in which thousands of evaluations are used. However, a low ADEN is representative of all those industrial applications when costly fidelity models are employed and computational efforts and time are in short supply.

All the figures report a specific suite problem with four main information: two plots of the Pareto approximation for dimensionality equal to 25 and two box-plots of the of D-metric and HV. The first plot illustrates the Pareto approximation at ADEN equal to 10 of a single run featuring a mean value of the HV, while the second one summarizes all the runs. Comparison between the two is intended to show how the low ADEN affects the coverage of the Pareto-optimal front. Sometimes, the lack in the full coverage of the front is not a matter of exploration capability, whereas reflects the stochastic nature of the EAs added with the low ADEN.

As far as the ZDT1, ZDT2 and ZDT3 bi-objective problems are concerned, ASEMOO seems to converge towards the Pareto-optimal front much better than GeDEA-II; furthermore, ASEMOO offers a high quality approximation after only ADEN=6, as suggested by both the HV and D-metric (Figure 2.2 to Figure 2.4). It is worth nothing that ASEMOO is almost insensitive to the problem scalability. Increasing the number of design variables slightly deteriorates ASEMOO performance at least for low ADEN values, while for ADEN=10 it seems that the effect is negligible. On the other hand, the difficulties experienced by GeDEA-II are related with the spread of solutions rather than the distance of solutions from the Pareto-optimal front. In fact, GeDEA-II converges toward the front only in some regions and clusters few solutions around these spots. Such behavior is well explained due to the adverse juncture of limited population size with the very low number of direct evaluations that are the worst-case scenario for a genetic algorithm. As the dimension gets higher, this situation is mitigate by a larger amount of direct evaluations, but still more of them will be needed to reach the competitor's Pareto front. As for repeatability, which is proportional to the width of the boxplots body, ASEMOO is superior to GeDEA-II due to the deterministic nature of EHVI function sampling criterion. At least for these three problems characterized by convex, non-convex and discrete fronts, Kriging approximation adopted in ASEMOO algorithm works properly. Bear in mind that ASEMOO was tuned on such problems [**64**]. The cooperation strategy adopted by GeDEA-II-K has the effect to accelerate the Pareto front coverage process in all the three-suite cases. GeDEA-II-K's performance gets really close to the one of ASEMOO, and this is even truer when the problem dimensionality gets higher. From the single run plot of Figure 2.4 it can be inferred that GeDEA-II-K has a poor uniformity in the front approximation compared to ASEMOO, but still it has found all the four chunks of the Pareto-optimal front. Moreover, GeDEA-II-K has improved repeatability.

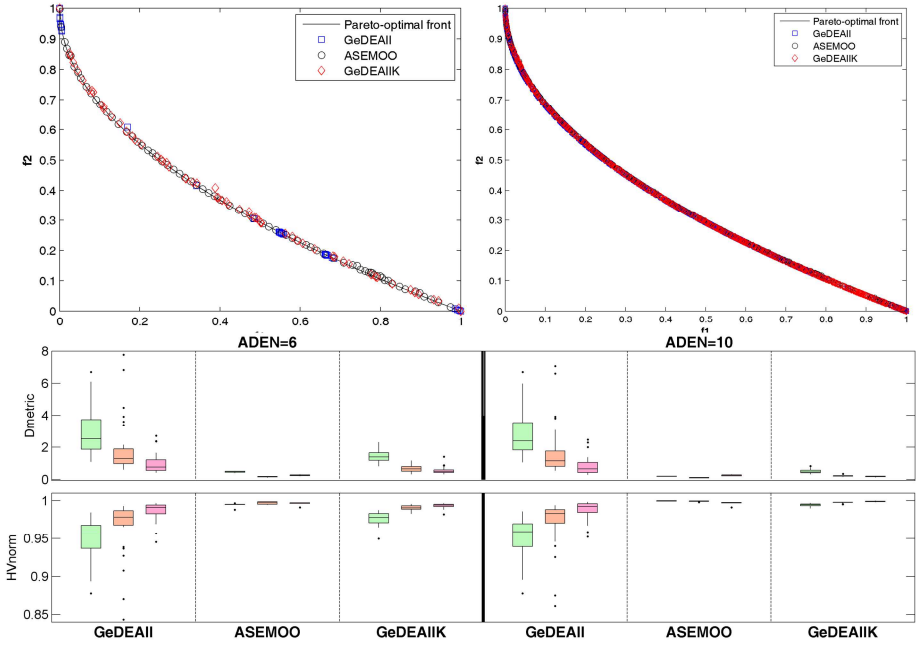


Figure 2.2: Test function ZDT1: Pareto front for single run (top left) and all runs(top right) for design space dimension = 25; at the bottom, box convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

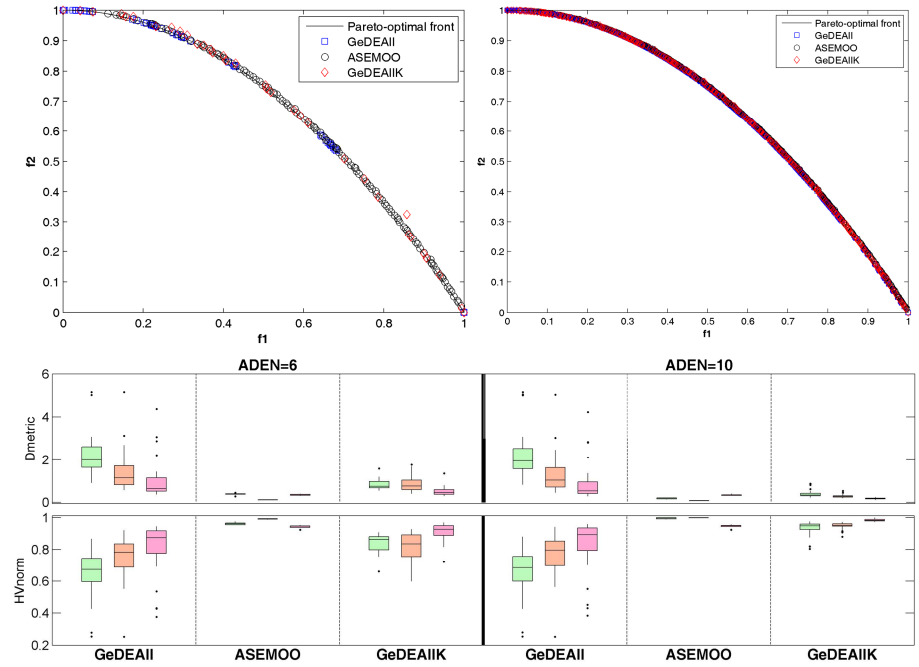


Figure 2.3: Test function ZDT2: Pareto front for single run (top left) and all runs(top right) for design space dimension = 25; at the bottom, box convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

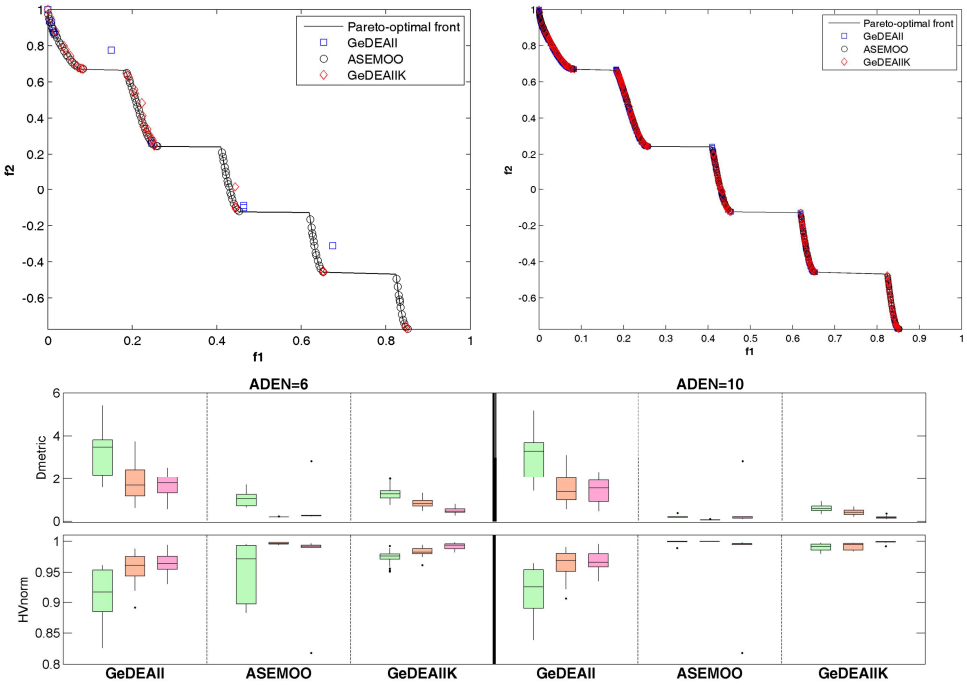


Figure 2.4: Test function ZDT3: Pareto front for single run (top left) and all runs(top right) for design space dimension = 25; at the bottom, box convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

The outcomes change when ASEMOO is applied to multi-modal and non-uniform fronts as those featured by ZDT4 and ZDT6 problems, respectively. For instance, the ZDT4 tests demonstrated an overturning of performance (Figure 2.5): GeDEA-II outperforms ASEMOO, and reveals a behavior specular to the one recorded by its competitor in the previous problems. On such multimodal landscape, ASEMOO is affected by dimensionality issues, which cause the boxplot width and median to get higher as the design variables increase. As for GeDEA-II, the enhancement of performance with higher dimensions observed before recurs for both ZDT4 and ZDT6, and it is even more pronounced. In such battlefield, GeDEA-II-K clearly improves the uniformity of the front converge as depicted in Figure 2.5 and Figure 2.6.

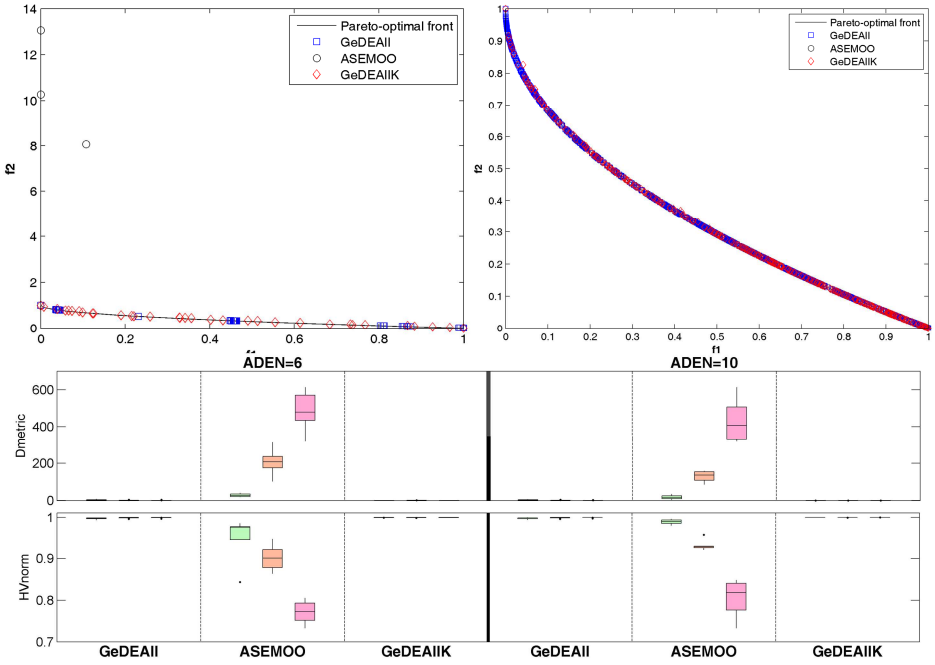


Figure 2.5: Test function ZDT4: Pareto front for single run (top left) and all runs(top right) for design space dimension = 25; at the bottom, box convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

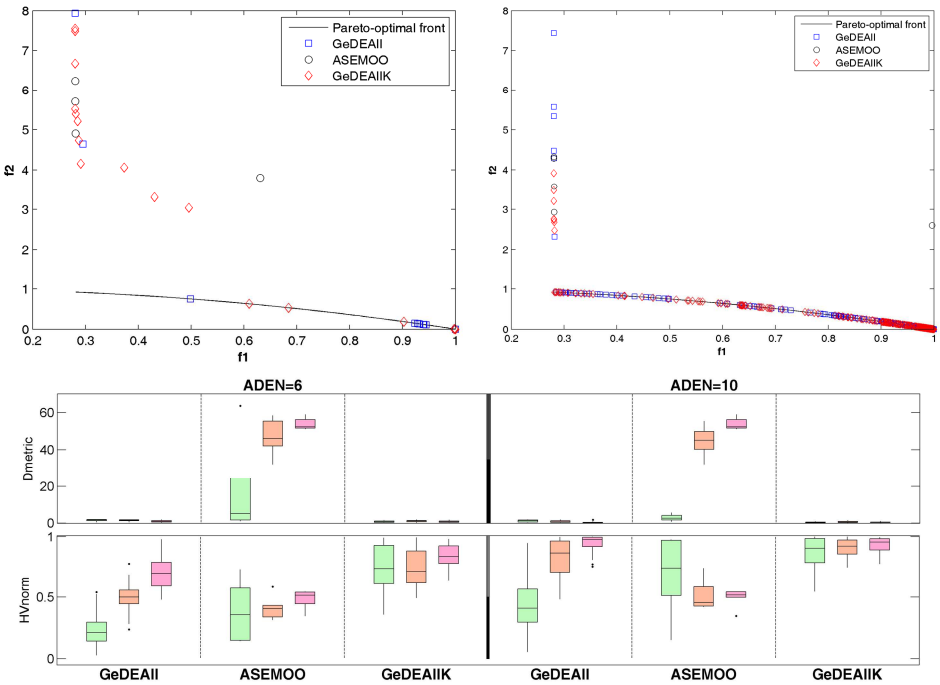
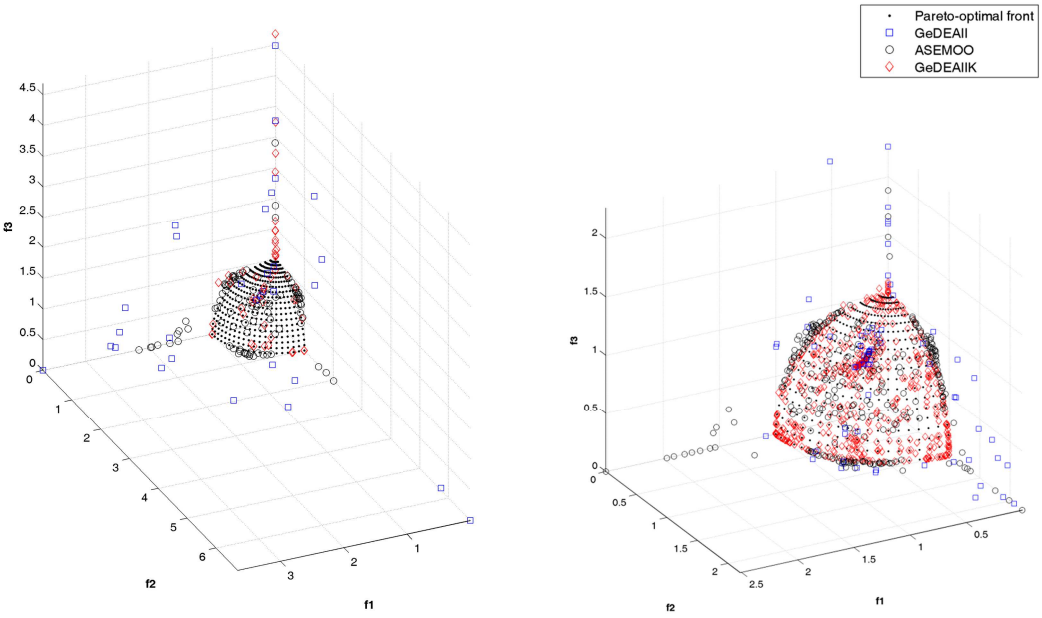


Figure 2.6: Test function ZDT6: Pareto front for single run (top left) and all runs(top right) for design space dimension = 25; at the bottom, box convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

As for the three-objective problems, DTLZs' class was adopted. Some common considerations can be extrapolated from the six DTLZs' test cases reported from Figure 2.7 to Figure 2.12. As the number of objective rises, the number of fitness evaluations needed to converge properly toward the Pareto-optimal front increases. The level of front coverage does deteriorate moving from bi- to three-objective problems, since the 3D topology of the Pareto-optimal front requires a higher number of points in order to have a comparable resolution. Exception made by the DTLZ5 that has a 3D line as Pareto-optimal front, and for which GeDEA-II and GeDEA-II-K reach a remarkable Pareto front approximation. In the other cases, limiting ADEN to ten leads to poor results. However, in such adverse conditions GeDEA-II seems to be slightly superior at high dimensions, while at low dimensions the performance is comparable with ASEMOO. Moreover, ASEMOO is very sensitive to the search space dimensions for all those test problems featuring a spherical Pareto-optimal front. Finally, repeatability is problem dependent and no generalization can be formulated.

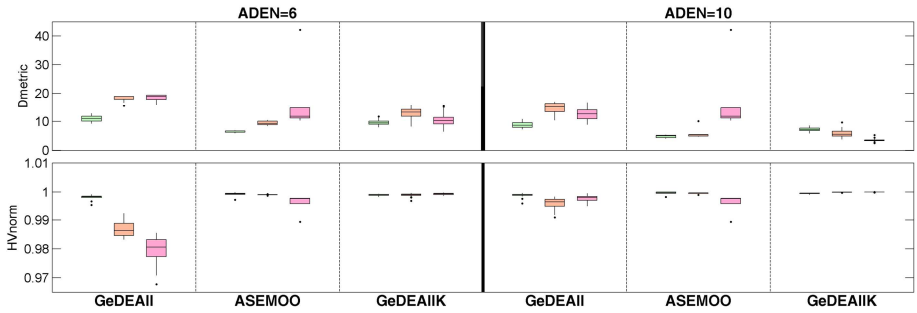
GeDEA-II-K enhances GeDEA-II performance over all the three-objectives tests. Both repeatability and the front coverage are improved. It is symptomatic DTLZ6 in Figure 2.12 where points of GeDEA-II and ASEMOO are clustered along edges of Pareto-optimal front patches; this is not the case for GeDEA-II-K that pushes the search inside all the patches. Such behavior it is even more evident in the plot comprising all the runs.

Some minor notes on DTLZ3 that seems to be impenetrable for all three the algorithms, at least for ADEN equal to 10. Despite GeDEA-II-K gets nearer to the spherical front, it is still too far. Moreover, although on DTLZ5 all competitors reach the Pareto-optimal front, on the modified version DTLZ-mod (Figure 2.11) ASEMOO gets trapped in some local front, whereas the other two algorithms collapse on the front, having GeDEA-II-K a superior spread of solutions.



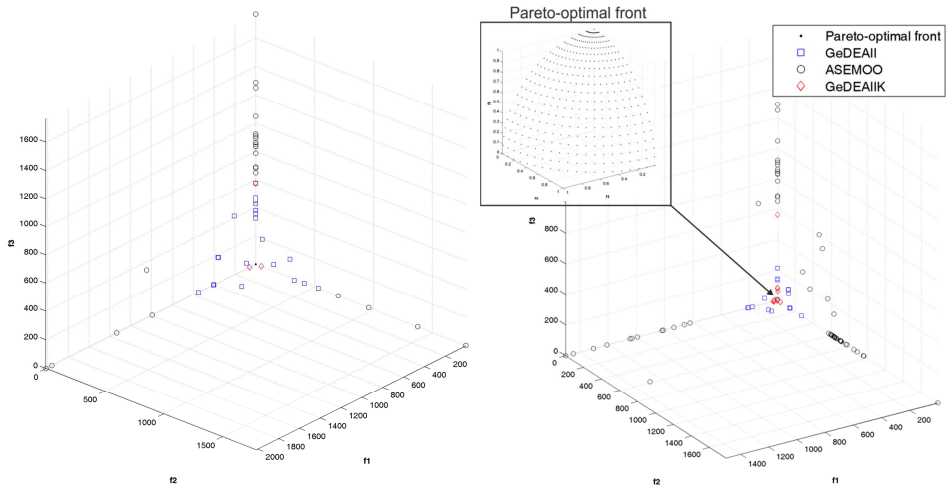
(a) Pareto front approximation of a single run with dimensionality = 25.

(b) Pareto front approximation of all runs with dimensionality = 25.



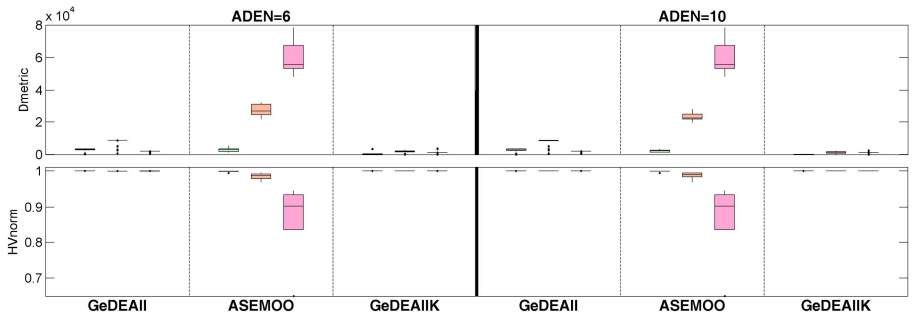
(c) Box-plot convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

Figure 2.7: Test function DTLZ2.



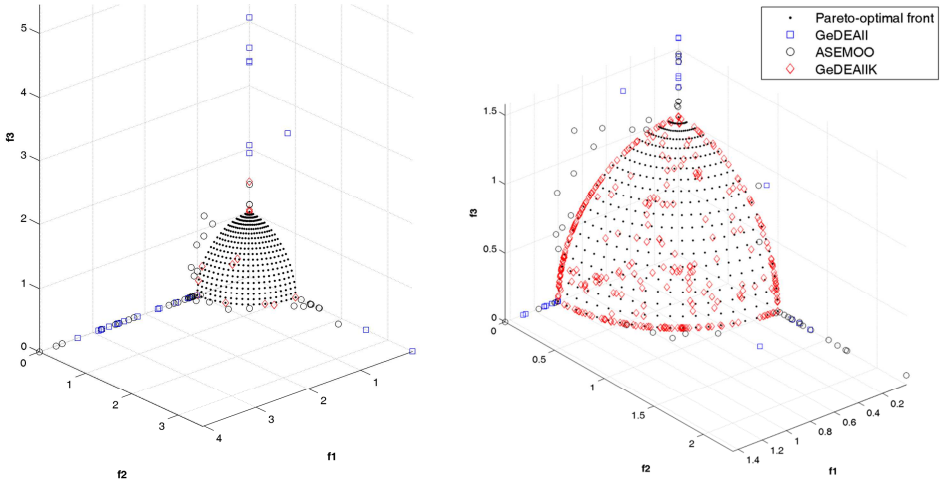
(a) Pareto front approximation of a single run with dimensionality = 25.

(b) Pareto front approximation of all runs with dimensionality = 25.



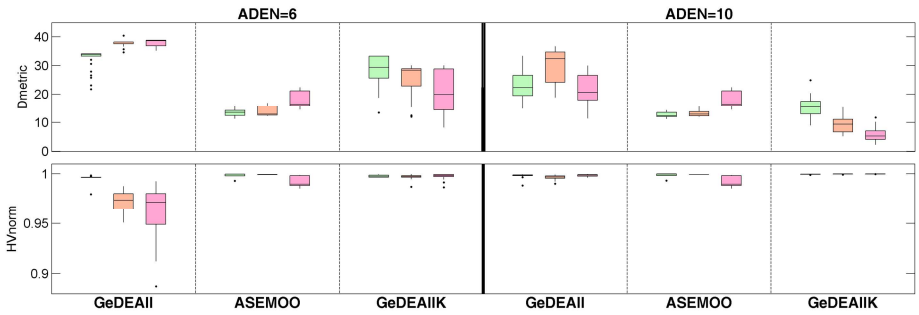
(c) Box-plot convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

Figure 2.8: Test function DTLZ3



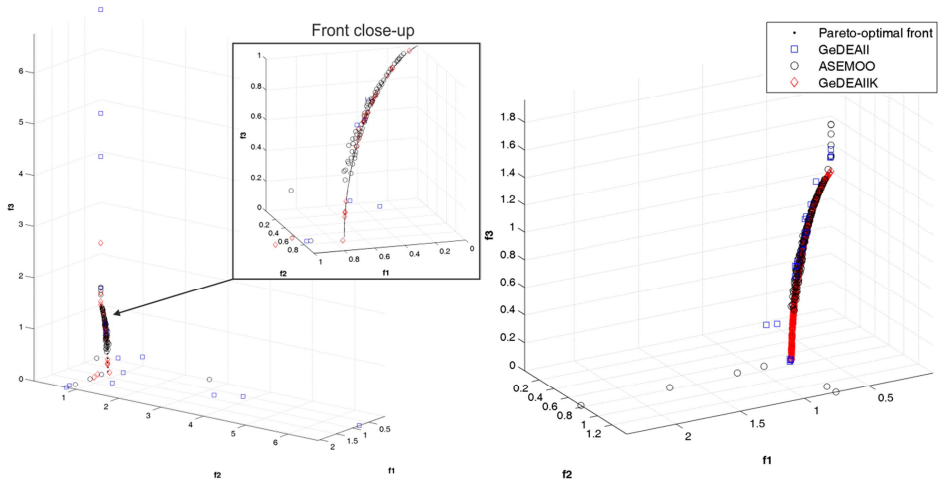
(a) Pareto front approximation of a single run with dimensionality = 25.

(b) Pareto front approximation of all runs with dimensionality = 25.



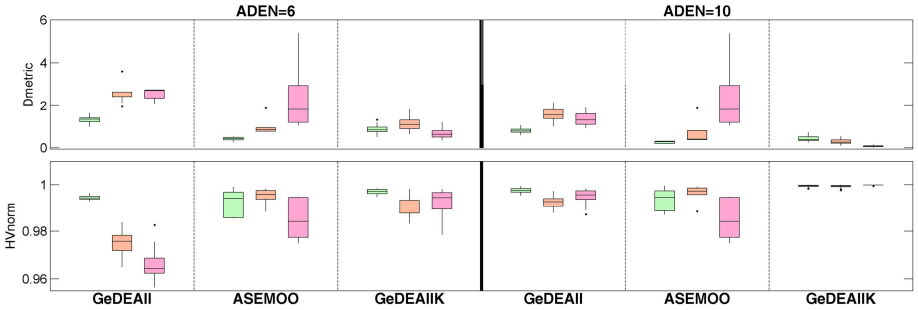
(c) Box-plot convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

Figure 2.9: Test function DTLZ4.



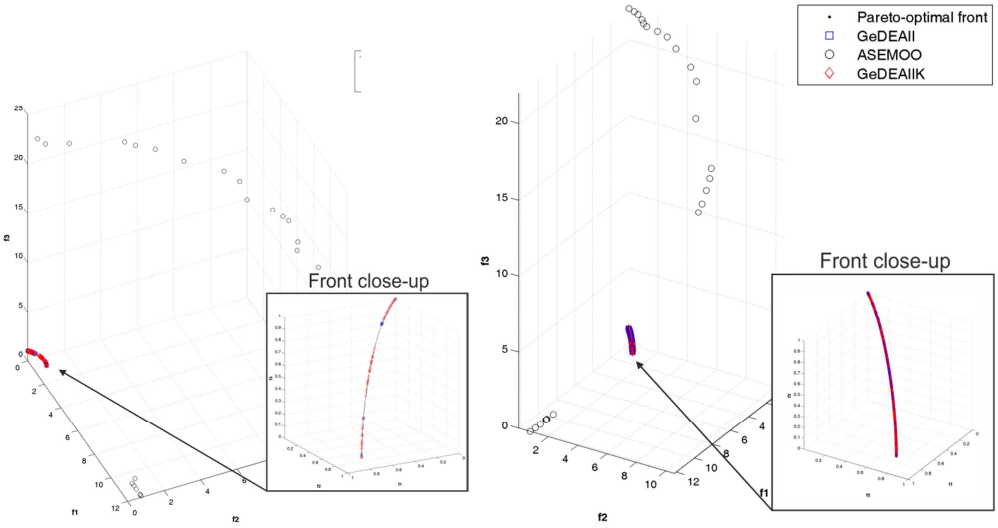
(a) Pareto front approximation of a single run with dimensionality = 25.

(b) Pareto front approximation of all runs with dimensionality = 25.



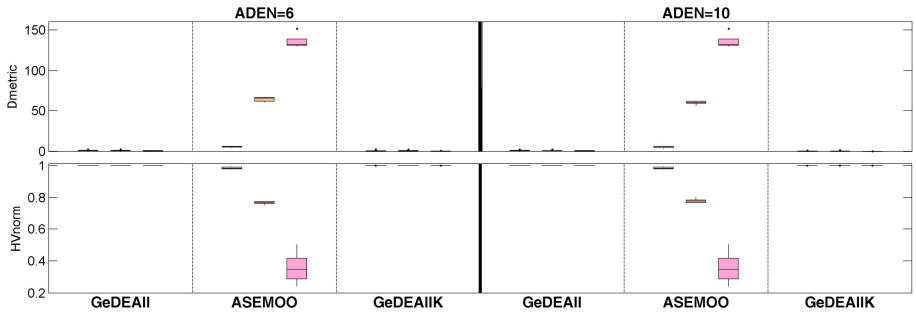
(c) Box-plot convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

Figure 2.10: Test function DTLZ5.



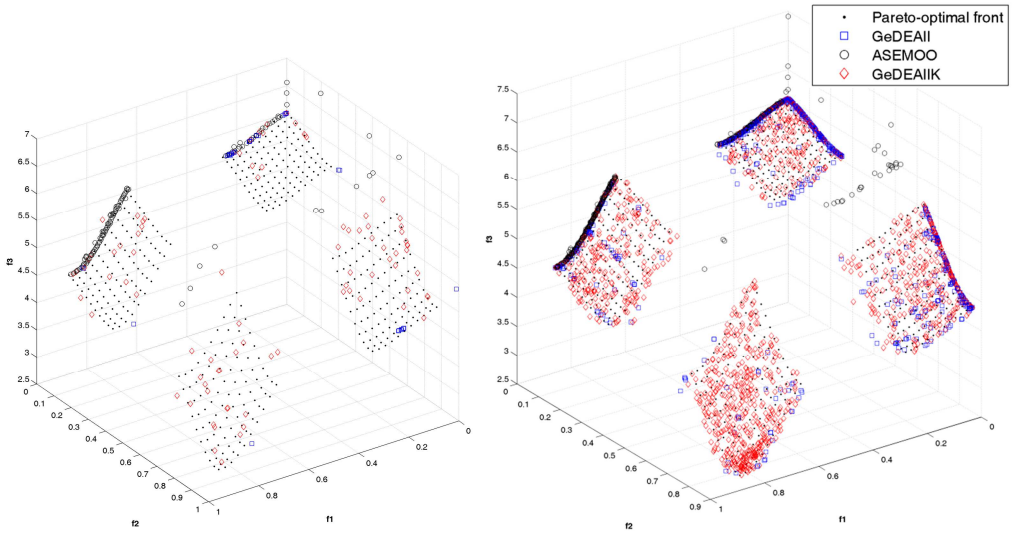
(a) Pareto front approximation of a single run with dimensionality = 25.

(b) Pareto front approximation of all runs with dimensionality = 25.



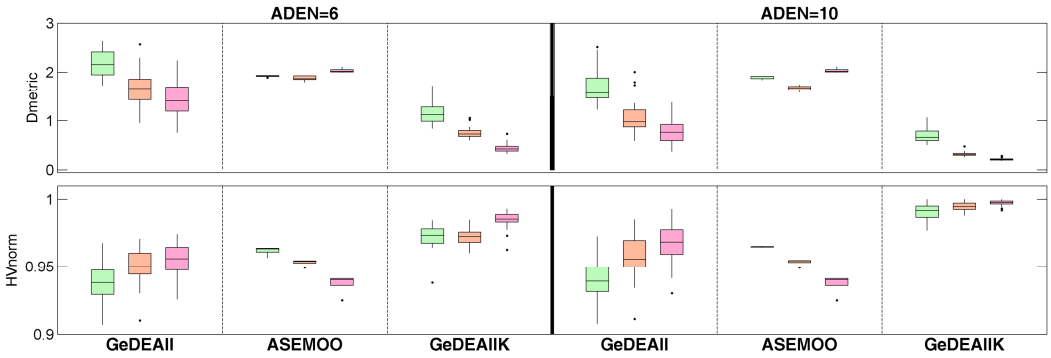
(c) Box-plot convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

Figure 2.11: Test function DTLZ5-mod.



(a) Pareto front approximation of a single run with dimensionality = 25.

(b) Pareto front approximation of all runs with dimensionality = 25.



(c) Box-plot convergence history reported at ADEN equal to 6 and 10 of D-metric and normalized Hyper-volume for different design space dimensions, 6 (green), 25 (orange) and 40 (pink).

Figure 2.12: Test function DTLZ6.

2.5 Conclusions

The cooperation between GeDEA-II genetic algorithm and ASEMOO Kriging-based algorithm is realized by means of the Kriging filter featured in the GeDEA-II-K algorithm. GeDEA-II-K shares the same framework of its predecessor (i.e., GeDEA-II) and adds the Kriging filter operator at the end of GeDEA-II's reproduction phase; the filter acts as a selection operator of the GeDEA-II's offspring; according to the inferred information coming out from the Kriging model, the filter decides which individual is more promising and so can

be evaluated. Furthermore, the information of the EHVI criterion is introduced in filter to exploit the Kriging exploration-exploitation balance capability.

The comparison over two- and three-objective test functions revealed the effectiveness of GeDEA-II-K, which is able to unite GeDEA-II and ASEMOO in such a way to avoid destructive interactions between the two and to synergistically intensify the strengths of both. The comparison made use of very few direct evaluations, which resembles the real life application where high-fidelity models are employed.

As two-objective test problems are concerned, surrogate-assisted based algorithms as ASEMOO could make the difference on convergence toward the Pareto-optimal front reaching an outstanding resolution with high repeatability when a very low number of direct evaluations are used. This is true if the surrogate model gets an appropriate insight on the fitness functions landscape, otherwise the performance rapidly deteriorates. Since GeDEA-II-K does not rely only on Kriging model, its performance is high-level even when ASEMOO discloses its weakness. Of course, when ASEMOO works properly, GeDEA-II-K cannot assure a so high Pareto front resolution but, still, its front approximation always reaches a “minimum level” that could be considered sufficient for real-life MOOPs. Moreover, GeDEA-II-K always assures a better performance than GeDEA-II regardless of the problem at hand.

Dealing with three-objective problems has a major impact on performance principally because of to the limited number of direct evaluations compared to the 3D topology of the Pareto-optimal front. All competitors get pour performance compared to bi-objective problems. However, GeDEA-II-K maintains that “minimum level” among test problems and greatly improves the capability of covering the Pareto-optimal front proven by GeDEA-II, compatibly with the number of direct evaluations.

The dimensionality of the design space affects in opposite directions the three algorithms: for ASEMOO the increase of dimensionality is detrimental on performance, while GeDEA-II and GeDEA-II-K experience benefits due to total amount of direct evaluations.

Chapter 3

Supersonic Compressor Cascades

3.1 Why Supersonic Flows

Since 1938 the need for compressor stages to obtain higher pressure ratios with large flow capacity has pushed the development of high-speed fans and compressors operating with supersonic relative inlet Mach numbers, the so called transonic compressor. The advantage to operate at high relative inlet velocity directly stems from Euler's momentum equation

$$h = (U_2 c_{\theta 2} - U_1 c_{\theta 1}) \qquad \text{Eq. 3.1}$$

It is clear that the two factors controlling specific energy transfer h within the compressor rotor are the blade speed U and the absolute flow turning, which cannot be increased above certain limits. For sure, increasing wheel speed, thereby the relative inlet velocity, is an effective way to augment the energy transfer, while keeping the aerodynamic loading unchanged. This is the main reason of high pressure ratios achieved in transonic compressors.

Unfortunately, near–tip rotor sections experience supersonic relative flows, shock waves phenomena arise in the inlet and passage regions drastically affecting efficiency of the machine due to both the entropy rise across shocks and the interaction of the shock waves with the boundary layer.

Research on supersonic cascades meant to be one piece of the puzzle to allow considerable insight on the complex flow inside transonic compressors. Despite flow phenomena are highly three dimensional in nature, systematic analysis of 2D supersonic cascades is a well posed starting point to understand how to mitigate losses related to shock waves while preserving advantages of their compression mechanism.

3.2 Historical Survey

In the late ‘40s, there were still many doubts on the possibility to design an efficient supersonic axial flow compressor due to shock waves losses. This diehard certainty was corroborated by the experience on isolated bodies, for which the large energy losses usually occur due to wave systems that extend far from the bodies themselves. In the milestone work by Kantrowitz (1950) [**70**] it was theoretically demonstrated that a cascade could entirely eliminate this extended wave system, or at least weakening its strength far from the cascade, thus allowing to efficiently exploit the shock waves compression mechanism inside the blade passage. Since then, much effort has been made to design an airfoil able to efficiently handle supersonic inlet flow and the inherent wave shock system. Nowadays, the *S-shape profile* seems to be the solution to this problem.

3.2.1 Classification Of Supersonic Compressor Cascades

Over the years, various cascade geometries have been designed for both rotors and stators. The way the rotor mechanical energy is transferred to the fluid marks the division between different kind of supersonic cascades classes [**71**] (Starken and Lichtuff 1970): *the pure impulse cascade* (Figure 3.1), the *high turning supersonic reaction cascade*, and the *low-turning supersonic reaction cascade* (Figure 3.2). Recalling Euler’s momentum equation, the specific energy transfer is proportional to the absolute flow turning, which can be realized in different ways such as decelerating the relative velocity, turning the relative velocity, or a mix of the two.

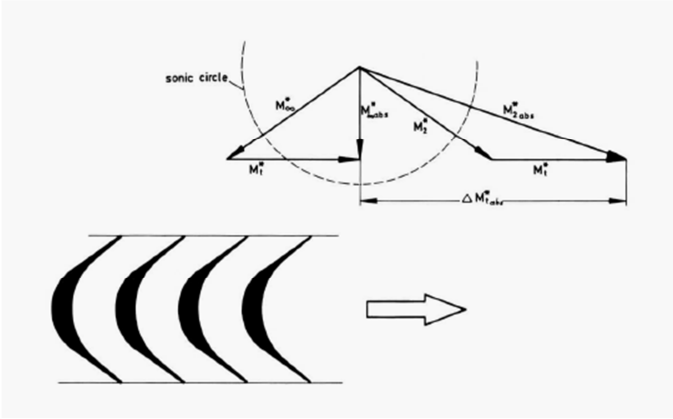


Figure 3.1: Supersonic impulse cascade (from Starken and Lichtuff 1970 [**71**]).

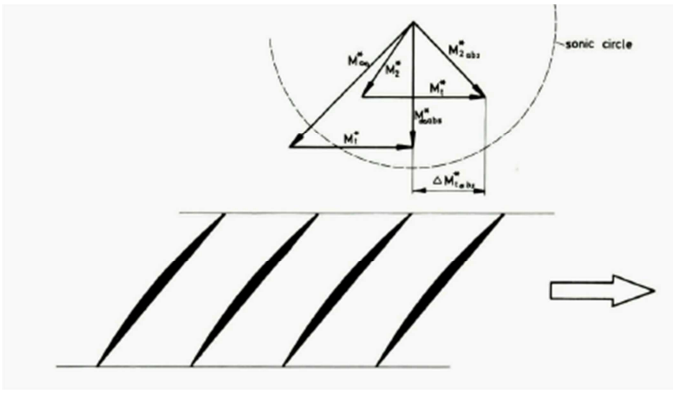


Figure 3.2: Low-turning supersonic reaction cascade (from Starken and Lichtuff 1970 [**71**]).

In the impulse cascade the mechanical energy of the rotor is mainly transferred into kinetic energy of the fluid without any static pressure increase. The entire change in the static pressure has to take place completely in the following stator at high supersonic velocity; the module of the relative inlet Mach number remains nearly unchanged across the row, while the flow undergoes to a high turning symmetrical to the inlet axial flow direction. Although various design procedures have been developed by Shapiro (1953) [**72**], Stratford (1962), Oswatitsch (1956) [**73**], all methods require a local acceleration and deceleration of the flow that may lead to separation of the boundary layer.

In order to accommodate the incoming flow out from impulse rotors, the high turning supersonic reaction cascade was designed to ensure high static pressure rise and flow turning. Details on such kind of design are found in Wilcox (1955) [**74**], Hartman (1953) [**75**], Klapproth (1952) [**76**], Shapiro (1953) [**72**], Johnson (1959) [**77**].

The velocity triangles of Figure 3.1 and Figure 3.2 show how the exit flow condition changes from supersonic to subsonic, going from an impulse to a low-turning supersonic cascade.

As for the *low-turning supersonic cascade*, the absolute flow turning (see Figure 3.2) is mainly due to deceleration of relative velocity through sound speed rather than turning of the relative flow, so that some part of the transferred energy is already converted into pressure energy within the rotor producing a high static pressure. The deceleration of the inlet supersonic flow can be realized in two different ways across the rotor cascade:

1. compression due to the flow area contraction with deceleration of supersonic flow through sound speed and further diffusion to subsonic

exit condition. Similar to a convergent-divergent nozzle, the supersonic deceleration is realized with a reduction in the flow area in the forward part of the blade passage until sonic throat, while in the aft part subsonic deceleration is continued by means of an increase in area. The main difference between a nozzle and a cascade is that in the latter such area variations can be realized only by flow deflection, due to the geometric periodicity constraint of the cascade passage itself. However, deflection has opposite sign regarding the supersonic and subsonic branch, thus could compensate the different deflections, and leave the flow direction unchanged between cascade inlet and outlet;

2. compression due to a normal shock wave in the blade passage.

In principle, such methods could actually provide flow deceleration without flow turning, or at least with a moderate turning. In practical application, however, both methods have severe limitations. Due to stability problems similar to those of a convergent-divergent engine inlet [**78**], the first method requires a normal shock wave behind the throat area in the diverging part of the passage. The exact shock location, from which the shock strength depends, is a function of on the exit back pressure, and so of the cascade operation. In the practice, there can be operating regimes where the shock strength becomes too severe. On the other hand, the second method is impracticable due to entropy rise across the shock, which becomes unacceptable for Mach numbers ahead of the shock above 1.6 [**79**] (Lichtfuss and Starken 1974). Nevertheless, the *S-shape profile* belongs to the *low-turning cascade* class and, from now on, particular attention will be paid on such class.

3.2.2 Towards The S-Shape Design

Supersonic cascades that adopt the shock wave compression mechanism are subject to two peculiar sources of loss: the entropy rise across the shock and the interaction mechanism of the shock waves with the boundary layer. Over the years, there has been the tendency to minimize the former loss acting on both the reduction of the Mach number ahead of the shock and on the shock wave pattern, moving from a normal to an oblique shock within the cascade passage. A weaken shock could, in principle, moderate the shock boundary-later interaction too.

In the Double Circular Arc profiles (DCA) the suction and pressure sides are defined by two circular arcs. The convex curvature of the suction side induces

acceleration to the incoming supersonic flow, which is deflected by supersonic expansion waves (Prandtl-Meyer turning) in order to follow the suction surface curvature. As a result, the inlet Mach number increases from the leading edge till the normal shock wave located at the entrance of the passage. The shock strength becomes overblown for Mach number ahead of the shock around 1.6, restricting the operating range to a maximum inlet Mach number around 1.3.

This limitation led to the concept of Multiple Circular Arc profiles (MCA), which incorporates and extends the DCA definition scheme to include the potential for defining the suction surface with two or more circular arcs of different curvature. A low curvature is usually adopted for the suction forward surface, which ensures reduced acceleration and shock losses of the flow in the entrance region. This concept was first proposed by Seyler and Smith (1967) [**80**]. Comparison of results obtained from both rotors (Gostelow *et al.* (1968) [**81**]) and linear cascade (Mikolajczak *et al.* (1971) [**82**]) tests with MCA and DCA profiles indicates that definite advantage by way of efficiency and maximum pressure ratio exists when MCA are used, confirming that excessive supersonic expansion in the cascade entrance region deteriorates performance. Further reduction in curvature of the entrance region suction surface characterize a new type of profile called Circular-Wedge profile (CW) investigated by Emery *et al.* (1960) [**83**]. Since CW profiles features an infinite curvature (i.e., flat surface) in the forward region, the upstream Mach number is kept almost constant up to the normal shock. Another similar profile is the J-shape profile developed by Hearsey and Wennerström (1970) [**84**].

Despite wide improvements on the reduction of Mach number ahead of the shock, the main issue regarding all the previous profile shapes is the normal shock wave at the passage entrance. Such shock topology inherently restricts the cascade operating range to an inlet Mach number around 1.6. To work at higher inlet Mach numbers while preserving efficiency, the normal shock wave must be avoided and replaced with several oblique shock waves. The ensemble of oblique shocks can produce the same amount of pressure rise in a more efficient way. In principle, the shock losses could be altogether avoided by using *homentropic compression* (isentropic compression), which was firstly proposed by Oswatitsch (1947) [**85**] for supersonic aircraft intakes. Isentropic compression can be applied on supersonic cascades in two different ways: *inner compression* and *outer compression*, depending on whether the isentropic compression takes place within the passage or in the entrance region; of course, a mixed solution is also possible.

Kantrowitz (1950) [**70**] was the precursor of the *internal compression* concept applied to transonic compressors, concept previously employed in supersonic diffusers for supersonic cascades. Actually, in the Kantrowitz's rotor the relative supersonic flow is decelerated through the speed of sound by a normal shock inside the passage, and then, in the passage subsonic region further diffusion is accomplished by area divergence. The profile proposed by Kantrowitz is depicted in Figure 3.3. The reduction of the Mach number ahead of the normal shock was the only way to improve rotor efficiency and, albeit with simplicity, it was pursued adding thickness on the suction side, as depicted in Figure 3.3. The experiments highlighted that the subsonic diffusion region diverged too rapidly leading to serious separation losses, which could also be exaggerated by shock-boundary-layer interaction. It was mandatory to reduce the rate of this divergence and also the annulus was made to converge downstream. As a result, the pressure ratio materially decreased compared to the design value and the Mach number leaving the blading was therefore about 0.97 instead of 0.68. It is worth nothing that the inclusion of a concave region on the suction side of the blades immediately behind the entrance region considerably improved the efficiency of the compressor.

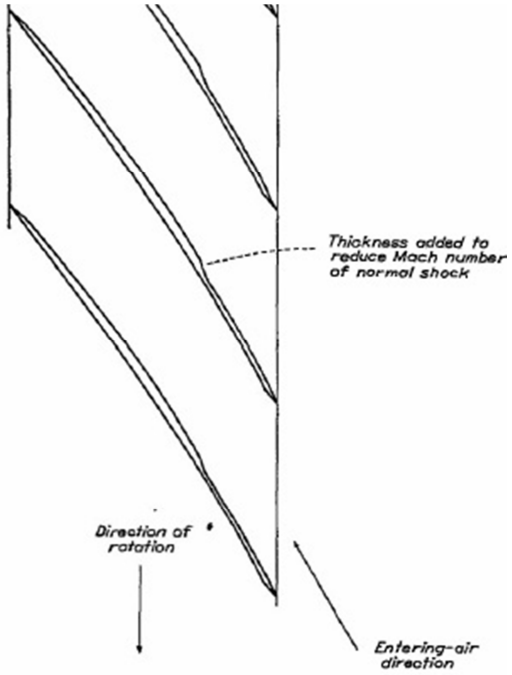


Figure 3.3: Supersonic profile proposed by Kantrowitz (1950) [**70**].

On the other hand, the *external compression* principle, or *spike-type diffuser* concept, was applied to rotors by Creagh and Klapproth (1953) [**86**] and by Lawrence and Melvin (1954) [**87**] (Figure 3.4). The compression takes place in the entrance region, before the flow gets into the passage. The effect of external compression is to decrease the streamline area in the blade inlet region, thus effecting a reduction in the in flow Mach number at the passage-inlet closure line (line c-d in Figure 3.4), where the Mach number should reach unity. According to the experiments, the sonic throat at the passage entrance is very sensible to inlet conditions and dastrically deteriorates at off-design operation.

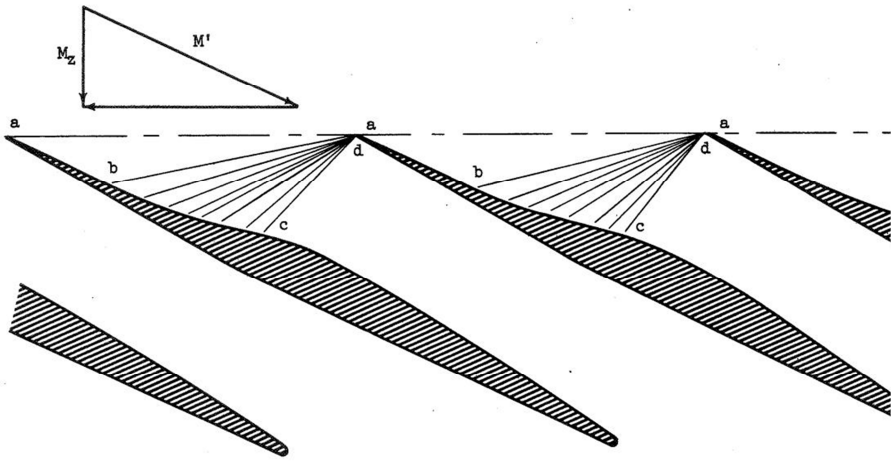


Figure 3.4: External compression principle applied to compressor cascades (Creagh and Klapproth (1953) [**86**])

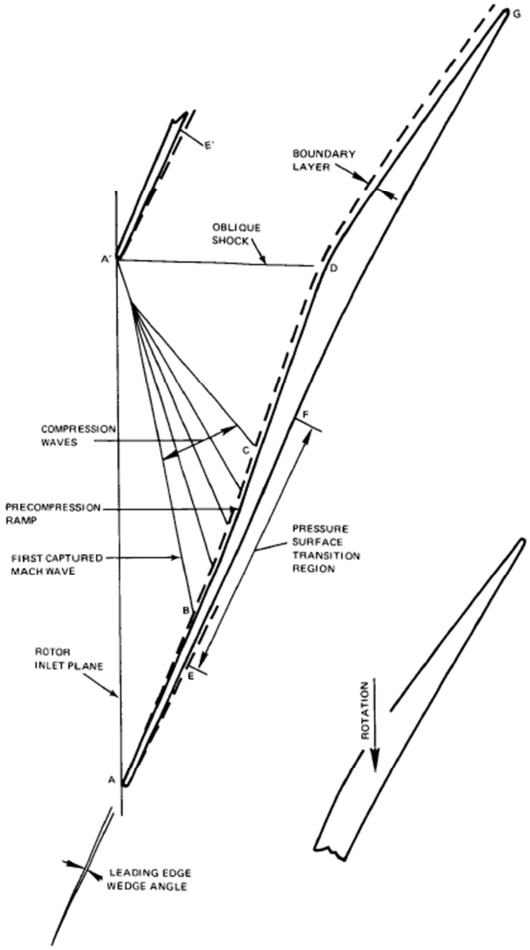


Figure 3.5: Pre-compression airfoil (Morris *et al.* 1972 [**88**]).

A more effective geometry, even in off-design operations, is the one proposed by Morris *et al.* (1972) [**88**] and depicted in Figure 3.5. The precompression design model assumes that the shock across the channel entrance must be oblique and attached to the leading edge of the airfoil.

The pre-compression is the result of the concave surface of the suction side (BC in Figure 3.5) that generates a series of compression waves that diffuse the supersonic flow. Channel flow downstream of the oblique shock is subsonic, turning and stream tube area are made compatible with the exit aerodynamic conditions.

3.2.3 S-Shape Profile

The *S-shape profile* belongs to the low-turning supersonic profile class and features the pre-compression mechanism. The PAV-1.5 cascade investigated by

Schreiber (1986) [**89**] at DFVLR has been taken as reference. The design inlet wave pattern, the real shock-wave pattern at design pressure ratio, and a Schlieren visualization are shown in Figure 3.6 and Figure 3.7.

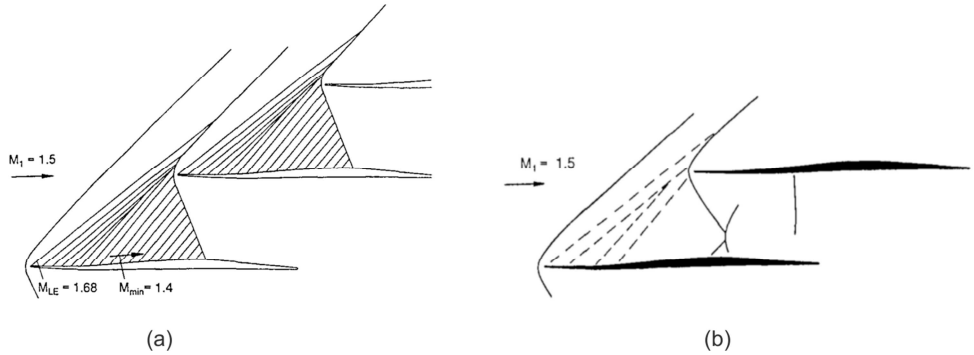


Figure 3.6: PAV 1.5 cascade: design inlet wave pattern (a) and experimental shock wave pattern at design pressure ratio (from Schreiber [**90**]).

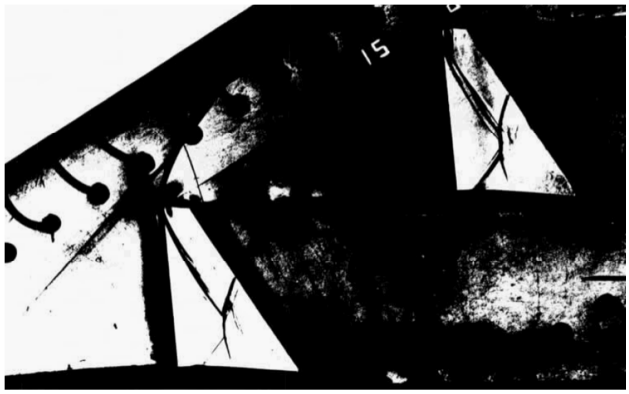


Figure 3.7: Schlieren photograph of the PAV1.5 cascade at inlet Mach of 1.5.

The curvature of both suction and pressure side and the wedge angle of the leading edge drive the main phenomena in the inlet region. First, the finite thickness of the leading edge develops a detached bow shock, which is normal at least in the very near region of the leading edge. Two oblique shock branches depart from the bow, a weaker one that extends into the upstream region and a stronger one that runs into the covered passage. The front portion of the suction surface features a concave curvature, similarly to the Miller's design, from which *left-running characteristics* depart towards the detached bow shock of the adjacent blade, and their coalescence forms the pre-compression shock wave responsible for the pre-compression mechanism. Although this shock is relatively weak, it significantly decelerates the incoming flow entering the covered passage,

from a peak value of about 1.67 to a level around 1.4. From the CFD analysis reported in Figure 3.8, it can be noted that there is a region of expansion waves between the bow shock and the pre-compression waves. In Figure 3.9 the contours of the Mach number for the same cascade are reported. The reason of such expansion fan is briefly explained. After the shock bow, the subsonic flow approaches the leading edge "upper side" (i.e., the one that leads to the suction side) with a certain incidence. In front of the leading edge curved surface, the flow accelerates until sonic velocity (point A in Figure 3.9). In order to follow the remaining part of the leading edge, which is characterized by a continuous lowering of the surface slope, the flow needs to turn more. However, since the flow is just above sonic Mach, such deflection can be realized only by means of the Prandtl-Meyer expansion waves. The turning, which is associated with supersonic flow acceleration, proceeds until the surface slope reaches its minimum. Such condition is realized in point B of Figure 3.9; point B identifies the end of the expansion fan and here the flow has its maximum Mach number of 1.8. Beyond point B, the variation of the suction surface slope changes from negative (i.e., such negative variation causes the strong expansion fan) to positive.

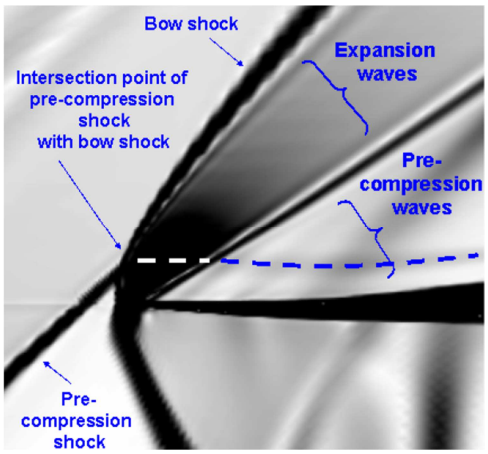


Figure 3.8: Simulated Schlieren picture of the leading edge of the PAV 1.5 cascade at inlet Mach number = 1.457 (Sonoda *et al.* [**91**]).

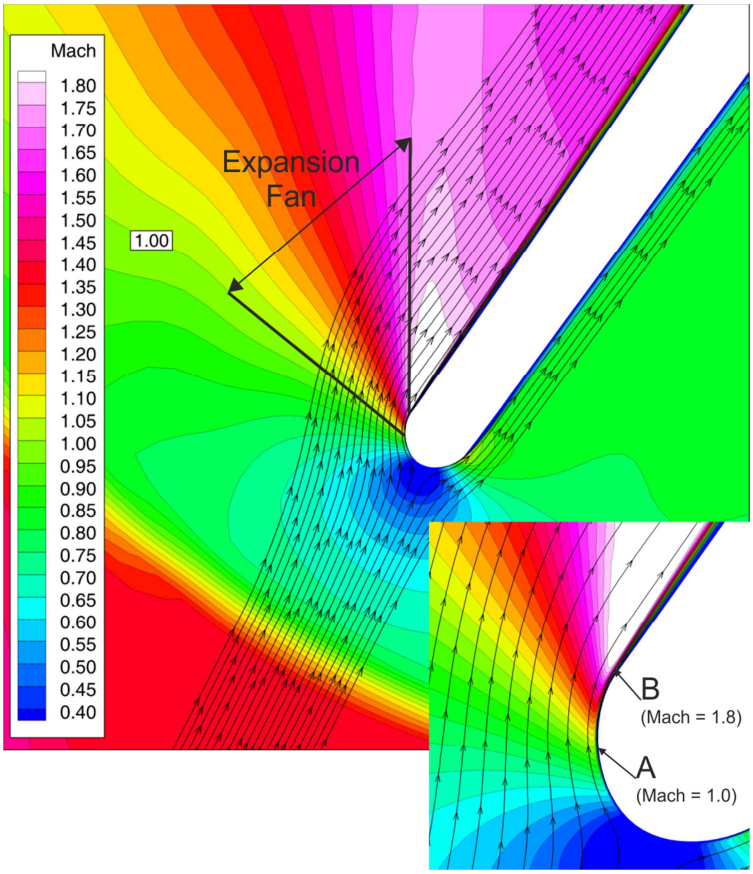


Figure 3.9: Contours of Mach of the leading edge region of the PAV 1.5 cascade at inlet Mach number = 1.457.

A change in the suction surface curvature appears at mid-portion, from concave to convex, inducing reacceleration of the flow to a Mach number of 1.52, before it encounters a shock system at around 60% chord. As it will be clear from the discussion of the cascade optimization in section §4.6, the suction surface curvature of the forepart is critical for the value of the Mach number. In fact, a concave curvature has the advantage to diffuse the supersonic flow, while a convex surface operates in the opposite direction. From this point of view, it is harmful to change curvature (i.e., from concave to convex) before the supersonic flow encounters the passage shock. In the case of PAV 1.5 cascade this was done on propose in order to have a higher pre-shock Mach number and thus a strong shock wave/boundary layer interaction.

In Schreiber's investigation [**89**], a proper value of static backpressure was imposed to achieve a strong boundary layer shock interaction, giving rise to the so-called *lambda shock system* (see Figure 3.10), which is composed by a leading oblique shock and a rear strong curved oblique shock. Such particular pattern can

be interpreted as a *Mach reflection phenomenon*, in which the intensity of the oblique passage shock increases to such a level that a regular reflection of the oblique passage shock is not possible (Schreiber and Starken 1992 [**92**]). In lieu of a simple Mach reflection, near the suction surface the oblique passage shock is replaced with a quasi-normal shock, whose interaction with the boundary layer gives rise to the lambda shock system, in addition to a severe boundary layer separation. It has been observed that the lambda shock develops when the oblique shock passages impinges on a relatively strong convex curved part of the profile surface.

As for the pressure surface, the convex curvature in the front region accelerates the flow until the 20% chord where a quasi-normal shock reduces the pre-shock Mach number of 1.15 to subsonic values. From this point to the trailing edge, the flow is subsonic and it is not clear how the pressure shape affects the flow diffusion.

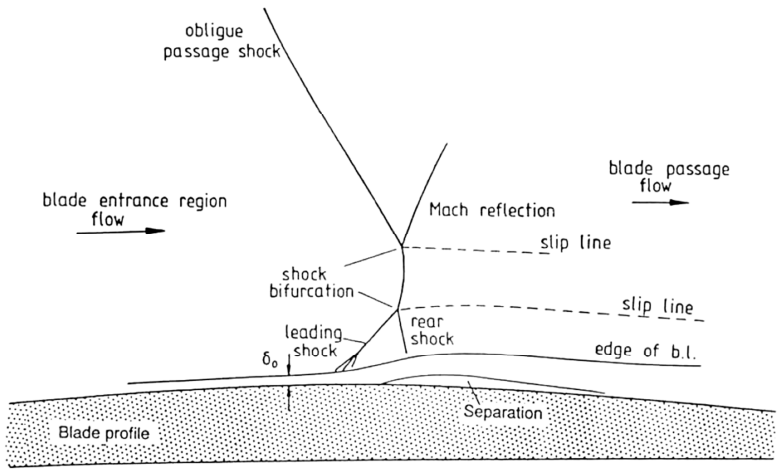


Figure 3.10: Flow structure of strong interaction in a cascade blade passage (from Schreiber [**89**]).

3.3 Supersonic Cascade Inlet Flow

The inlet Mach number M_1 is a fundamental parameter to discriminate different flow configurations of linear supersonic cascades. Beside the inlet Mach number itself, it is of paramount importance the axial component of the inlet Mach number M_{1x} , as illustrated in Figure 3.11. Two different flow configurations exist:

1. As long as the inlet axial Mach number is supersonic $M_{1x} > 1$, the upstream flow field is not influenced by disturbances emanated from the cascade. The shock and expansion waves caused by the leading edges and the suction surfaces are enveloped by the blade passages (left sketch in Figure 3.11).
2. On the other hand, with a subsonic axial Mach number (but supersonic upstream Mach number $M_1 > 1$) the perturbations coming from the cascade propagates to infinity in the upstream direction and influence the incoming flow. This second type of flow is of practical interest for transonic compressor.

All issues related with the *starting* of supersonic cascades, unstarted flows, *unstarted* and choked flows and supersonic axial Mach number flows in supersonic cascades are discussed in several references, as Lichtfuss and Starken 1974 [**79**] and Schreiber *et al.* [**93**]. For brevity, the focus of this dissertation is only devoted to *started* supersonic cascades with subsonic axial Mach numbers and subsonic exit Mach number, due to their appealing for transonic compressor rotor applications. In this mode, if the inlet Mach number is sufficiently high, the cascade operates along the so called *unique incidence* curve.

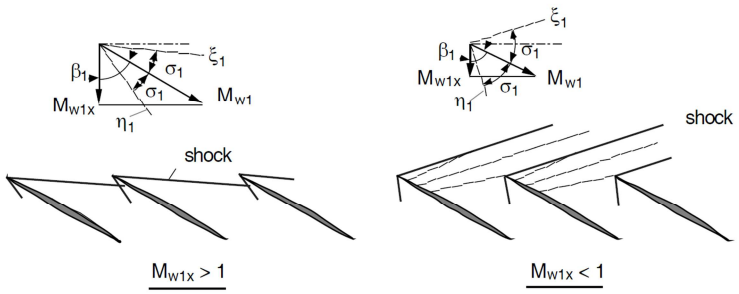


Figure 3.11: Supersonic flow in blade row (Bölch and Suter 1986) [**94**].

3.4 Unique Incidence flow

As far as the supersonic cascade with inlet subsonic axial Mach number is started, the inlet Mach number and flow direction are dependent one another, at least within a range of static back pressures and inlet Mach numbers. It is not possible to change M_1 without not affecting β_1. The cascade operation is possible only along the so called *unique incidence curve*.

The unique incidence behavior is the result of three conditions:

1. subsonic axial Mach number $M_{1x} < 1$, for which the information can travel upstream along the axial direction; the existence of the cascade affects the incoming flow by means of weak compression and expansion shock waves, which mainly depends on stagger angle, solidity and suction surface profile.
2. Honor the cascade *periodic condition* at inlet, namely, the approaching Mach number and flow direction must repeat among all profiles.
3. The idealized passage shock wave is attached to the leading edge of the airfoil and meets the suction surface of the adjacent one, serving as a "stopper" for information coming from downstream of the cascade.

How these three conditions imply the unique incidence flow can be rigorously demonstrated applying the characteristic method to cambered profile with sharped or blunt leading edge profiles (Lichtfuss and Starken 1974 [**79**]). Nevertheless, a simple reasoning on flat-plate cascade could intuitive convince the reader on the veracity of the existence of the unique incidence flow.

3.4.1 On The Existence Of The Unique Incidence

When infinite flow (M_∞, β_∞) approaches a semi-infinite flat-plate cascade with an inlet angle β_∞ diverse than the stagger angle γ, two different wave patterns exist depending on the incidence, positive or negative, of the incoming flow over the first (lowest) blade, as illustrated in Figure 3.12. In the case of positive incidence, Prandtl-Meyer expansion waves centered on the leading edge of the first blade develop in the upper region, and accelerate the flow up to M_1 and turn it into the flat plate direction $\beta_1 = \gamma$ (Figure 3.12 - a). Due to subsonic axial Mach number, this expansion fan perturbs the flow ahead of all the other blades, which experience a uniform inlet flow (M_1, β_1). From a mathematical point of view, the left-running Mach lines emanate from the suction surface of the first blade extend in front of all the other blades. Therefore, downstream of the first blade, the incoming flow is characterized by (M_1, β_1) and approaches all the other blades with null incidence.

As for negative incidence, at the leading edge of the first blade the Prandtl-Meyer fan is replaced by a compression shock wave, which is in charge to turn the flow up to $\beta_1 = \gamma$ and decelerate the Mach number to M_1 (Figure 3.12 – b).

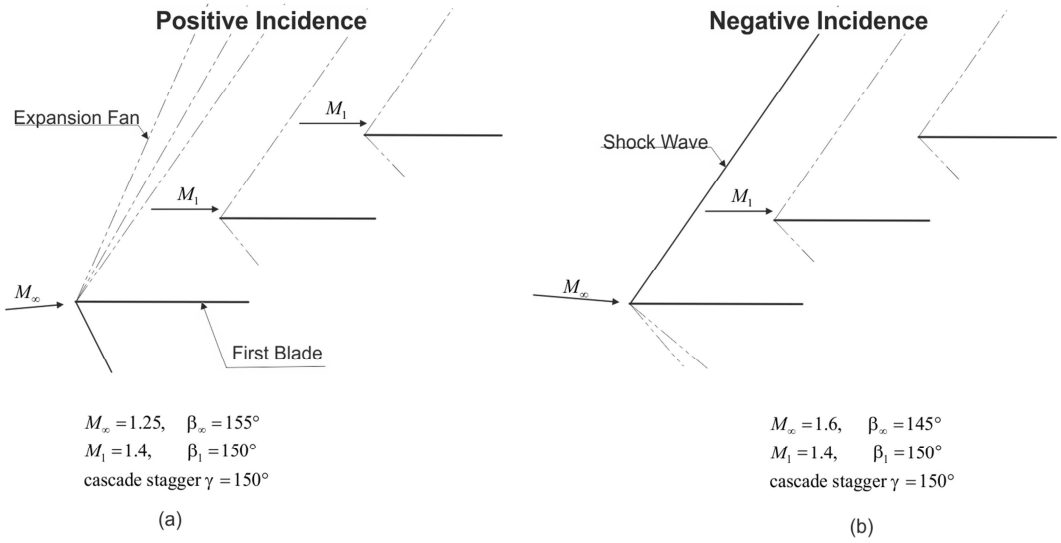


Figure 3.12: Semi-infinite flat-plate cascade with subsonic axial inlet-flow Mach number.

As a result, for a semi-infinite cascade the blades can be distinguished into two groups according to inlet flow conditions:

1. the first blade, which experiences an inlet flow of (M_∞, β_∞) and sets the incoming flow condition (M_1, β_1) for the remaining blades; it assures the periodic condition for the entire cascade.
2. All the blades except the first one, for which an inlet flow of (M_1, β_1) is applied.

What happens if the infinite flow angle β_∞ varies? The first blade adapts the shock system centered on its leading edge in such a way to turn the flow into the direction of the flat plate, i.e. $\beta_1 = \gamma$. Even though the angle of the infinite flow changes, the incoming flow applied to the second and all the other blades has always the same direction, owning to the straight plates.

An infinite cascade has no first blade, so that in this case only (M_1, β_1) is possible as periodic solution of the inlet flow. The unique incidence relationship between M_1 and β_1 is of the form $\beta_1 = \gamma, \forall M_1$, which is valid as long as straight plates are used as blade profiles.

It is worth nothing that there is a substantial difference in the incoming flow between the semi- and infinite cascade. In the first case, the flow could assume values of (M_∞, β_∞) different than (M_1, β_1), at least in the upstream region before the flow meets the shock wave system emanated from the first blade. On the other hand, the infinite cascade influences the entire flow area upstream of the cascade

as shown in Figure 3.13: the left-running Mach lines, which emanate from the front part of the blade suction surface and go in front of the cascade, influence the inlet flow and, consequently, are the reason for the existence of the unique incidence.

Although simplicity of flat-plate profile cascade, moving to S-Shape supersonic cascades makes the analysis more complicated and changes the shape of the unique incidence curve. The inlet-flow behavior can be summarized in a diagram, in which the inlet flow angle β_1 is plotted against the inlet-flow Mach number M_1. In Figure 3.14 and Figure 3.15 a comparison of flat-plate and S-shape cascade is reported. A very detailed explanation of the unique incidence phenomenon is reported in several references, e.g. Levine(1957) [**95**], Novak(1967) [**96**], Lichtfuss and Starken (1974) [**79**], York and Woodard(1976) [**97**], and Bölsc and Suter (1986) [**94**].

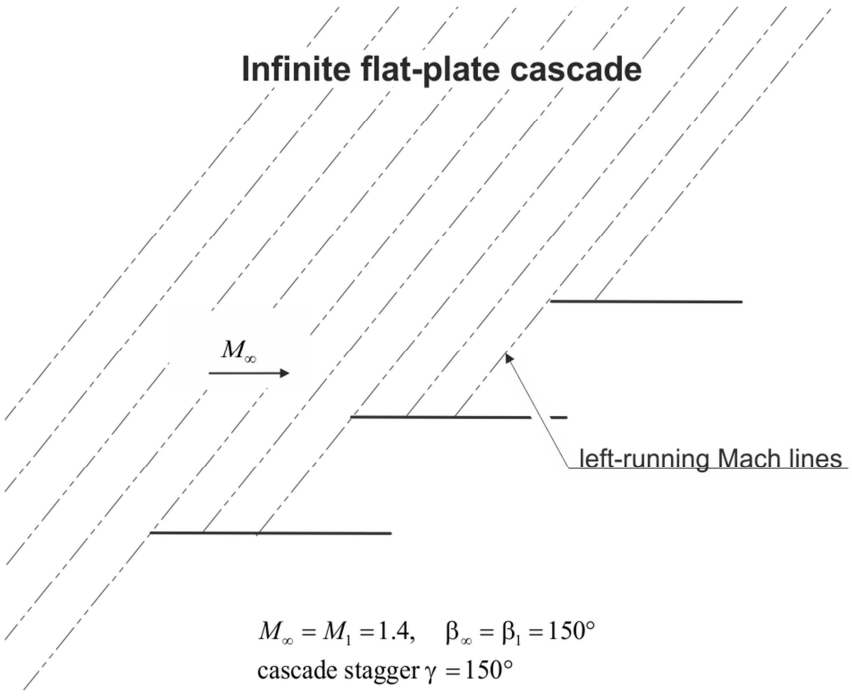


Figure 3.13: Infinite flat-plate cascade at subsonic axial inlet-flow Mach number.

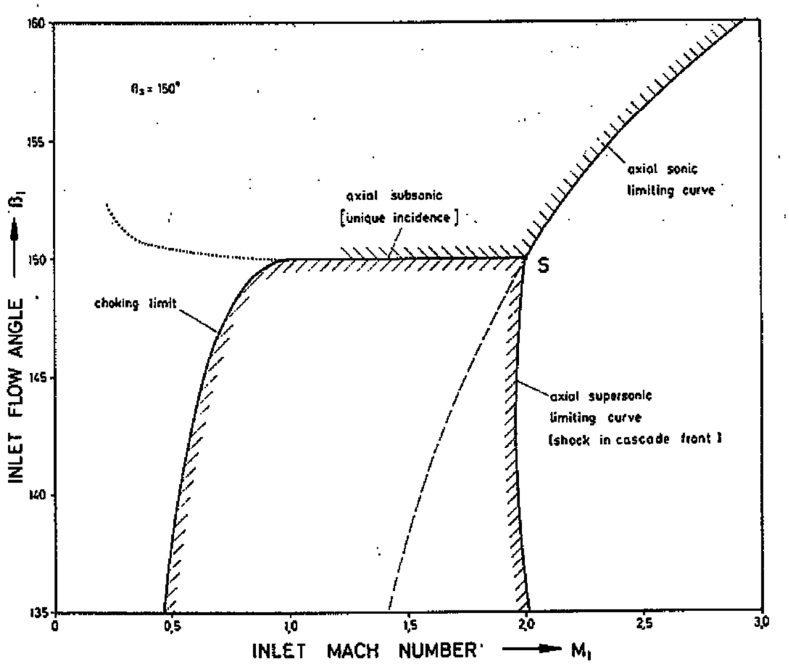


Figure 3.14: Inlet-flow angle boundaries of a flat-plate cascade (Lichtuff and Starken (1974) [**79**]).

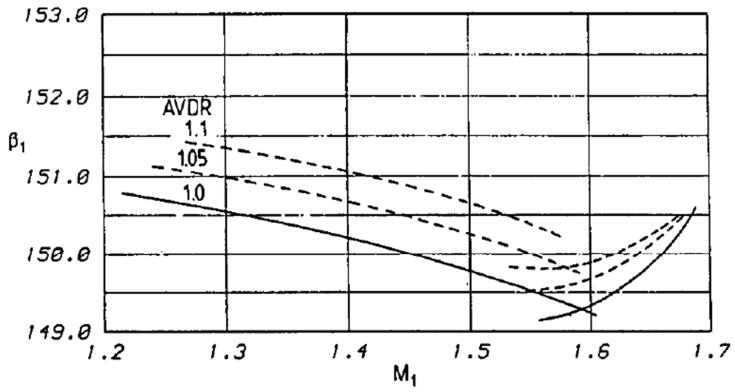


Figure 3.15: Inlet-flow angle boundaries of a S-Shape cascade (Schreiber (1996) [**90**]).

3.4.2 Unique Incidence Operating Condition

The unique incidence relation holds within a confined operating range of Mach numbers and static back pressures, for which the existence of an attached shock wave at the leading edge is possible. As depicted in Figure 3.14 the curve starts from an inlet Mach number around 1.1 and ends at point S, where the axial Mach number reaches unity. Beyond S point the cascade has no influence on the upstream flow, thereby the Mach number and flow direction are independent, within certain limits.

In Figure 3.16, the schematic flow field well reproduces what happens at any point of the unique incidence curve when the static backpressure is increased and the inlet flow conditions are fixed.

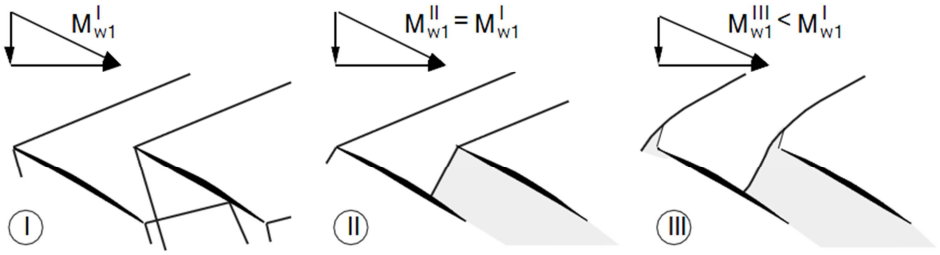


Figure 3.16: Supersonic flow in a compressor cascade for different operating points (Bölch and Suter 1986 [**94**]).

For as far as the shock wave stays inside the cascade passage, this is condition (I) and (II) of Figure 3.16, the change in static backpressure affects only the shock pattern within the blade passage and exit plane, whereas the upstream flow field is not influenced. Precisely, the increase in static backpressure forces the shock wave to move towards the passage entrance. There is a limit in static backpressure for which the shock is exactly at the passage entrance (condition (II)); a further increase would give rise to a detached shock in front of the passage entrance (condition (III)) leading to the so called *unstarded or spill condition*, in which the static backpressure has an influence on the inlet flow. The unique incidence does not hold any more and a new relation between the inlet Mach number and flow direction is established and, unlike unique incidence, it becomes parametric with the static backpressure. The operating condition (II) is the inception of the unique incidence and represent the highest static pressure ratio obtainable for a cascade working in the unique incidence regime at given inlet Mach number. For lower inlet Mach numbers, the unique incidence relation holds until a specific inlet Mach number below which an attached shock wave at the entrance passage is no more possible, irrespective of the static backpressure. Although a sharp leading edge is concerned, the leading edge wedge angle requires a minimum Mach number for the shock to be attached. Under this value, the supersonic cascade works in the unstarted mode.

In the unique incidence operations, the cascade is chocked. Each point of the unique incidence curve is characterized by a specific mass flow. In fact, the

variation in static back pressure does not affect the inlet flow, thereby the mass flow remains unchanged.

3.4.3 Experimental And Numeric Considerations

Generally, a supersonic cascade adopted in a wind tunnel experiments can be assimilated with a semi-infinite cascade. The cascade first blade establishes the appropriate periodic inlet flow condition (M_1 , β_1) for the remaining blades according to the unique incidence relation. Figure 3.17 shows a sketch of a wind tunnel test section.

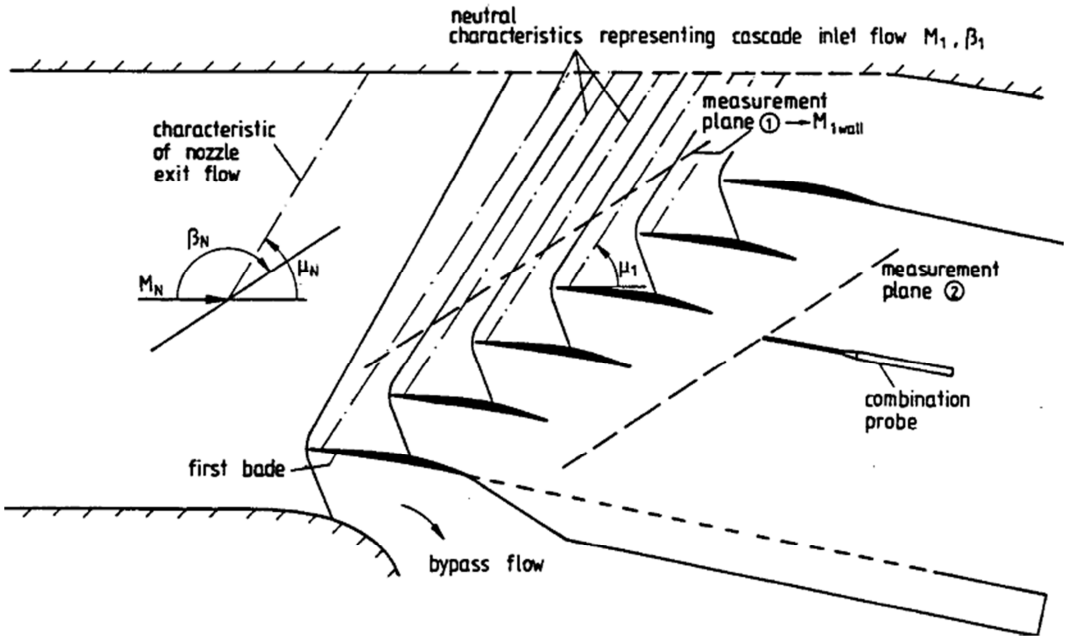


Figure 3.17: Nozzle exit flow and cascade inlet floe with periodic wave pattern behind neutral characteristic of first blade (Schreiber *et al.* (1993) [**93**]).

The incoming flow could be considered with good approximation two-dimensional, irrotational, and isentropic up to the strong shock wave inside the blade passage. In fact, supersonic blades are usually characterized by thin leading edges and low front chamber, which minimize total pressure losses related with the inlet wave pattern. This assumption is corroborated by experimental results as those reported in Figure 3.18, which justify the outstanding performance of simplified numerical methods, for instance those based on *simple-wave theory*, in the prediction of the inlet flow region ahead of the passage shock wave.

Under the assumption of irrotational and isentropic flow, the inlet region obeys to the *simple-wave theory* and, thereby, to the Prandtl-Meyer relation, also known as Riemann invariant,

$$\beta_\infty + \nu(M_\infty) = \beta_1 + \nu(M_1)$$
$$\nu(M) = \left(\frac{k+1}{k-1}\right)^{0.5} \arctan\left[\frac{k-1}{k+1}\left(M^2-1\right)\right]^{0.5} - \arctan\left[M^2-1\right]^{0.5} \qquad \text{Eq. 3.2}$$

, which establishes the existence of an invariant quantity between two points of the flow domain, for instance at the nozzle exit and ahead of the cascade entrance plane.

Another consequence of these hypotheses is the existence of *characteristics*, curves along which the Mach number and flow angle are constant. As for supersonic cascades, characteristics usually have a linear pattern as that reported in Figure 3.17. Among all characteristics, the *neutral characteristic*, emanating from the front portion of each blade, is a special curve because it represents the cascade inlet flow periodic condition (M_1 , β_1). All Mach lines being upstream of the neutral inlet characteristic attenuate the shock wave arising at the sane profile, whereas all downstream Mach lines interfere with the shock wave of the following blade.

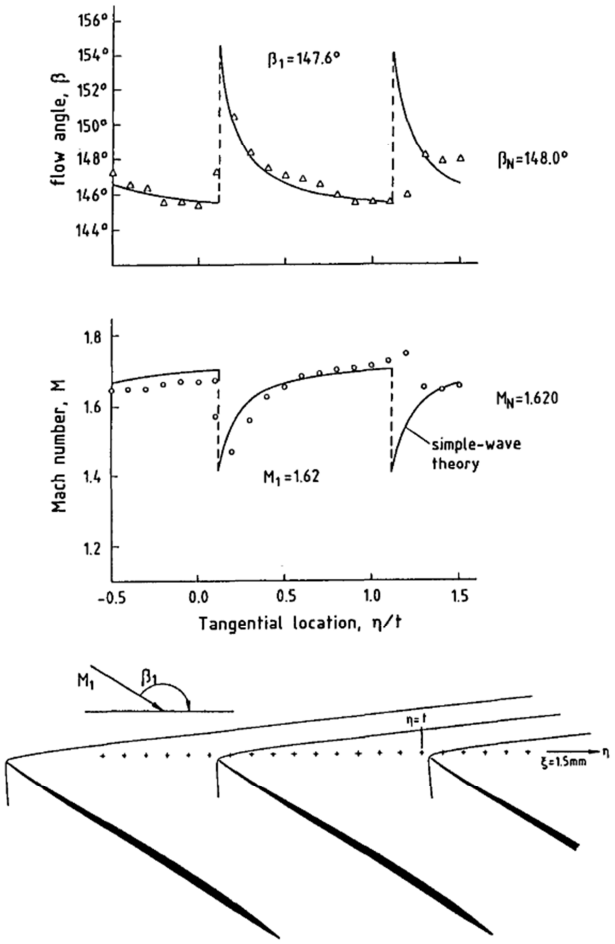


Figure 3.18: Flow angle and Mach number distribution in the entrance region at supersonic inlet flow conditions (Tweedt DL *et al.* (1998) [**98**]).

When the nozzle flow (M_∞, β_∞) differs from the unique incidence condition, the first blade adjust the flow with a compression or expansion waves system centered at the leading edge, depending on whether β_∞ is higher or lower than a β_1, respectively. Both turning mechanism adjust the flow, but with different repercussions on the assumptions of irrotational and isentropic flow. The expansion fan accelerates the flow in a *quasi-isentropic* way, whereby the compression shock waves deteriorates the flow entropy. It is to prefer the first mechanism because it guarantees the validity of the Riemann invariant assumption, which makes it possible to estimate the cascade inlet condition given the nozzle flow (M_∞, β_∞) and the unique incidence relation $\beta_1 = f(M_1)$. Although the unique incidence condition is not known *a priori*, the relationship can be obtained from theoretical calculations for the interesting Mach number range. Such unique incidence calculation also should include the leading edge blockage

effect and the shock losses of the entrance wave pattern (Starken *et al.* (1984) [**99**]).

3.5 Cascade Influence Parameters

3.5.1 Inlet Mach Number Influence

The relative inlet Mach number of a cascade is strictly connected with the wheel speed U of a compressor rotor, which indeed is the key parameter behind high specific energy transfer: the higher the relative inlet Mach number, the higher the static pressure ratio achievable in supersonic compressor cascades. The upper plot in Figure 3.19 shows the maximum static pressure ratio achieved with ARL-SL19 cascade over the inlet Mach number range 1.23 to 1.72; corresponding total pressure losses are reported in the lower plot. On the other hand, increasing static pressure ratio is made to the detriment of total pressure losses.

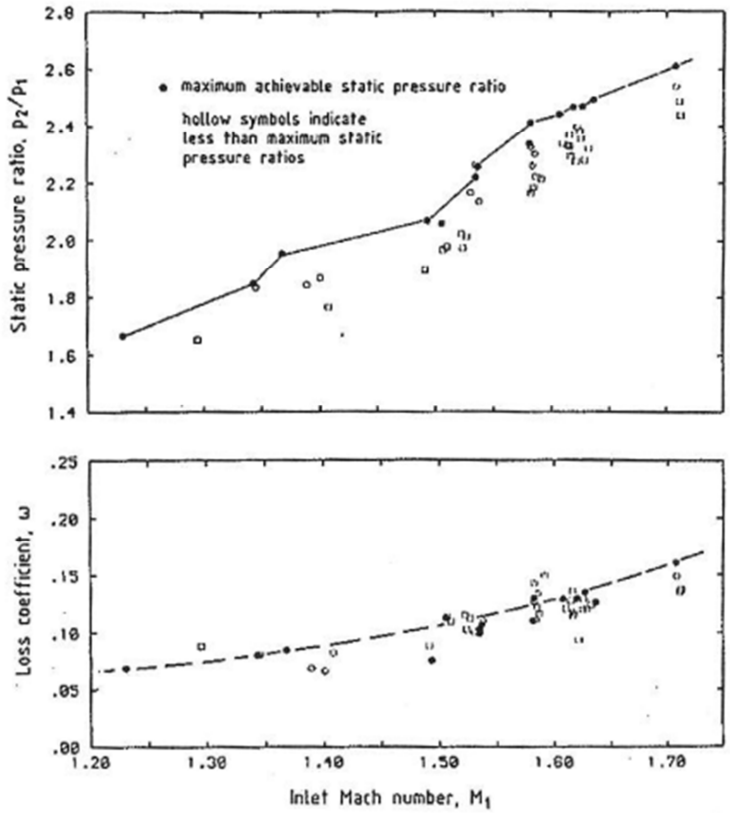


Figure 3.19: Influence of the inlet Mach number on the maximum achievable static pressure ratio and the corresponding total pressure losses.

3.5.2 Static Pressure Ratio Influence

On the unique incidence operating line, the variation of static backpressure (cascade throttling) affects only the flow inside the passage, modifying the shocks pattern and their interaction with the boundary layer, while the inlet region remains unaltered. According to Tweedt *et al.* (1988) [**98**], increasing static backpressure, from moderate to high static pressure ratios, causes an increment of losses. In particular, there is a reduction in the shock loss, but with corresponding increases in the viscous loss, which can be attributed mostly to a change from weak to strong suction surface boundary layer separation. The losses coefficient can be expected to be approximately 0.10-0.15. Moreover, in throttling the cascade the exit flow angle can vary at most by 2 or 3 [deg] and the exit Mach number shifts to supersonic values for low values of static back pressure .

3.5.3 AVDR Influence

The Axial Velocity Density Ratio (AVDR) has a strong influence on the total pressure losses, the flow exit angle and the shocks pattern. Increasing AVDR means to increase the spanwise stream tube convergence that, for a supersonic flow, tends to lower the Mach number in the passage, thereby reducing the shock loss and the losses from the shock boundary layer interaction region. The sensitivity of AVDR on total pressure losses strictly depends on the level of static backpressure, with the loss reduction being more pronounced at higher static pressure ratios. As for the exit flow angle (or flow turning since the inlet conditions are unchanged), an increase in AVDR always reduces the exit flow angle and thereby increase the flow turning. The variation of the shocks pattern with AVDR is depicted in Figure 3.20.

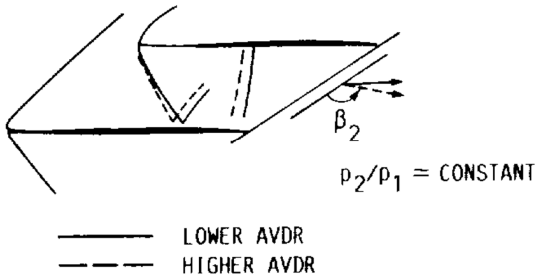


Figure 3.20: Influence of AVDR on shock wave position at constant back pressure (Schreiber and Starken (1992) [**92**]).

3.6 Relation Between Linear Cascade And Rotor Blade Element

The experiments performed in wind tunnel are usually carried out at unique incidence condition because it is the condition which can be established most easily in linear cascades with supersonic inlet flow. What happens to the behavior of a liner cascade foil when its profile is adopted inside a transonic compressor rotor? Could the experimental test on linear cascades be somehow useful during the design phase of 3D rotors?

Such problem was handled for the first time by Mikolajczak *et al.* [**82**]. Performance of three airfoil shapes tested in linear cascade were compared to the performance of similar airfoils tested in rotor. In particular, a "J" profile, a circular arc CA profile and a multi-circular arc (MCA) profile were selected, keeping the same chord (3.75 in.), same camber (10 deg), and comparable maximum thickness. Results for the three blades are shown in Figure 3.21. The cascade results denoted by a solid line are taken at unique incidence condition for different static back pressures till the spill point. In symbols there is the rotor incidence which is essentially constant and in good agreement with cascade results for the MCA and CA cascade, at least for static pressure ratios near spill point. As for the "J" section, in the rotor this section chocked at a different incidence condition at all pressure ratio perhaps for the different values of AVDR between cascade and rotor test, highlighting the paramount importance of such parameter on blade performance. In fact, for similar values of AVDR both turning angle and total pressure losses are comparable.

Beyond the spill point, the compressor rotor seems to be able to operate at higher static pressure ratios than the cascade, representing a subcritical operation beyond spill point since the periodicity requirement is automatically satisfied in the annular geometry of the rotor.

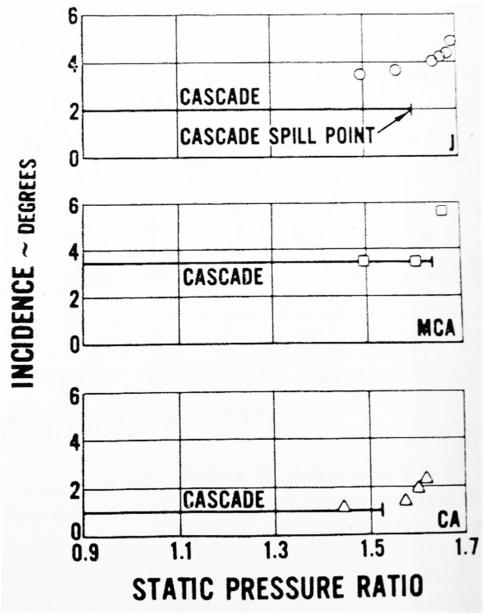


Figure 3.21: Incidence versus static pressure ratio for supersonic blades. The solid lines show results for cascades and are all at unique incidence condition, the points for results obtained in the rotor (Mikolajczak *et al*. [**82**]).

Also in the work of Schreiber and Starken [**100**] a comparison between rotor blade section and its homologous linear cascade was conducted for a Mach number in the range 0.82-1.1. Figure 3.22 shows the total pressure loss coefficient as a function of the inlet flow angle for different inlet Mach numbers. Both the cascade and the rotor section performance are plotted in the diagram. The overall loss behavior in terms of both the shape and the values of the loss curves are in good agreement. However, the various Mach numbers covered by the rotor due to its operating condition, in addition with the higher AVDR of the rotor tests, could explain the discordance of the choking angle. In conclusion, the experimental tests on 2D linear cascade are a convenient tool to gain insight on the complex transonic axial compressor blade element within the whole operating range.

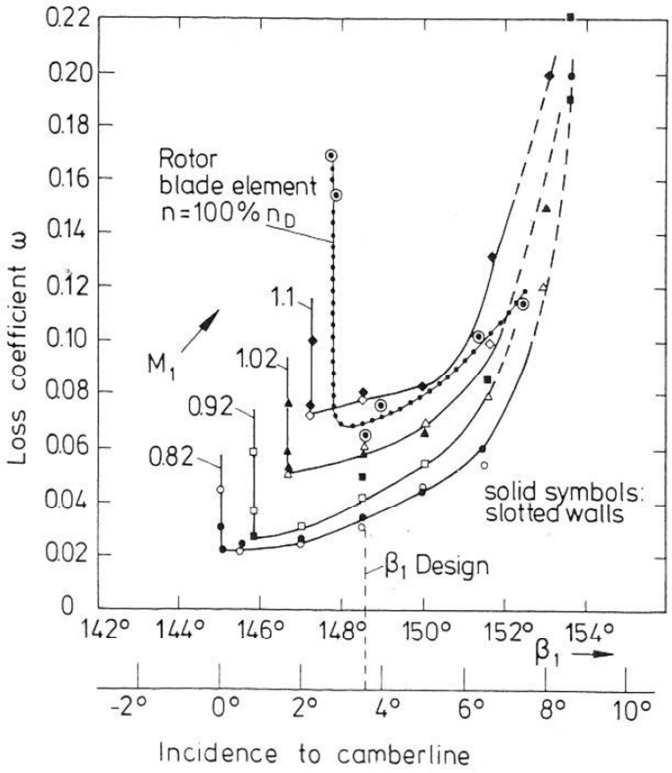


Figure 3.22: comparison of total pressure loss coefficient of rotor blade element and cascade (AVDR of the cascade tests 1.1 -1.17).

Chapter 4

Shape Optimization of a Supersonic Compressor Cascade

The DLR-PAV-1.5 supersonic compressor cascade tested at DLR by Schreiber [**89**] is a very well documented example of blading featuring the *pre-compression mechanism* and it has been taken as *baseline geometry* for the scope of this work. The cascade was especially designed for investigations on strong shock-wave boundary layer interaction. The geometry is typical for a tip section of a highly loaded transonic fan operating with an axial Mach number of 0.6 and a relative inlet Mach number of 1.5. The cascade had to provide a static pressure ratio of more than 2.0 with little flow turning.

4.1 Baseline Cascade

4.1.1 Airfoil Geometry

The shape of the airfoil employed in the DLR-PAV-1.5 cascade is given by points in Schreiber and Starken [**92**] and reported in Table 4.1. Unfortunately, in the open literature pressure and suction sides are reported with only 32 points for each side; moreover, both leading and trailing edges are missing. It must be stressed that the lack of information, at least for the leading edge zone, is of major concern due to the importance of such region in establishing the first shock wave, thereby influencing the overall cascade performance. This matter and its repercussion will be discussed more in depth in the grid validation section §4.2.4.

In order to heal the gaps at the front and rear part of the foil, leading and trailing edges were reconstructed with cubic splines safeguarding the continuity of first and second order derivatives at the junction points, i.e. the extreme points of the suction and pressure side. As a result, the reconstruction of the baseline foil

looks smooth, as it is clear from Figure 4.2, in which a close-up of leading and trailing edge is illustrated.

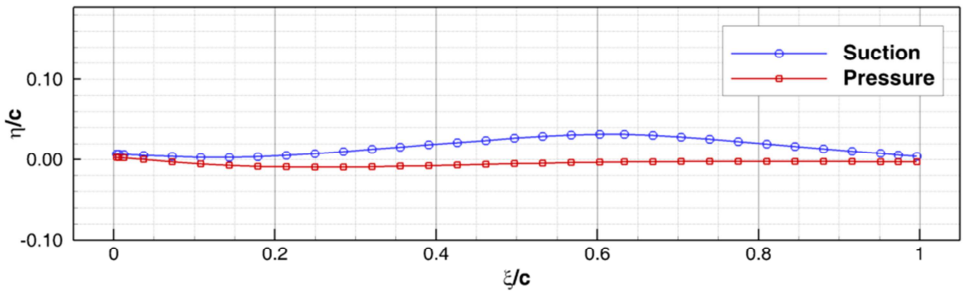


Figure 4.1: Geometry of DLR-PAV-1.5 supersonic cascade (Schreiber and Starken [**92**]).

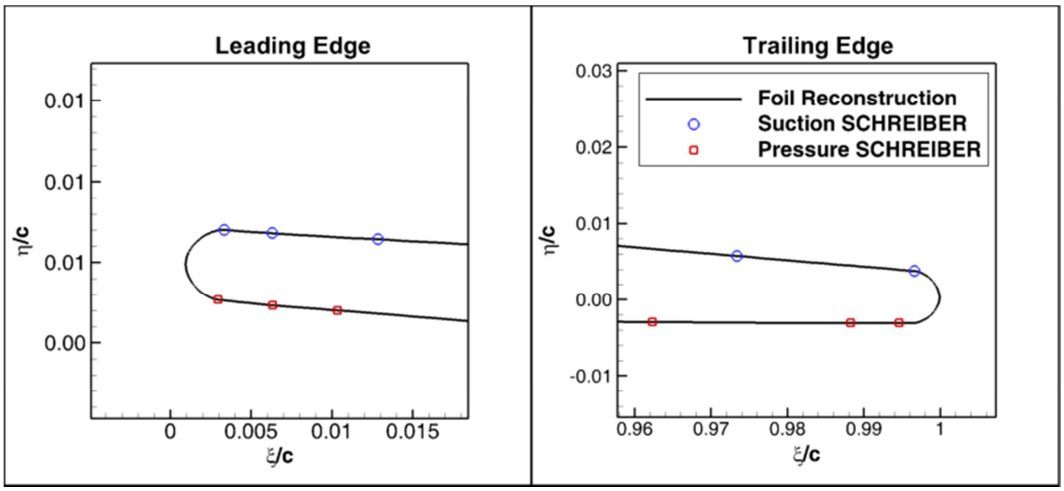


Figure 4.2: Reconstruction of the DLR-PAV-1.5 cascade: close-up of leading and trailing edge.

Suction		Pressure	
x/c	y/c	x/c	y/c
0.00334	0.00702	0.00294	0.00267
0.00631	0.00681	0.00633	0.00233
0.01286	0.00643	0.01034	0.00201
0.03720	0.00518	0.03681	-0.00024
0.07257	0.00372	0.07205	-0.00315
0.10796	0.00277	0.10725	-0.00566
0.14337	0.00242	0.14246	-0.00751
0.17883	0.00325	0.17768	-0.00874
0.21422	0.00510	0.21290	-0.00941
0.24959	0.00750	0.24815	-0.00968
0.28495	0.01017	0.28337	-0.00951
0.32030	0.01296	0.31862	-0.00911
0.35564	0.01578	0.35388	-0.00856
0.39097	0.01862	0.38913	-0.00790
0.42631	0.02138	0.42440	-0.00715
0.46166	0.02413	0.45966	-0.00632
0.49698	0.02685	0.49494	-0.00549
0.53227	0.02912	0.53022	-0.00467
0.56754	0.03078	0.56549	-0.00410
0.60277	0.03166	0.60960	-0.00354
0.63365	0.03171	0.64487	-0.00319
0.66884	0.03064	0.68015	-0.00289
0.70411	0.02834	0.71542	-0.00269
0.73940	0.02560	0.75069	-0.00260
0.77471	0.02250	0.78596	-0.00261
0.81005	0.01942	0.82122	-0.00258
0.84537	0.01644	0.85650	-0.00263
0.88070	0.01343	0.89177	-0.00270
0.91602	0.01054	0.92705	-0.00281
0.95135	0.00763	0.96231	-0.00294
0.97340	0.00576	0.98825	-0.00308
0.99660	0.00380	0.99459	-0.00307

Table 4.1: Airfoil geometry of DLR-PAV-1.5 supersonic cascade (Schreiber and Starken [**92**])

4.1.2 Cascade Geometry

After airfoil reconstruction, it follows scaling and rotation operations in order to meet the required cascade geometric parameters adopted by Schreiber [**89**]. In particular, chord c, pitch spacing s and stagger angle γ are prescribed according to Figure 4.3 and Table 4.2.

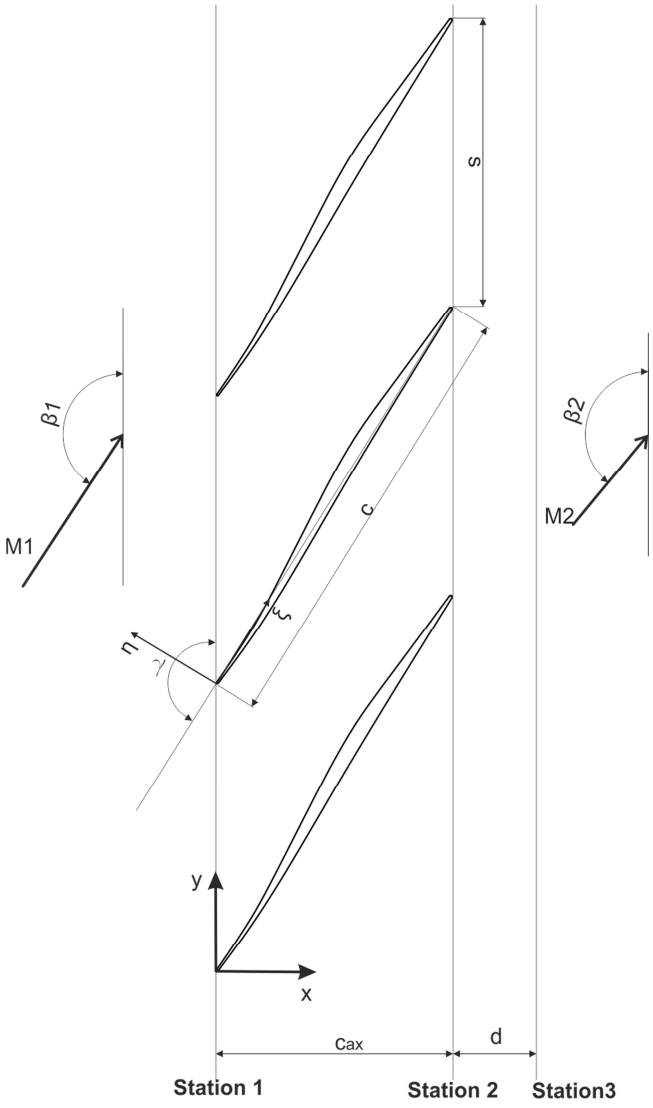


Figure 4.3:Definition of cascade geometric parameters.

Cascade Geometric Parameters		
Chord	c	170 [mm]
Picth-chord spacing	s/c	0.65
Stagger angle	γ	148.1 [deg]
Maximum thickness	$t_{\max}/c$	0.035
LE radius spacing	r_{LE}/c	0.0025

Table 4.2: DLR-PAV-1.5 geometric parameters.

4.2 Flow Solver And Computational Domain

In this section, the numerical model setup is described, which comprise the computational grid and the CFD solver setup.

4.2.1 Grid

The computational grid was carried out by means of ANSYS® ICEM, which is a powerful tool in the construction of structured grids; moreover, all procedures can be journalized, making its employment suitable within automatic optimization loops.

A multiblock grid with one O-grid around the blade was used to simulate the cascade passage. As depicted in Figure 4.4, the computational domain extends from $-1.88 \cdot c_{ax} < x < 2.77 \cdot c_{ax}$, where $x = 0$ corresponds to the blade leading edge. The length of the computational domain is slightly bigger than the one reported by Küster and Schreiber [**101**] and Sonoda *et al.* [**91**]. Moreover, the periodic edges were kept unchanged among all simulations, since deformations applied during optimization was very small compared with the spacing *s* dimension.

The grid quality was assessed with the quality metrics routines embedded in ANSYS® ICEM CFD [**102**]. All grids developed in this work respect the metrics limits summarized in Table 4.3.

The final grid dimensions are outcomes of the grid sensitivity analysis, which is reported in §4.2.4. As a result, a coarse grid of 23k elements were identified for the optimization, whereas a more refined grid of 77k were adopted for high resolution calculations at the end of the optimization procedure.

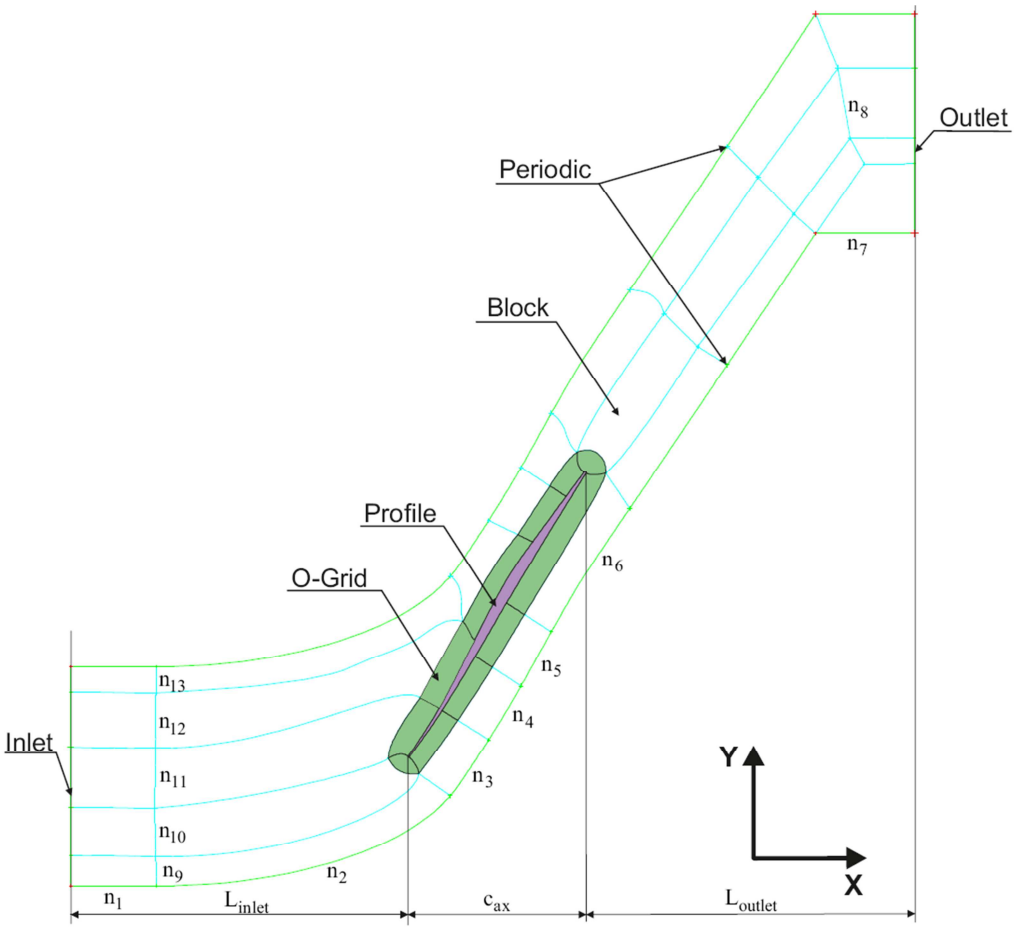


Figure 4.4: Computational grid topology.

ANSYS ® ICEM Grid quality metrics		
	min	max
Quality	0.5	1
Orthogonal Quality	0.5	1
Equiangle Skewness	0.3	1
Aspect Ratio	4e-04	1
Skew	0.3	1
Determinant	0.5	1
Min Angle [deg]	30	90

Table 4.3: Quality metric limits for grid quality assessment.

4.2.2 Flow Solver Setup

All calculations were carried out by means of ANSYS® Fluent v14 [**103**], in which the Reynolds Average Navier Stokes (RANS) equations coupled with a

turbulence model are solved adopting the finite volume method approach. Two-dimensional steady state simulations were performed for a fully turbulent compressible ideal gas in double precision. The main setups of the density-based solver are summarized in Table 4.4.

As for the boundary conditions (BCs), a *pressure-far-field* at the inlet with specification of turbulence intensity and length scale was prescribed, while at the outlet a *pressure outlet* BC was imposed. Blade walls were treated as no slip walls. The turbulence model is the two equation shear-stress transport k-ω model (kω-SST) proposed by Menter [**104**].

Convergence was established when all residuals went under 1e-06 and oscillation of the inlet and exit Mach number, flow angle and total pressure were below a certain threshold.

ANSYS© Fluent solver setup		
Formulation	Implicit	
Flux Type	Roe-FDS	
Spatial discretization	Gradient	Least squares cell based
	Flow	
	Turbulent Kinetic Energy	Second Order Upwind
	Specific Dissipation Rate	

Table 4.4: ANSYS® Fluent solver setup.

4.2.3 Calculation Of The Variables Of Interest

In order to avoid any kind of misunderstanding, the procedure employed to calculate the variables of interest (e.g., inlet and exit Mach number, flow angle, loss coefficient) and the survey stations are hereafter illustrated.

First, all global quantities, except for total pressure, were calculated with a *mass-weighted average* surface integral evaluated at the specific survey stations. On the contrary, total pressure calculation employed the *area-weighted average* surface integral.

Second, there are two survey stations referred as inlet and outlet stations. The former corresponds to the computational domain inlet (see Figure 4.4); whereas the outlet station is reported in Figure 4.3 as station 3 and it is located at 28% axial chord behind the trailing edge.

The loss coefficient was defined as follow:

Eq. 4.1

$$\omega = \frac{p_{01} - p_{02}}{p_{01} - p_1}$$

4.2.4 Grid Sensitivity Analysis And Validation

In this section, the main outcomes of the grid sensitivity and validation study carried out on the baseline cascade will be illustrated. Basically, the sensitivity study is aimed at obtaining a reliable CFD model while saving computational time and resources. Three grid sizes and two turbulence models were investigated.

As for the grid size, the number of elements was increased from 23k of the coarse grid up to 77k of the refined one; a medium grid size of 33k has also been tested (see Figure 4.5). A detailed list of the topology nodes distribution is reported in Table 4.5, whereas the topology variables are defined in Figure 4.4. Elements were added in those region considered critical for the physic phenomena, such as the zone right ahead of the cascade entrance, the leading edge, the fore passage zone where the oblique shock establishes, and the wake.

As for the turbulence models, the one equation approach developed by Spalart and Allmaras (SA) [**105**] and the two equation shear-stress transport k-ω model (kω-SST) proposed by Menter [**104**] were applied to each of the three grids. For both models, the first grid spacing normal to the blade surface was fixed to 0.0008 [mm] in order to yield values of $y^+ < 1$ and 15 nodes inside the physical boundary layer were guaranteed.

Variable description	Topology Var.	Coarse C1	Medium C2	Refine C3
Inlet 1	n1	5	5	10
Inlet 2	n2	27	36	65
PS fore zone	n3	30	35	60
PS middle1	n4	16	20	35
PS middle 2	n5	15	19	35
PS aft zone	n6	40	53	60
Outlet	n7	15	20	25
Trailing edge	n8	18	20	35
Spanwise lower	n9	15	15	27
Leading edge	n10	20	25	35
SS fore zone	n11	33	40	50
SS middle	n12	25	30	65
Spanwise upper	n13	16	16	27
O-grid layers	-	46	53	74
O-grid GrowthRate	-	1.2	1.16	1.1
O-grid height first layer [mm]	-	0.0008	0.0008	0.0008
	Total nodes	**23k**	**33k**	**77k**

Table 4.5: Topology parameters adopted in the grid sensitivity study. The variables "Topology Var." are defined in Figure 4.4

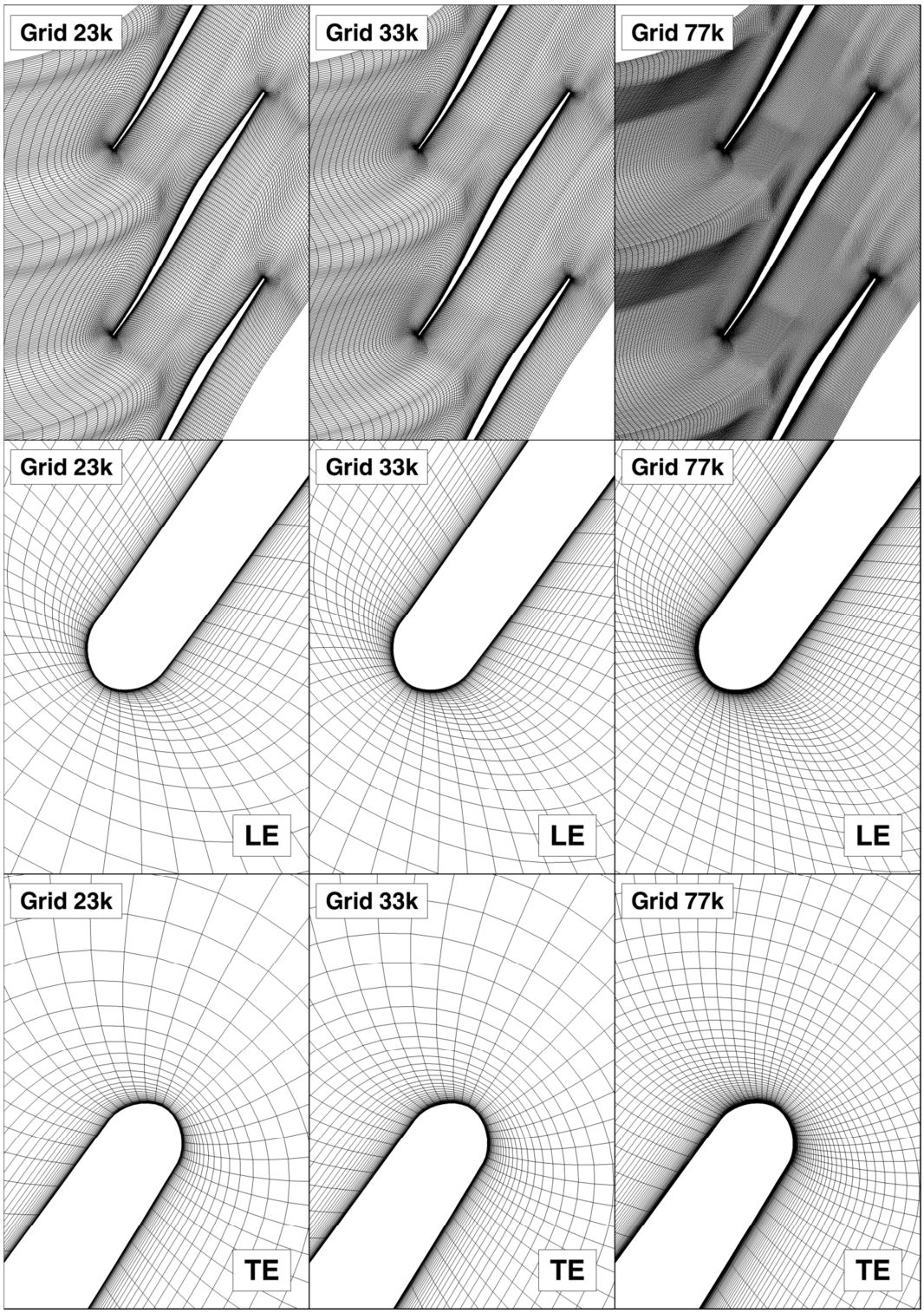


Figure 4.5: Computational grid of the baseline cascade passage. Comparison of three different grid sizes: an overall view of the passage (top), a close-up of the leading edge (middle) and a close-up of the trailing edge (bottom).

The CFD results were compared against Schreiber's experiment [**90**], in which the baseline cascade was operated at unique incidence condition with an inlet Mach number of 1.45 and a static pressure ratio of 2.2. Similar numeric calculations were also carried out by Schreiber [**101**] and Sonoda *et al.* [**91**]. It is worth nothing that Schreiber itself suggested that the best way to reproduce the flow field of a supersonic cascade is to set the exit pressure and the linear stream tube thickness variation (AVDR) in such a way that the boundary layer loading within the interaction region should be nearly identical for the experiment and the numerical simulation [**101**]. In other words, the numeric boundary conditions could slightly be adjusted with respect to the experimental ones, in order to obtain the best agreement of shock waves layout between numerical simulation and the experiment. Since the AVDR is not a parameter in our numerical model (i.e. AVDR = 1.00), only the Mach number and the static back pressure could be modified. The best results were obtained for an inlet Mach number of 1.456 and a static pressure ratio within the range 2.20-2.22. These values are also in good agreement with the analysis operated by Sonoda *et al.* [**91**].

In Figure 4.6 a sketch of the shock wave pattern by means of the magnitude of density gradient is reported for all six configurations (3 grids times 2 numerical models), whereas in Figure 4.7 the experimental pattern obtained from Schreiber [**90**] is illustrated. The pre-compression shock wave originated by the coalescence of the left-running characteristic, which are emanating from the concave forward portion of the blade suction surface, is well calculated by both turbulence models. However, it seems that kω-SST gives a better resolution of the intersection region of the pre-compression shock with the bow shock. As the gray scale indicates, the pre-compression shock is weaker than the oblique passage shock, but it makes the difference by reducing the Mach number ahead of the oblique shock wave from 1.75 to 1.35 (see Figure 4.8). The oblique shock wave departs from leading edge and meets the adjacent blade at about 60% of the suction surface; the shock is strong enough that the interaction with the boundary layer causes its separation and a peculiar lambda shock establishes above the separation region. The lambda shape is much well captured by the kω-SST model with the refined grid rather than SA computation. On the other hand, it seems that the Mach reflection branch is peculiar only of the SA calculation, being the Mach reflection just a blend gray band in the kω-SST visualization, even with refined grid. Moreover, the slope of the oblique shock is more prominent for the SA than the kω-SST.

Regardless the turbulence model, the isentropic Mach number distribution reported in Figure 4.8 is not affected at all by the grid size (Figure 4.8), except for the resolution around the shock wave peaks. Furthermore, the experimental behavior on the suction side from 22% to 50% of chord could be caused by experimental disturbances, which were reported by Schreiber and Starken [**92**], originated in the sidewall region of the test section and faintly influencing the blade mid-span test section. Except for this portion, the kω-SST seems to be in very good agreement with the experiments, whereas SA overestimates the isentropic Mach ahead of both the shock waves (i.e. oblique and quasi normal, see Figure 4.7) and it is even in pore accuracy in estimating the shock position.

In Figure 4.9 and Figure 4.10 the pitchwise distribution of exit flow angle and loss coefficient are reported. The experimental data were taken from Küsters and Schreiber [**101**]; unfortunately, the data is referred to the baseline cascade tested at different boundary conditions, i.e. inlet Mach number of 1.43, pressure ratio of 2.19 and AVDR 1.06. As a result, such data could be exploited only for a qualitative comparison. Effectively, the loss coefficient curves were aligned on the peak, and also the exit flow angle were ordered similarly. As for the exit flow angle, the variation inside the wake is not captured at all. Such behavior seems to be related with the mixing phenomenon that is inherently unsteady and not predictable by a steady state simulation [**101**]. As for the loss coefficient, the kω-SST captures the “bulge” patter on the right side of the bell that is originated by the shock wave boundary layer interaction (detachment + lambda shock). On the other hand, from the wide bell base it can be inferred that the SA diffuses the wake much more than the kω-SST; furthermore, there is no track of the “bulge” pattern.

The variations of cascade global performance as a function of grid size, turbulence model, and static pressure ratio, are summarized in Figure 4.11. The lack of knowledge of the leading edge geometry could be responsible for the important difference of the inlet flow angle, the prediction of which is overestimated of about 0.5 [deg] regardless grid size and turbulence model. As a result, also the outflow angle is shifted to higher values than experimental one, however the mean flow deflection is well captured by the SA model and slightly less by the kω-SST. As for the loss coefficient, the kω-SST estimation is in good agreement with the experiment, being the variation of 0.5% against the 5% of the SA model. It is worth nothing that the loss trend as a function of static back pressure is opposite for the two models. The reason of this depends on the

operating condition of the two cases. From a numerical point of view, it seems that the kω-SST case is beyond the spill point, in contrast with what is happening at the SA case, which indeed is working within the unique incidence range but with low static back pressure. This hypothesis is corroborated by tests operated applying the Unique Incidence Control Loop (UICL) to CFD calculations, in which the cascade is forced to work at the unique incidence point. The increase of static back pressure affects the kω-SST performance, which already was working beyond the unique incidence condition and now is forced to work at an even more unfavorable condition, whereas pushes the SA case nearer to the unique incidence condition, this explaining the increase in cascade efficiency.

Overall, the CFD model based on the kω-SST seems to well capture and predict in good accuracy both the local and global performance. The sensitivity on the grid dimension suggested that for optimization purpose the medium or even the coarse grid could be adopted, and that the refined one should be used for detailed analysis after the optimization is over.

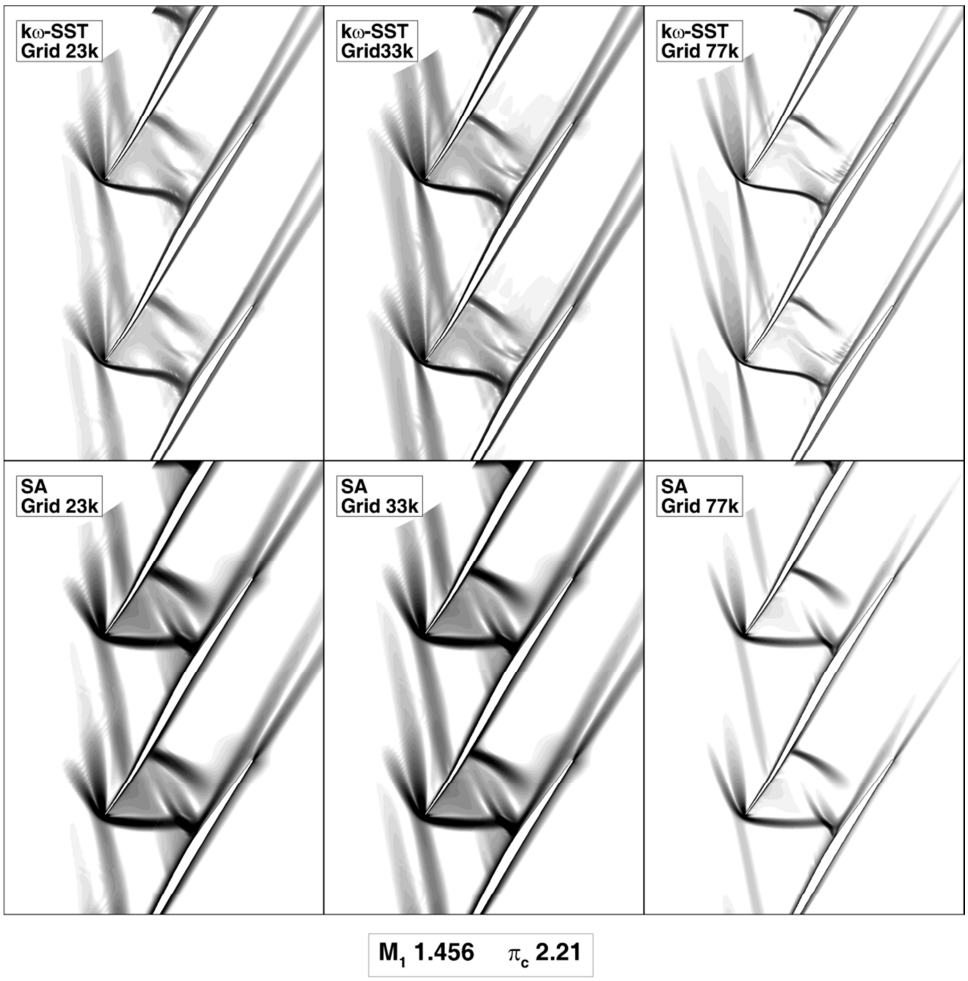


Figure 4.6: Density gradient magnitude for the baseline cascade evaluated via CFD. Comparison between different turbulence models and grid sizes. Same gray scale among images.

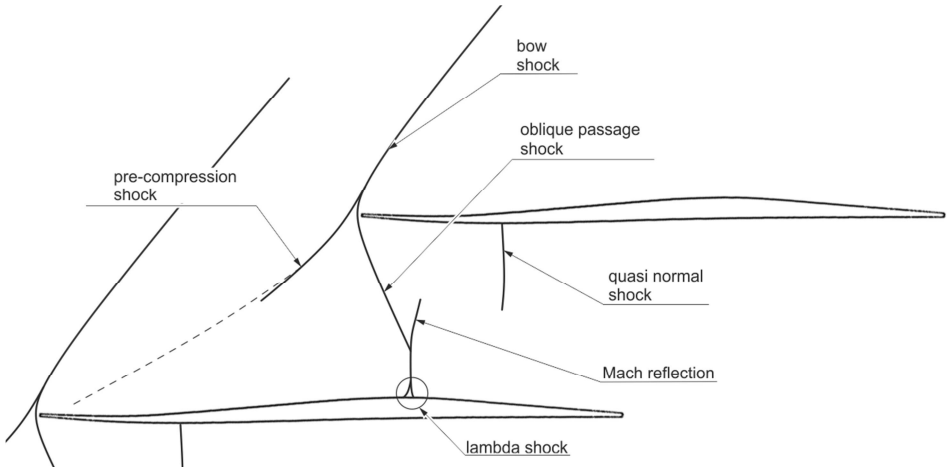


Figure 4.7: Experimental shock wave pattern for the baseline cascade at near design condition, M1=1.44, AVDR 1.01 and π_c 2.21, obtained from Schreiber [**90**].

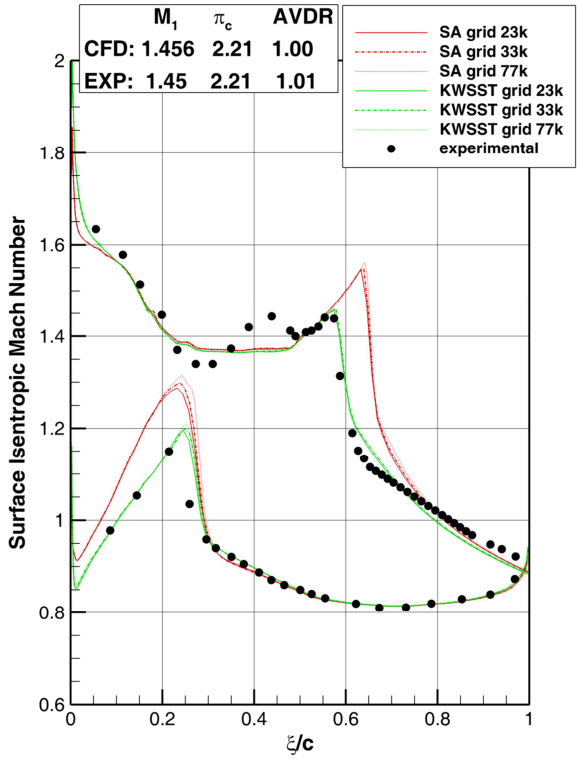


Figure 4.8: Isentropic Mach number distribution over the baseline cascade: the kω-SST against SA turbulence model for different grid sizes.

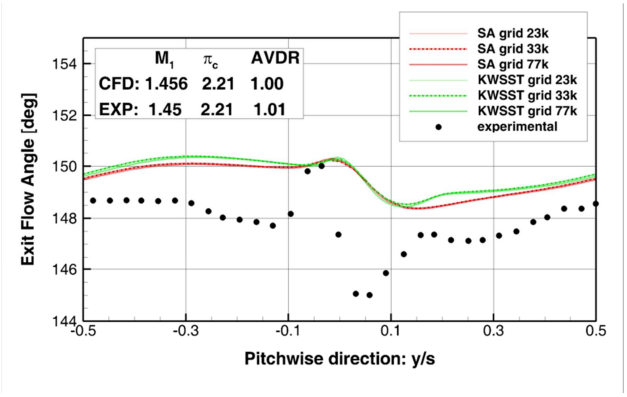


Figure 4.9: Pitchwise distribution of the outlet flow angle for the baseline cascade: the kω-SST against SA turbulence model for different grid sizes. The survey section is at 28% axial chord downstream the trailing edge

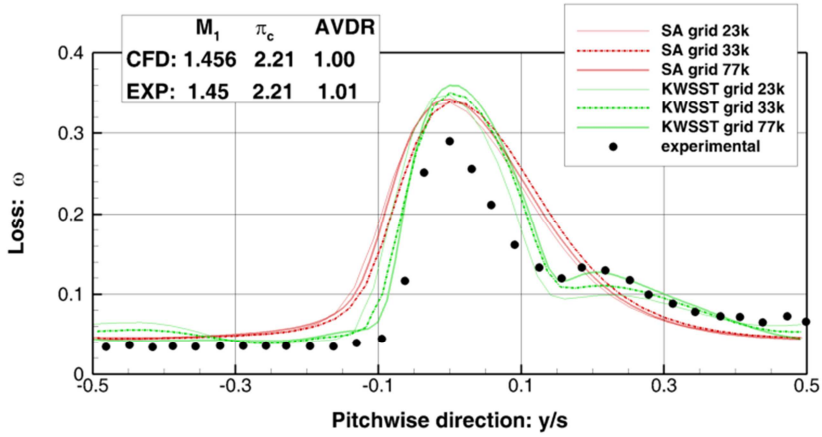
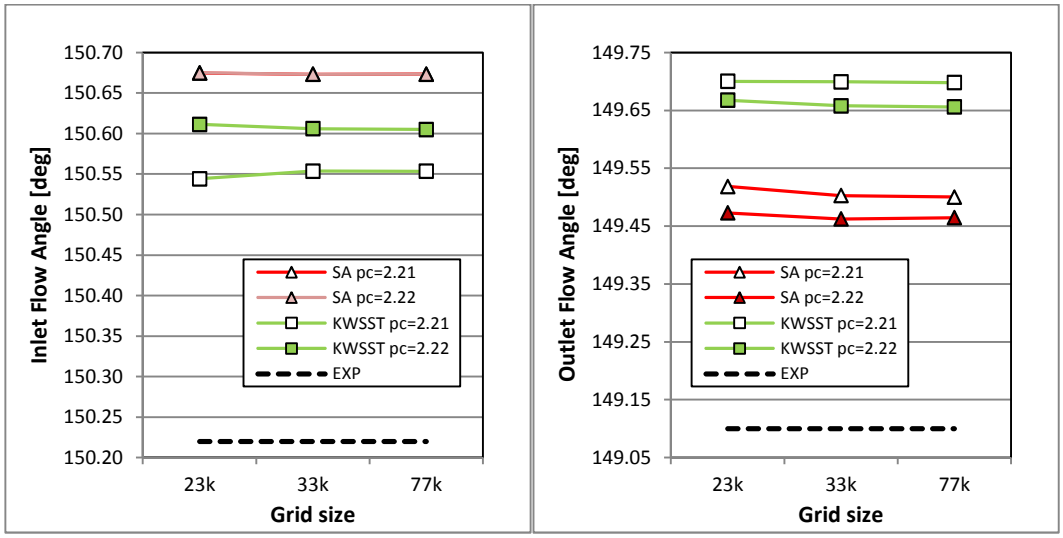
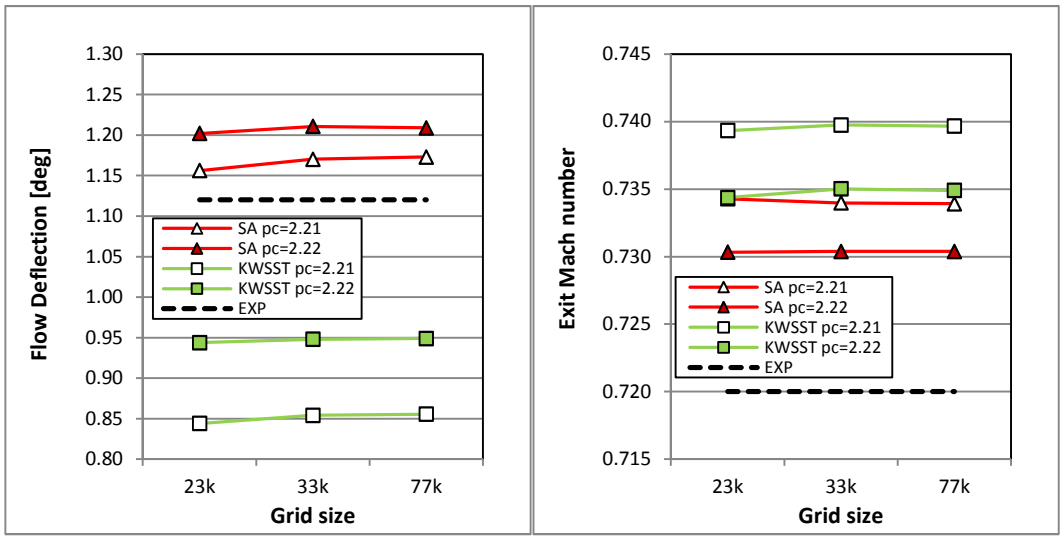


Figure 4.10: Pitchwise distribution of the loss coefficient for the baseline cascade: the kω-SST against SA turbulence model for different grid sizes. The survey section is at 28% axial chord downstream the trailing edge.


Inlet Flow Angle [deg]
150.70
150.65
150.60
150.55
150.50
150.45
150.40
150.35
150.30
150.25
150.20
23k
33k
77k
Grid size
SA pc=2.21
SA pc=2.22
KWSST pc=2.21
KWSST pc=2.22
EXP
Outlet Flow Angle [deg]
149.75
149.65
149.55
149.45
149.35
149.25
149.15
149.05
23k
33k
77k
Grid size
SA pc=2.21
SA pc=2.22
KWSST pc=2.21
KWSST pc=2.22
EXP


Flow Deflection [deg]
1.30
1.25
1.20
1.15
1.10
1.05
1.00
0.95
0.90
0.85
0.80
23k
33k
77k
Grid size
SA pc=2.21
SA pc=2.22
KWSST pc=2.21
KWSST pc=2.22
EXP
Exit Mach number
0.745
0.740
0.735
0.730
0.725
0.720
0.715
23k
33k
77k
Grid size
SA pc=2.21
SA pc=2.22
KWSST pc=2.21
KWSST pc=2.22
EXP

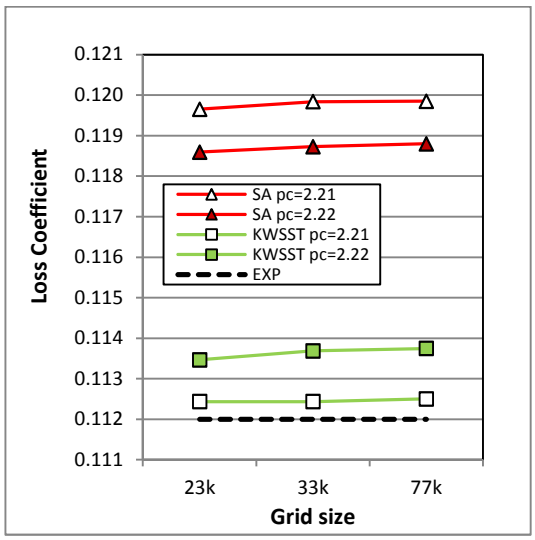

Loss Coefficient
0.121
0.120
0.119
0.118
0.117
0.116
0.115
0.114
0.113
0.112
0.111
23k
33k
77k
Grid size
SA pc=2.21
SA pc=2.22
KWSST pc=2.21
KWSST pc=2.22
EXP

Figure 4.11: Global performance of the baseline cascade as a function of grid size, turbulence model and static pressure ratio. The inlet Mach number is 1.456 and the static pressure ratio 2.21 for all calculations.

4.3 Unique Incidence Control Loop (UICL)

The unique incidence (UI) point is a well-defined operating condition at which supersonic cascades could gain the best performance in terms of cascade loss coefficient, for a fixed inlet Mach number. The UI point is within a flow regime where there exist a strong relationship between the inlet Mach number M_1 and the inlet Mach angle β_1. The reader is referred to §3.4 for a detailed explanation of UI condition.

The Unique Incidence Control Loop (UICL) is in charge to control the boundary conditions of CFD computation bringing the cascade to work at its UI condition, for a prescribed inlet Mach number. The UICL were developed in C code and linked with ANSYS® Fluent [**103**] by means of the User Defined Function (UDF) library.

4.3.1 Why UICL?

The need for the UICL originates directly from the impossibility to prescribe a specific value of (M_1, β_1) at the inlet boundary due to the physical behavior of a supersonic cascade ($M_1 > 1$) with a subsonic axial Mach number ($M_{1x} < 1$), for which the information can travel upstream along axial direction, thereby allowing the cascade to affect the incoming flow. Approaching the cascade, the flow at infinite (M_∞, β_∞) is deflected till (M_1, β_1) by a shock wave system released from the cascade fore region; deflection is essential in order for the incoming flow to honor the cascade periodicity (§3.4). This is the reason why (M_1, β_1) cannot be treated as a boundary constraint for the simulation, but must be considered as a result of the calculation.

From a numerical point of view, the *pressure far field* adopted as inlet boundary condition allows one to specify the incoming nonlinear supersonic Riemann invariant $\beta + \nu(M) = \cos t$ through the prescription of the flow at infinite (M_∞, β_∞). The Riemann invariant is the joining link between the infinite and the inlet boundary. Owning to the constancy of the Riemann invariant in the entire inlet region, the UICL can affect (M_1, β_1) by tuning (M_∞, β_∞). Be aware that a variation of the only M_∞ would affect the Riemann invariant, which, in turn, would have repercussions on both M_1 and β_1. A novel recursive procedure was

developed to find the correct values for (M_∞, β_∞) in order to achieve the required M_1, avoiding trial and error approaches.

4.3.2 UICL Strategy

For a given cascade geometry and total quantities at the inlet boundary (i.e., P_0 and T_0), there are only two variables which influences the UI point: M_1 (or β_1) and the outlet static pressure p_2. This holds under the hypothesis (1) of supersonic inlet flow with subsonic axial Mach number and (2) the cascade is *started* and does not work in *spill* conditions.

The UICL is based on a two-step strategy:

1. First, a very low value for p_2 is prescribed at the outlet; in such condition, the inlet flow domain is completely independent from the static back pressure p_2 and the UICL can search, regardless p_2, for an appropriate value of (M_∞, β_∞) in order to achieve the prescribed M_1. At this stage the solution (M_∞, β_∞) is unique and sets the starting point (M_1, β_1) for the following step.
2. In the second step, the main goal is the identification of a suitable value for p_2 in order to push the cascade to operate at the UI condition. Starting from the condition (M_1, β_1), p_2 is increased until a variation of the inlet quantities (M_1, β_1) is detected. The operating condition just before the inlet flow variation is observed, it is regarded as the cascade UI condition.

4.3.3 UICL Implementation

The UICL code implementation retraces the *status machine* paradigm. At the end of each CFD iteration, a UICL iteration begins: the CFD code transmits the flow domain to the UICL (Figure 4.12), which applies specific actions based on the analysis of the flow domain and, when necessary, modifies the domain boundary conditions; the UICL iteration ends returning the flow of control to the CFD code, ready to start a new CFD iteration.

The code is divided into two parts, the Mach Loop Status Machine (MLSM) in Figure 4.13 and the Unique Incidence Status Machine (UISM) reported in Figure 4.13 and Figure 4.14, respectively.

The *status 'S-#'*, which is specified at the right-top corner of each block, is stored and retained among UICL iterations, and drives UICL to the appropriate

block at the beginning of the next iteration. Beside the status variable, there is the Restart From Status (RFS) variable, which is an additional memory for the status machine. RFS is used in those circumstances in which a certain block needs to differentiate its own output *status*, depending on the calling block. For instance, the *CFD CONVERGENCE TEST* block in Figure 4.13 implements such strategy: when the conditions inside the block are satisfied, the output *status* changes according to the value of RFS.

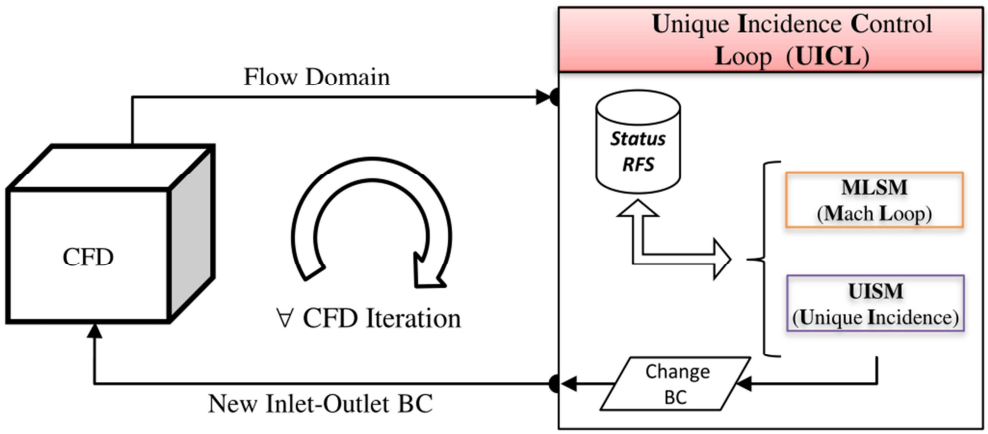


Figure 4.12: Framework flowchart of the Unique Incidence Control Loop (UICL).

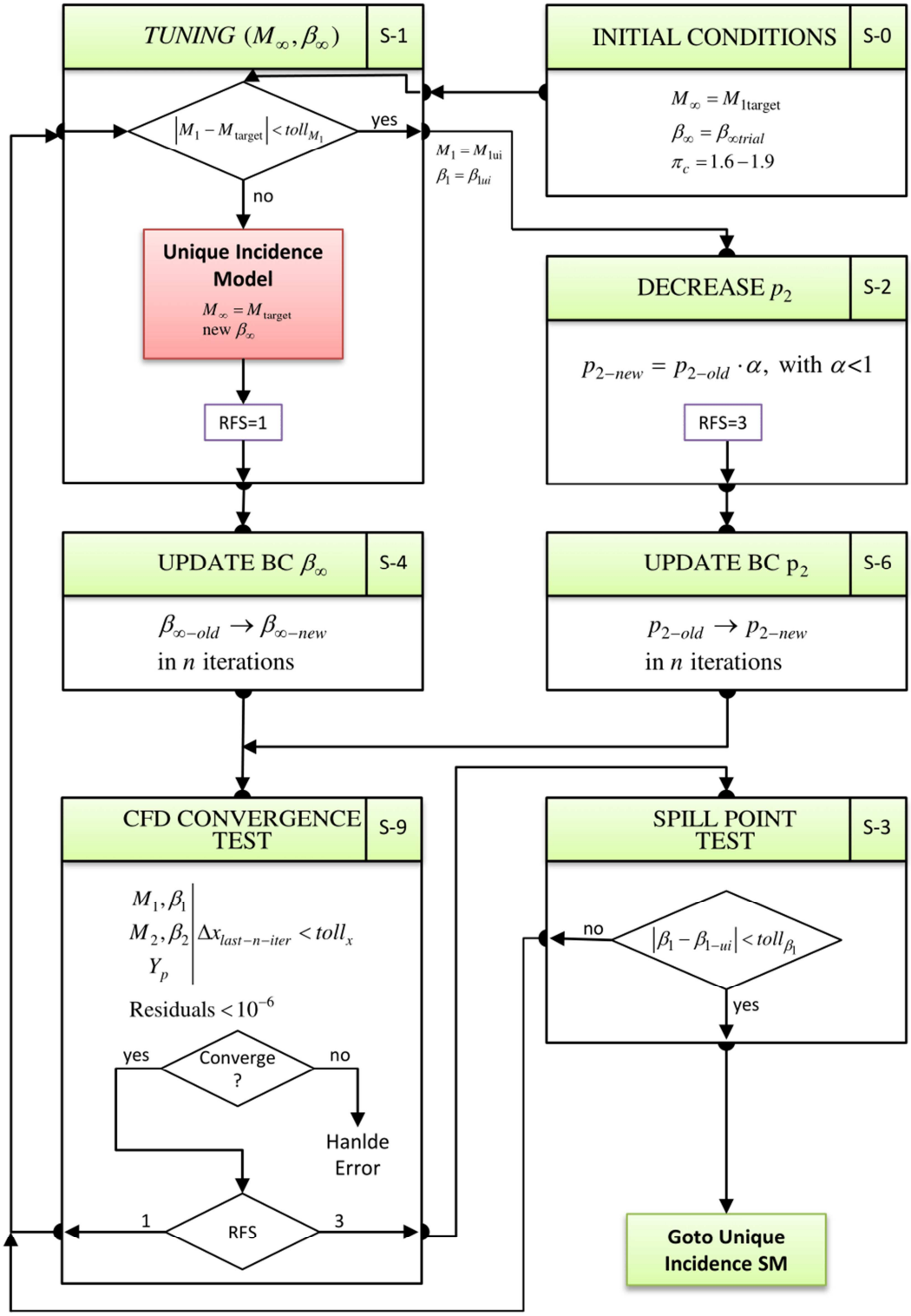


Figure 4.13: Flowchart of the Mach Loop Status Machine (MLSM).

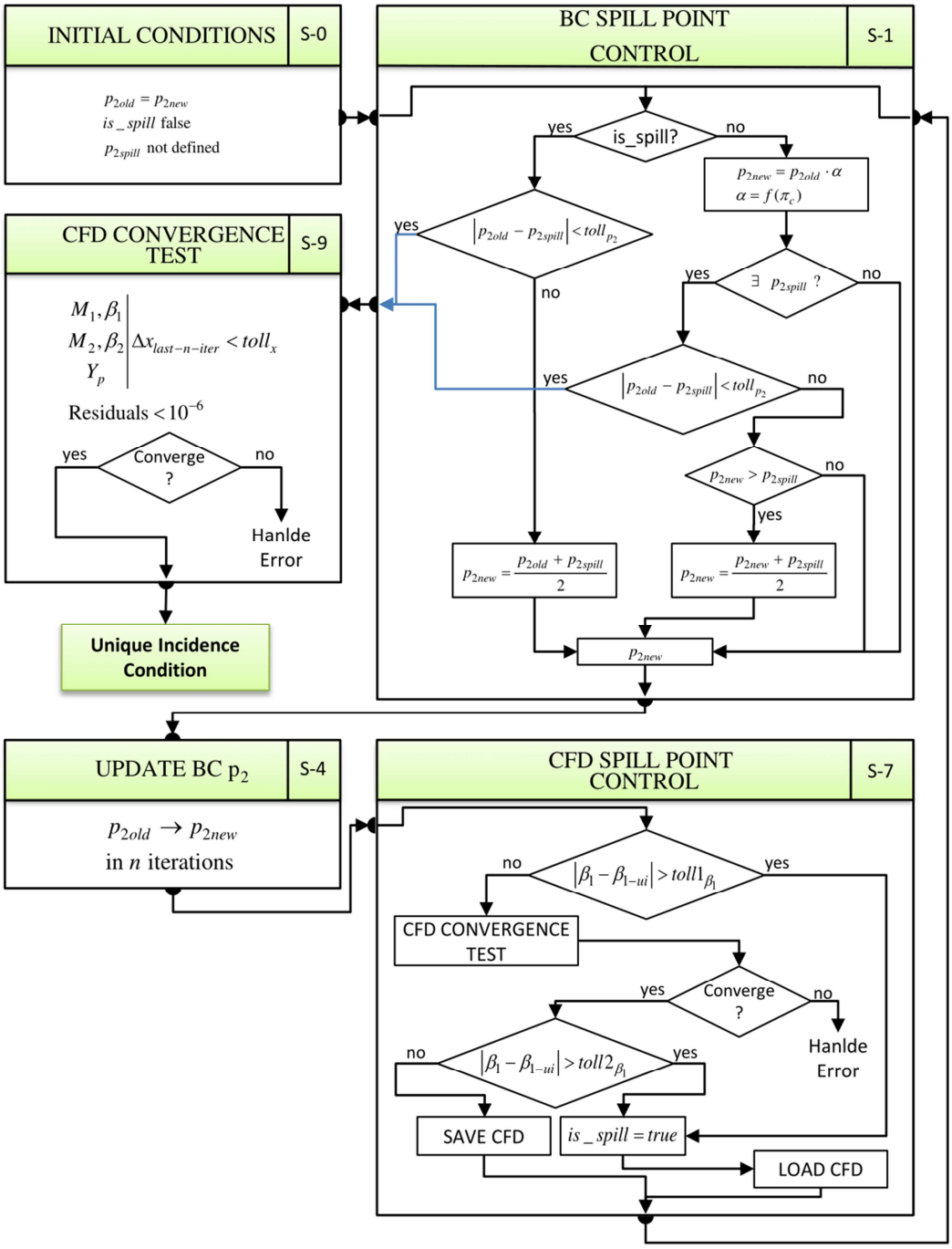


Figure 4.14: Flowchart of the Unique Incidence Status Machine (UISM).

4.3.3.1 Mach loop status machine

In the Mach loop status machine, there are six main blocks (Figure 4.13). The most important are hereafter described:

a. **Tuning** (M_∞, β_∞) **block**: (M_∞, β_∞) are tuned in order to meet the required $M_{1target}$. As shown in Figure 4.15, M_∞ is kept constant and equal to the target value during all the iterative process, whereas β_∞ is adjusted in such a way that, iteration after iteration, M_1 approaches $M_{1target}$. How β_∞ is updated is essential for this procedure. According to the analysis of supersonic cascades presented in §3.4, it seems a fairly strong hypothesis to approximate the Unique Incidence Curve to a linear model, at least within a neighborhood of $M_{1target}$.

   The curve linear model is defined by the last two points (M_1, β_1) evaluated via CFD, circle symbols in Figure 4.15. At iteration one, the initial point $(M_\infty, \beta_\infty)_{iter1}$ is arbitrarily selected, bearing in mind that β_1, thereby β_∞, is related with both the geometry of the leading edge profile and the inlet Mach number. When iteration one is completed, the issue of a second CFD point is avoided providing a fair value of the curve slope. As a result, the dashed line in Figure 4.15 represents the *a priori* unique incidence curve, being the slope value based on *a priori* knowledge on supersonic cascades rather than related to the specific cascade under examination. The intersection of unique incidence curve with the vertical line passing through the target Mach number identifies the new infinite condition for the next iteration. From the second iteration on, only CFD results are used to build the linear model. The iterative process is carried on until a tolerance $|M_1 - M_{\text{target}}| < toll_{M_1}$ on the inlet Mach number is met. In the experiments, the iterative process usually converges after only four or five iterations with a tolerance of 5.0E-05.

b. **CFD convergence test block**: when the boundary conditions are modified, there are a certain number of iterations needed in order to detect the perturbation on the flow domain. This is even more true when the attention is paid on the inlet quantities and the variation of boundary condition happens at the outlet section.

Depending on grid dimension, flow filed Mach number, CFD relaxation coefficients, and all those intrinsic simulation parameters, the number of iteration could vary from some dozens to some hundreds. Low residuals and flat pattern of the variable of interest are necessary conditions, but not sufficient to guarantee numeric convergence. By adopting smart locations for "numerical probes", it is possible to understand when the solution is completely propagated within the domain, thus to drastically reduce the total amount of iterations.

c. **Spill Point Test**: at this point, MLSM has identified a suitable value for (M_∞, β_∞) in order to meet the required Mach target. The main hypothesis is that the initial static back pressure is low enough guarantee a chocked flow condition inside the cascade. In such situation, the inlet field is completely independent from the static back pressure.

   The spill point test block is activated after the static back pressure is diminished, and it controls the variation of the inlet flow angle. If the test is positive, the back pressure is diminished and the Mach Loop is started again.

   The simplest solution would have been to adopt a very low value of static back pressure; however, experiments suggests that a such low value could induce numeric instabilities, at least for certain cascade geometries, besides the large amount of CFD iterations required to reach high pressure ratios typical of the unique incidence operations. The spill point test block introduces low complexity in the algorithm, while adding robustness the searching procedure and keeping to a minimum the CFD iterations.

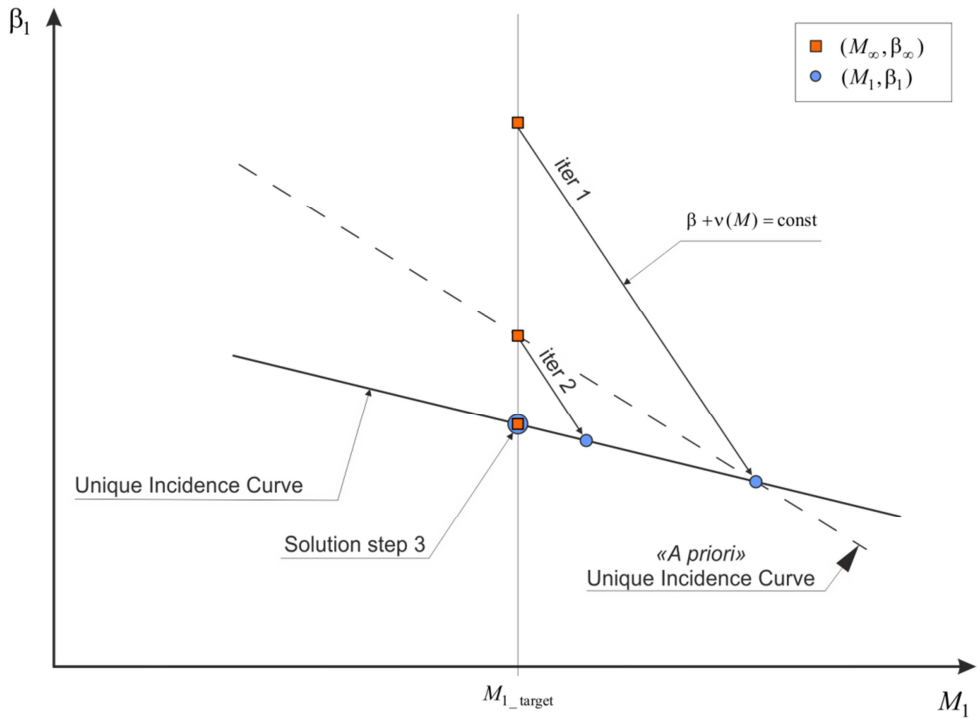


Figure 4.15: Procedure for tuning (M_∞ , β_∞) in order to achieve a prescribed M_1 .

4.3.3.2 Unique Incidence Status Machine

The Unique Incidence Status Machine (Figure 4.14) is composed by two main blocks:

a. **BC Spill Point Control**: this block is in charge to increase the static back pressure in order to force the cascade to work just before the spill point; this is the condition in which the shock wave detaches from the leading edge and the influence of the back pressure over the inlet region becomes remarkable: unique incidence relation does not hold any more.

   The low back pressure coming out from MLCL is increased by steps. At a certain point, the cascade overcomes the spill point for a specific value of the back pressure; such value is not acceptable because pushes the cascade beyond the spill point and so it is regarded as a superior limit for the following back pressure attempts. This superior limit and the last acceptable back pressure represent the starting points for a bisection method, which ends when a tolerance on the pressure ratio is achieved.

In the experiments, the tolerance was set to 0.05%, which represents the resolution on the UI pressure ratio. For such tolerance, the number of back pressure attempts are around ten.

b. **CFD Spill Point Control**: when a high value of static back pressure is imposed, the cascade could overcome the spill point and even reach its maximum pressure ratio, beyond which the only feasible solution is for a lower Mach number and a different inlet flow angle. In this situation, after convergence is obtained, even though the static back pressure is decreased to the previous value, an *hysteresis* phenomena appears: the cascade will not return to the previous operating point. In fact, the cascade is now operating in its subsonic regime or, in other words, the cascade is *unstarted*.

   The CFD control on the spill point monitors the variation of the inlet flow angle and roughly stops the CFD simulation when such variation overcomes a prescribed tolerance. In this way, the time spent on a worthless numeric solution is avoided, but still the information of the back pressure upper limit is retained and exploited in the bisection method.

   It is worth noting that experiments reveal a great variation in the margin of back pressure beyond the spill point among different geometries.

4.3.4 Validation

The UICL were employed in the CFD calculation of the baseline cascade for both turbulence models, i.e., SA and the kω-SST. In the previous validation phase, the numeric boundary conditions were tuned in order to achieve good agreement with the experimental shock wave pattern and the isentropic Mach distribution over the blade surface. In principle, there is any guarantee that such boundary conditions are specifically those for which the cascade operates in its unique incidence condition. Here, the inlet Mach number is kept constant equal to the one adopted in the validation phase (i.e. 1.456), while the static back pressure is free to vary in order to be tuned by the UICL.

The main results in terms of shock wave pattern and isentropic Mach number are reported in Figure 4.16 and Figure 4.17, respectively. The oblique shock wave slope is now similar between the two configurations (this is not the case in the

validation phase), this is confirmed also by the position of the isentropic Mach peak around 60% of the blade suction side. The main difference in the shock pattern can be appreciated in the strength and position of the "quasi normal" shock, which is stronger for the kω-SST than the SA model and it is located at 28% of chord against 18% of the SA case. As for the static pressure ratio needed to reach the unique incidence condition, the kω-SST works with a 2.186 pressure ratio compared to a 2.269 of the SA. The former under-predicts the experimental pressure ratio of -1%, while the latter overestimates the experimental by 2.7%. On the other hand, the loss coefficient is captured with very good accuracy by the kω-SST with a discrepancy of 0.5%, while SA differs by a 2.7%.

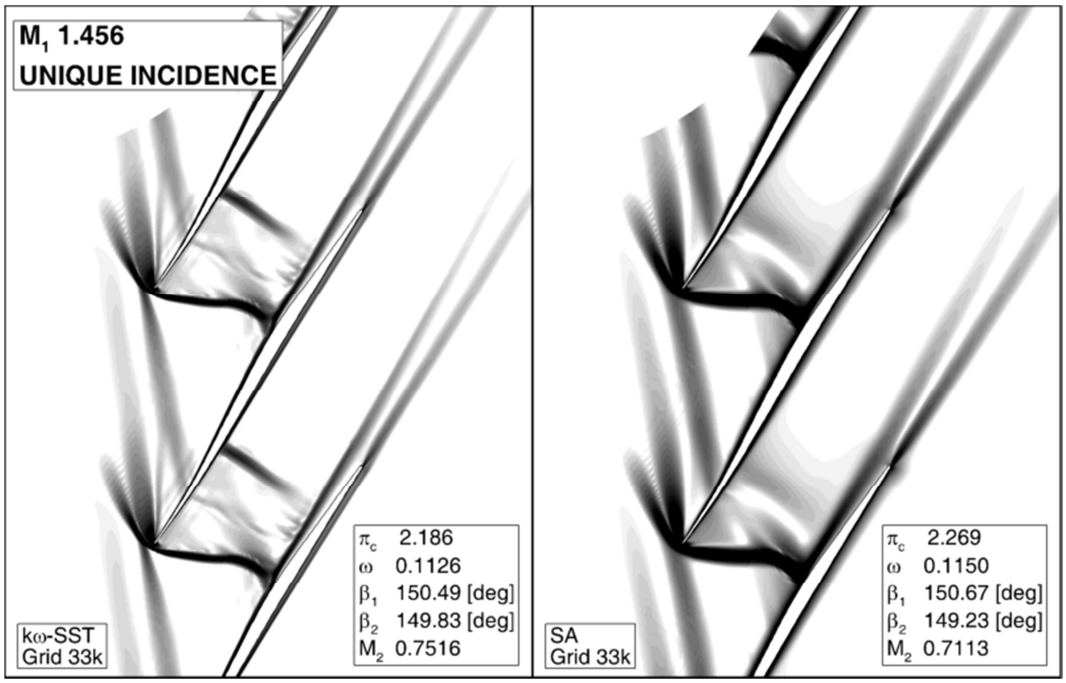


Figure 4.16: Magnitude of the density gradient. CFD calculation of the unique incidence condition by means of the Unique Incidence Control Loop (UICL). The kω-SST (left view) and the SA (right view) were adopted as turbulence models.

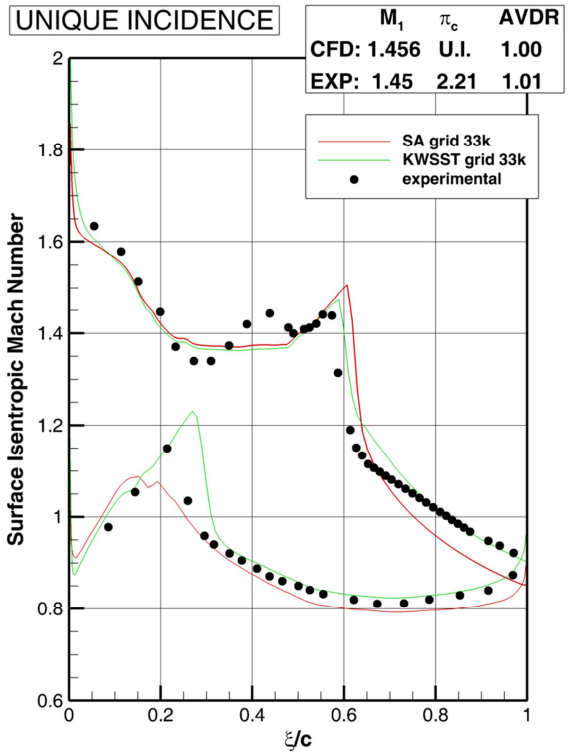


Figure 4.17: Isentropic Mach number distribution over the baseline cascade at unique incidence point: the kω-SST against SA turbulence model.

4.4 Parameterization

Shape deformations were accomplished by superimposing a displacement field to the baseline geometry points. The main advantage lies in the simplicity of such approach because it does not require the approximation of the baseline geometry with the curve related with parameterization, for instance a spline with control points. This does not mean that the baseline geometry is not approximated at all, but that the approximation could be made with any kind of curve, complex splines with hundreds of control points or higher order curves, regardless the purpose of the parameterization itself. The baseline reconstruction could be so accurate that it is like to include the exact shape of the baseline cascade within the design space. Thereby, the parameterization is focused on the displacement field, which defines the *dx* and *dy* displacement for each point of the baseline profile. In this work, B-spline curves were selected to parameterize the displacement field.

4.4.1 B-spline Curves: A Brief Review

A B-spline curve is a piecewise polynomial curve defined by a set of control points which the curve ordinarily does not interpolate. The reader is referred to the work of Mortenson [**106**] for a detailed explanation of B-spline curves and their numeric implementation. Hereafter, the main formulas are reported. The B-spline curve is defined as follows:

$$\boldsymbol{p}(u) = \sum_{i=0}^{n} P_i B_{i,p}(u) \qquad \text{Eq. 4.2}$$

parametric in the natural parameter *u*, the curve $\boldsymbol{p}(u)$ is defined by the control points P_i and the basis functions polynomials $B_{i,p}(u)$, which depends on the specific i-th control point, but its degree *p* is independent on the number of control points. The basis functions are defined iteratively:

$$B_{i,1}\begin{cases} = 1 & \text{if } t_i \le u \le t_{i+1} \\ = 0 & \text{otherwise} \end{cases} \qquad \text{Eq. 4.3}$$

$$B_{i,p}(u) = \frac{(u - t_i) B_{i,p-1}(u)}{t_{i+p-1} - t_i} + \frac{(t_{i+p} - u) B_{i+1,p-1}(u)}{t_{i+p} - t_{i+1}} \qquad \text{Eq. 4.4}$$

where t_i are the knot values that relate the parametric variable $u \in \langle 0, n+1-p+1 \rangle$ to the control points P_i. In the case of a uniform B-spline the knot values are equal distributed with an appropriate multiplicity at the extremes:

$$\begin{aligned} t_i &= 0 && \text{if} && i < p \\ t_i &= i - p + 1 && \text{if} && p \le i \le n \\ t_i &= n - p + 2 && \text{if} && i > n \end{aligned} \qquad \text{Eq. 4.5}$$

Compared with Bézier curves, B-spline curves have two main advantages: first, control points do not influence the curve globally, but affect only a neighbor region (*local control*); second, the curve degree does not depend on the number of control points, so adding control points to augment the control on deformation is always possible.

4.4.2 Geometry Reconstruction

A specific MATLAB® [**107**] application with graphic interface were developed with the goal to easily import, handle and parameterize a generic 2D airfoil. The rough data is manipulated in such a way to build a very detailed closed spline $\boldsymbol{p}_{base}(u)$, which loops from the trailing edge (u=0) forth to the leading edge (u=1) and back again to the trialing edge (u=2). A linear relation between the parameter u and the curvilinear length of the pressure and suction sides were imposed, and a uniform distribution of knots were used.

4.4.3 Displacement Field Definition

The displacement field were modeled by a cubic B-spline curve $\boldsymbol{p}_{displ}(u)$, which defines the *dx* and *dy* displacement for each point of the baseline. Similar to the baseline B-spline, $\boldsymbol{p}_{displ}(u)$ loops around the foil and the parameter u behaves in the same way, yielding a perfect match between the two B-spline representations. As a result, the following can mathematically represent a generic deformed geometry:

$$\boldsymbol{p}_{\mathrm{mod}}(u) = \boldsymbol{p}_{base}(u) + \boldsymbol{p}_{displ}(u) \qquad \text{Eq. 4.6}$$

The main difference between $\boldsymbol{p}_{base}(u)$ and $\boldsymbol{p}_{displ}(u)$ is the number of control points used. The former has hundreds of control point to well approximate the baseline geometry, while the latter has as many control point as the parameterization requires.

It was arbitrarily decided to prevent deformation of the leading and trailing edge profiles during optimization. In order to obey to such constraint, inactive (or fixed) control points were added in the leading and trailing edge regions, as illustrated Figure 4.18. In this way, the local influence of the active control points on the not deformable region were avoided.

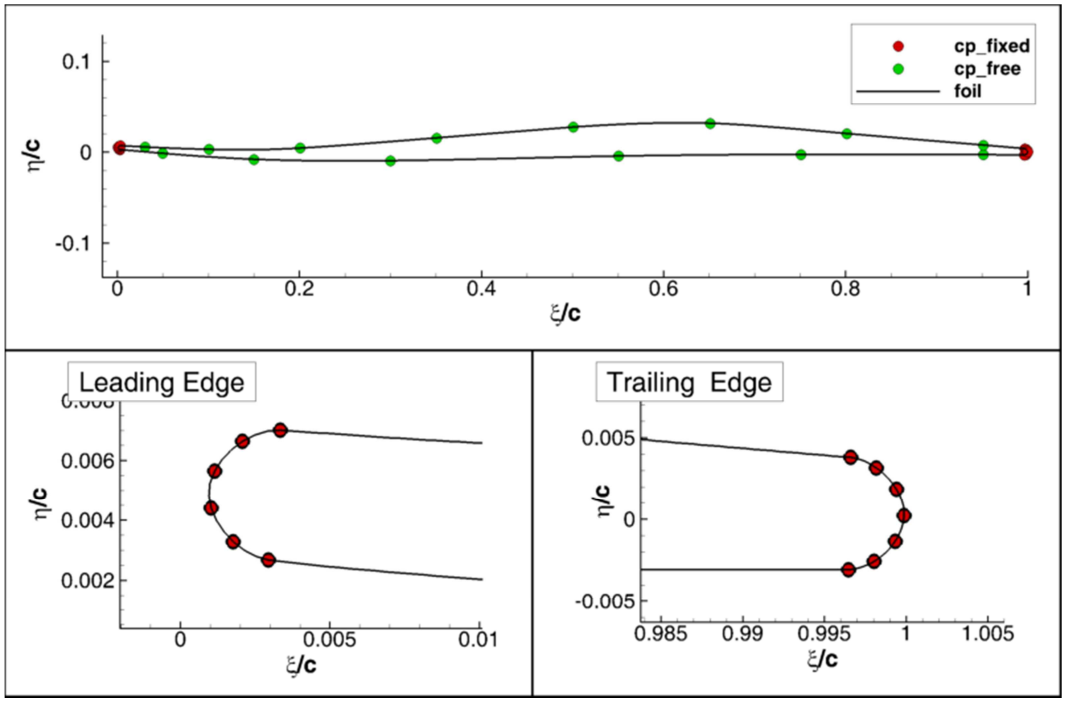


Figure 4.18: B-spline control points: active control points in green, while fixed ones in red.

4.4.4 Geometric constraints

Feasible geometries are not guaranteed by the employment of the displacement field. For instance, pressure and suction sides could even intersect depending on the control points range, at least in the fore portion of the blade. Curbing the control points range in order to avoid intersection extremely tapers the search space; these geometries were just rejected.

On the other hand, tapered fore portion geometries are exceptional to accommodate the incoming flow, but less appreciated by the manufacturing. The minimum thickness and the covered area in the first 35% of chord were monitored during optimization.

4.5 Formulation Of The Optimization Problem

The aim of the multi-objective optimization problem was to minimize the two-objective function:

$$\boldsymbol{F}(\boldsymbol{P}) = (f_1, f_2) = (\omega, -\pi_c) \qquad \text{Eq. 4.7}$$

where $\boldsymbol{P}$ is the vector of the optimization parameters or decision variables of the design (i.e., the B-spline control points of the displacement field), ω is the cascade loss defined in Eq. 4.1 and π_c the static pressure ratio.

The constraints complete the problem formulation. A penalty function approach is adopted to handle geometric constraints. The penalty term $\boldsymbol{Q}$ depends linearly on the level of violation of the constraints:

$$\begin{aligned} &\boldsymbol{Q}=(\mathrm{Q}_{1,}Q_2)=(r_1,r_2)\cdot\sum_{i=1}^{2}a_iG_i \\ &G_1=\max\{0,t_{\%}-t_{\%\mathrm{limit}}\} \qquad t_{\%\mathrm{limit}}=-5\% \\ &G_2=\max\{0,A_{35\%}-A_{35\%\mathrm{limit}}\} \qquad A_{35\%\mathrm{limit}}=-10\% \end{aligned} \qquad \text{Eq. 4.8}$$

where $t_{\%\mathrm{limit}}$ is the lower bound of the minim thickness percentage variation, $A_{35\%\mathrm{limit}}$ is the lower bound of the percentage variation of the covered area from leading edge till 35% of chord, a_i is a linear weight and (r_1,r_2) are the transformation coefficients in order to properly scale the violation to the objective functions.

The reader is referred to Appendix A for a brief review of the optimization framework adopted.

4.6 Discussion Of Results

Two multi-objective optimizations were carried out by means of the Kriging-assisted genetic algorithm GeDEA-II-K (see §Chapter 2). In the first analysis, namely OPT1, the 14 design variables having control of the forepart of the foil were activated, while in the second one, namely OPT2, the entire geometry was allowed to be modified by the 21 degrees of freedom embedded in the parameterization. Although the CFD model was composed by only 33k elements, for each computation 1.5h was required in order to attain the UI condition. For both the optimizations, a population composed by 20 individuals was evolved for 12 generations.

Figure 4.19 shows the Pareto-optimal front of both optimizations. The OPT1 accounts for more than 95% of the losses reduction achieved in OPT2, corroborating the fact that the forepart of the airfoil geometry is of utmost importance when supersonic flows need to be accommodated.

As far as the OPT1 is concerned, two curves are reported in Figure 4.19: the green front takes into account designs that strongly respect the geometric

constraints, while the red front is constructed including designs that did not completely fulfill the constraints, but with a weak level of violation. Since the constraints violation is representative of airfoils featuring a tapered shape in the first 35% of chord, it is clear why the red front could extend its left branch towards regions of lower losses compared to the green front. Even with the second analysis such levels of losses (i.e., point B1) could not be reached.

On the other hand, the blue front in Figure 4.19 refers to OPT2 and, compared to OPT1, it enlarges the covered area toward high static pressure ratios, whereas little improvement in loss reduction is detected.

The results are presented by means of total pressure losses contours, Mach number contours, density gradient magnitude contours (which resembles the Schlieren pictures) and isentropic Mach number distribution. The few optimal points reported in Figure 4.19 are taken into account for comparison.

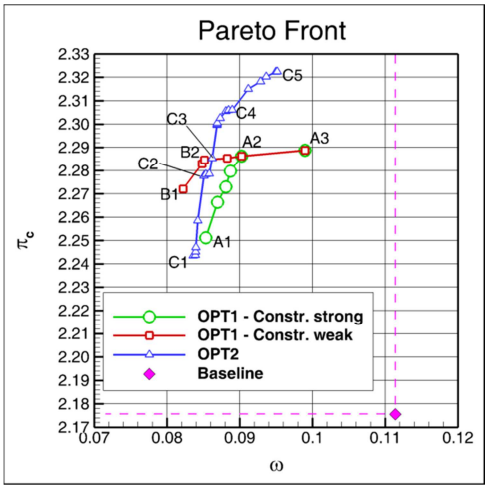


Figure 4.19: Pareto-optimal front.

4.6.1 General Considerations

The disappearance of the strong shock wave boundary layer interaction explains the greatest improvement of the loss coefficient among all the Pareto designs. The isentropic Mach number distribution (Figure 4.20) reveals that somehow the optimal designs are able to lower the incoming shock wave Mach number, which is reduced from 1.5 for the baseline to 1.32-1.38, and slightly shift toward aft (i.e., from 58% of chord of the baseline to 62%) the impinging point of the passage oblique shock over the suction side. A lower pre-shock Mach number

reduces the passage shock strength and the shock wave boundary layer interaction. This phenomenon will be analyzed ahead. Although with different degrees of reduction, the flow reacceleration on the pressure side of the baseline is damped, thus the second shock wave is weakened; in addition, the position of the shock moves towards the leading edge at about 20% of chord.

The optimal designs featuring a higher pressure ratios belong to the OPT2 analysis, in which the entire airfoil was modified including the rear part. The flow is first decelerated to subsonic velocity in the forepart of the passage and then it is deflected furthermore in the rear part, thanks to the increased exit solid angle (see Figure 4.23). As a result, between the 60% and 95% (see Figure 4.20) the isentropic Mach number makes evidence of the higher loading in agreement with the flow pitchwise distribution of the outlet flow angle (see Figure 4.21). Moreover, OPT2 optimal designs have the maximum thickness location shifted aft around 68% of chord, while the maximum value is slightly higher than the baseline (see Figure 4.24).

As the pitchwise distribution of the loss coefficient is concerned (Figure 4.22), it can be observed very well the absence of the “bulge-pattern” caused by the boundary layer detachment, which is the result of the strong interaction between the passage oblique shock and the boundary layer on the suction side (lambda shock). Generally, the main differences between optimal designs lay on the region outside the “bell”. However, it is singular how the B1, B2 and A1 can achieve slightly lower losses under the “bell”. The explanation of such behavior is not simply related with the tapered shape of the airfoil forepart (i.e., this could be true only for B1 and B2 which do not respect the geometric constraint), but must be searched on the interaction between the bow shock wave in front of the leading edge and the inlet flow domain. It will soon be clear.

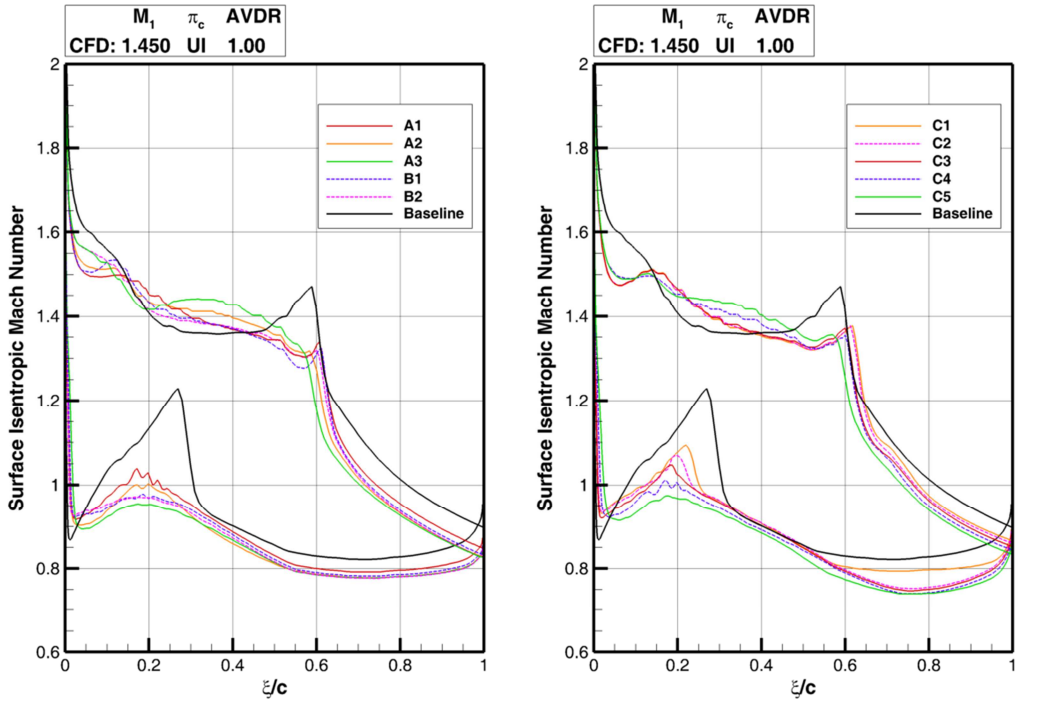


Figure 4.20: Isentropic Mach distribution over Pareto-optimal front designs.

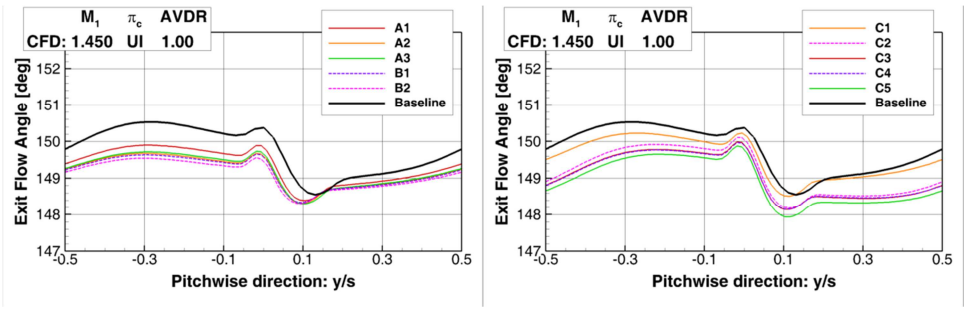


Figure 4.21: Pitchwise distribution of the outlet flow angle for the Pareto-optimal front designs.

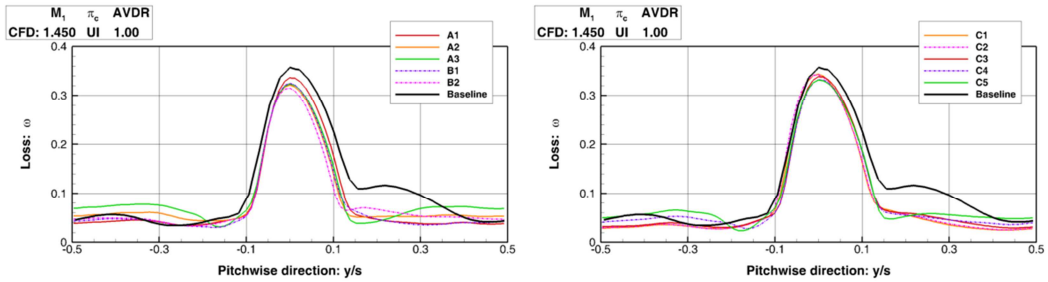


Figure 4.22: Pitchwise distribution of the loss coefficient for the Pareto-optimal front designs.

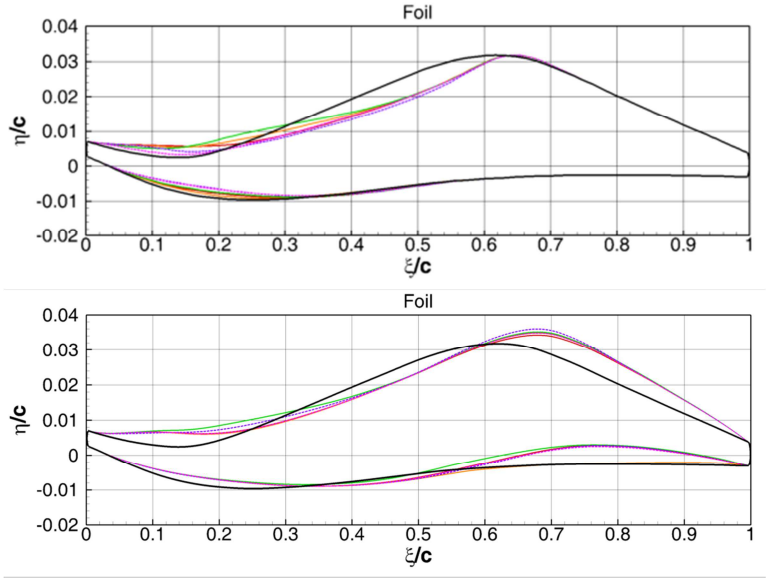


Figure 4.23: Suction and pressure side of the Pareto-optimal front designs.

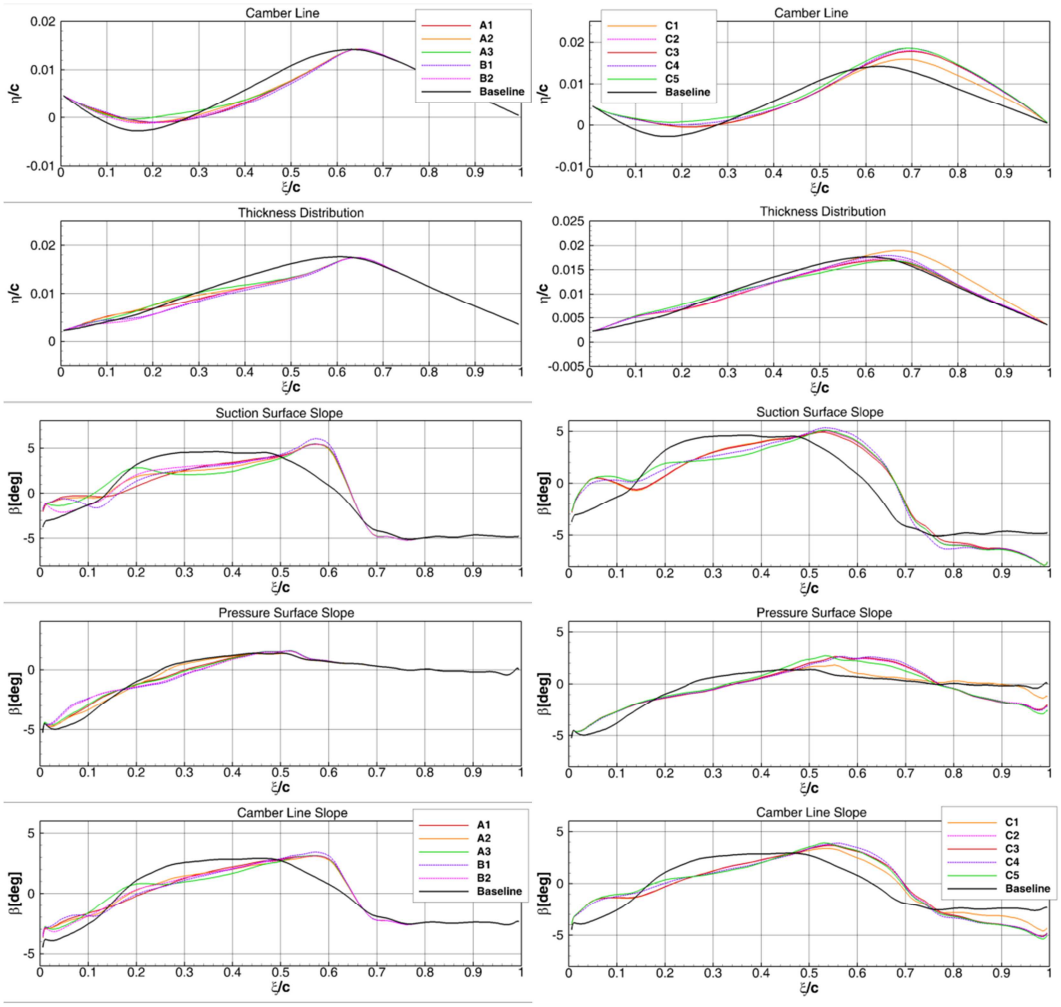


Figure 4.24: Geometric features of the Pareto-optimal front designs.

4.6.2 Pre-shock Mach reduction

Two main mechanisms are responsible for the pre-shock Mach number reduction: the reduced flow acceleration on the suction surface near the leading edge and the pre-compression mechanism.

4.6.2.1 From 0% to 20% of chord

All the optimal designs have a lower negative inlet camber (see Figure 4.24) that results in a decreased incidence flow angle. Moreover, compared to the baseline, although the camber line continues to have the "bowl shape" (i.e., the camber has negative tangent from 0% to 15% of chord and then increases toward positive values), its depth is less pronounced and thus the acceleration of the incoming flow is restrained. Because of both the reduced incidence and the softer "bowl shape", the flow experiences a limited acceleration, which can be appreciated in both the isentropic Mach number distribution (see Figure 4.20) and in the Mach number contours (Figure 4.26 and Figure 4.29).

Following the flow development on the suction surface, after the expansion, the flow enters the region were the pre-compression mechanism is established. The variation of the suction surface slope imposes to the flow a continuous deflection, which is realized by compression Mach lines of weak strength departing the suction surface. The baseline geometry (

Figure 4.25) is designed in such a way that these Mach lines coalescence towards a narrow region (i.e., vertex of the compression fan) that is the result of the intersection of the Mach lines with the leading edge bow shock of the previous blade. Approaching the vertex, the contribution of the whole Mach lines becomes significant and gives rise to a finite, but still not too strong, compression wave. However, the intersection of compression Mach lines with shock waves is not isentropic and, as a consequence, the compression fan induces a strong deficit in total pressure that remains confined in a narrow strip called *linguina*. The *linguina* departs from the intersection point and goes downstream (see close-up of the baseline geometry leading edge Figure 4.27). In the baseline, the *linguina* total pressure loss is around 8%, which is remarkable compared to a 4% of mean total pressure loss occurred through the passage oblique shock.

The modifications of the camber "bowl shape" change the direction of the Mach lines departing the suction surface, lines that no more coalescence towards a narrow region. The "vertex" of the compression fan becomes a wide spread region that covers a wide part of the bow shock or points towards directions beyond the leading edge of the previous blade, depending on the optimal design considered. The simulated Schileren pictures reported in

Figure 4.25 and Figure 4.28 show in dark grey strong variations of density that correspond to compression or expansion regions; the information of the Mach contour resolve the ambiguity.

In the baseline configuration, it is not clear to what extent the *linguina* induces the formation of the lambda shock system, but is clear from the close-up of Figure 4.27 that the existence of a low total pressure strip exactly at the top of the lambda shock system has some drawbacks. To answer the question further investigations are required.

As far as the pressure side is concerned, the negative slope does not vary near the leading edge (see Figure 4.24), contrary to the suction side where the slope had a jump of +2 [deg], thus leading to a wedge angle larger than the baseline. However, this effect is mitigated by a faster slope increase that produces a lower acceleration of the flow and a very weak second shock wave. The slope has small variation among optimal design, at least until the 50% of chord, exception made for B1 and B2 designs that do not respect the geometric constraints and feature a tapered shape.

4.6.2.2 From 20% of chord to passage shock wave

Proceeding along the suction surface, the baseline has a flat suction surface slope (see Figure 4.24) between 28% and 48% of chord (position of the passage oblique shock wave), thereby no more Mach Lines are generated and the supersonic Mach number remains more or less constant around 1.38, before a remarked acceleration up to 1.48 just before the shock. On the other hand, all the optimal designs features a peculiar profile of the suction surface slope, which continues to increase and, therefore, additional Mach lines are generated and thus the flow is furthermore decelerated. The Mach lines are quasi-parallel one another and shatter on the passage oblique shock wave, avoiding any influence outside of the cascade passage. The flow reaches the shock wave front at the 57-62% of

chord with a Mach number of 1.32-1.38, depending on the optimal design contemplated.

4.6.1 The Tradeoff Imposed By Unique Incidence

The Unique Incidence (UI) condition introduces a further degree of freedom in the optimization process related with the passage shock wave obliquity. When the static back pressure is increased, the passage oblique shock wave modifies its obliqueness, tending to be more “normal”. The shock movement can be assimilated with a rigid rotation around a pivot point (i.e., the hinge of the shock) located nearby the leading edge. The rotation can proceed until the UI is achieved, i.e., until the position of the shock wave does not disturb the incoming inlet flow. Precisely, the passage shock influences the shape and position of the bow shock lower branch, which, beyond a certain rotation, transmits this variation at the upper branch of the bow and, from there, the disturbance propagates in front of the entire cascade. It can be observed that optimal designs with high pressure ratio (HPR) have passage shocks much more “normal” than those performing a low loss coefficient (LLC). In fact, as described by the losses of Figure 4.27 and Figure 4.30, the wakes released by the passage shocks of HPR designs have a higher total pressure deficit, even though the pre-shock Mach distribution is comparable with that of the LLC. It seems that there is a limit in the pre-compression mechanism, probably due to the fixed cascade solidity, beyond which the pre-shock Mach number cannot be decreased furthermore. In this condition, the designs that can support a greater rotation of the shock wave are the HPR, whereas those that are more susceptible to the shock rotation flow into the LLC designs set. From this point of view, we can say that the UI condition introduces a tradeoff between HPR and LLC.

The mechanism behind the maximum rotation of the passage shock wave is a complex phenomenon related with both the bow shock in front of the leading edge and the location of pre-compression fan vertex. From the Mach number contours (Figure 4.29) and the numerical Schlieren visualizations (Figure 4.28) of HPR designs as A3 and C5, it can be inferred that there is a noticeable shape discontinuity in the wake front where the bow shock meets the passage shock: this point turns out to be the hinge of the passage shock. In addition, a very short expansion/compression branch departing the pressure surface at 0.5% of chord attaches to the hinge point. This interaction gives rise to a *linguina* characterized by a very low total pressure loss. It seems that the *linguina* establishes a

decoupling effect between the passage and the bow shock, letting the former to rotate around its hinge for greater degree without affecting the shock bow.

The tough questions is in which manner the airfoil geometry can control the existence of the expansion/compression branch, and thereby the UI condition. There is no strong evidence that suggests that this occurrence is somehow related with the pressure surface profile only. We think that the explanation should be searched in the pre-compression mechanism. It seems that the HPR designs features a higher Mach number ahead of the bow shock lower branch and a weaker pre-compression fan near its vertex.

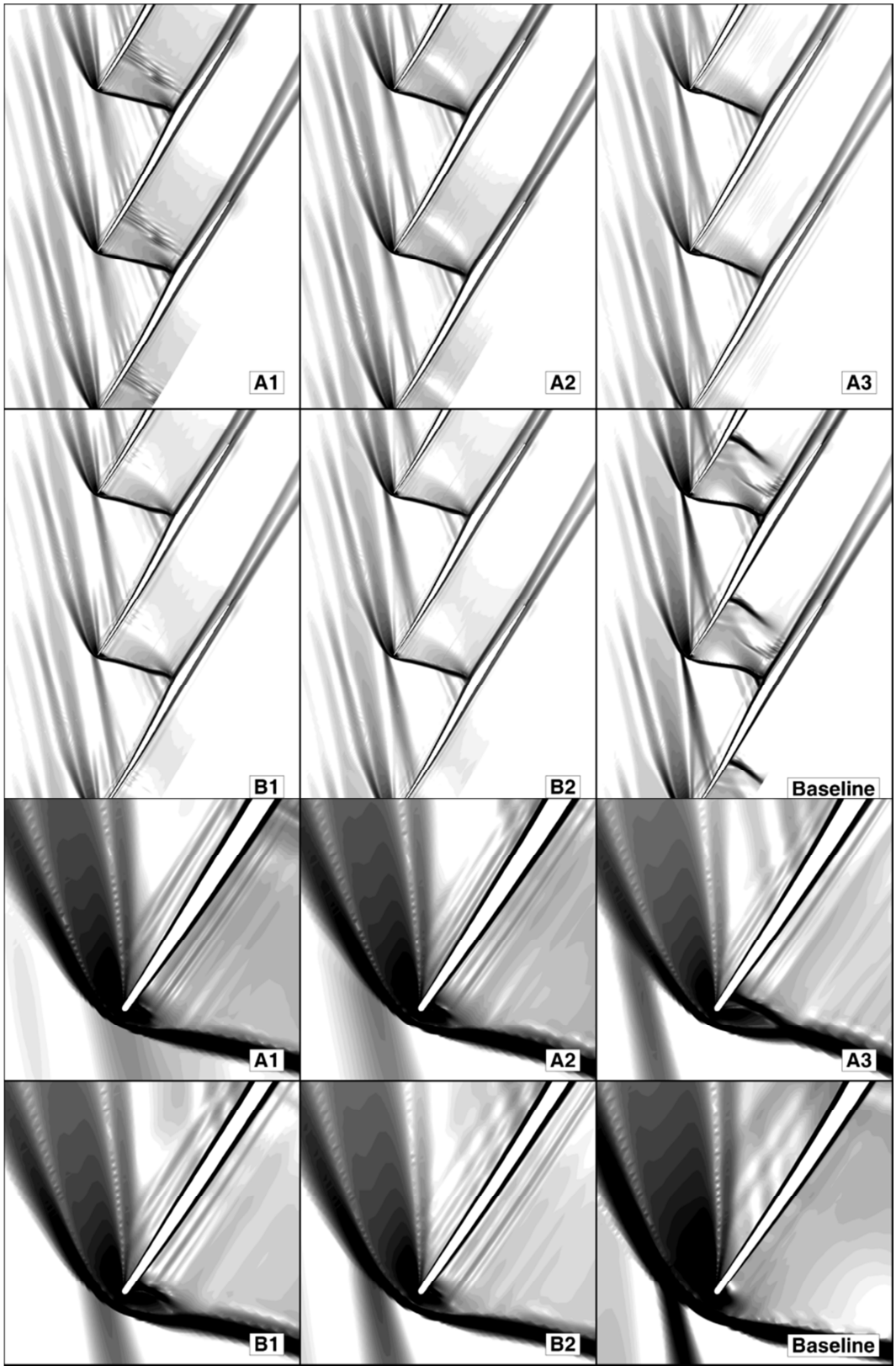


Figure 4.25: Density gradient magnitude of the OPT1 optimal designs. Passage global view (top) and a close-up on the leading edge (bottom).

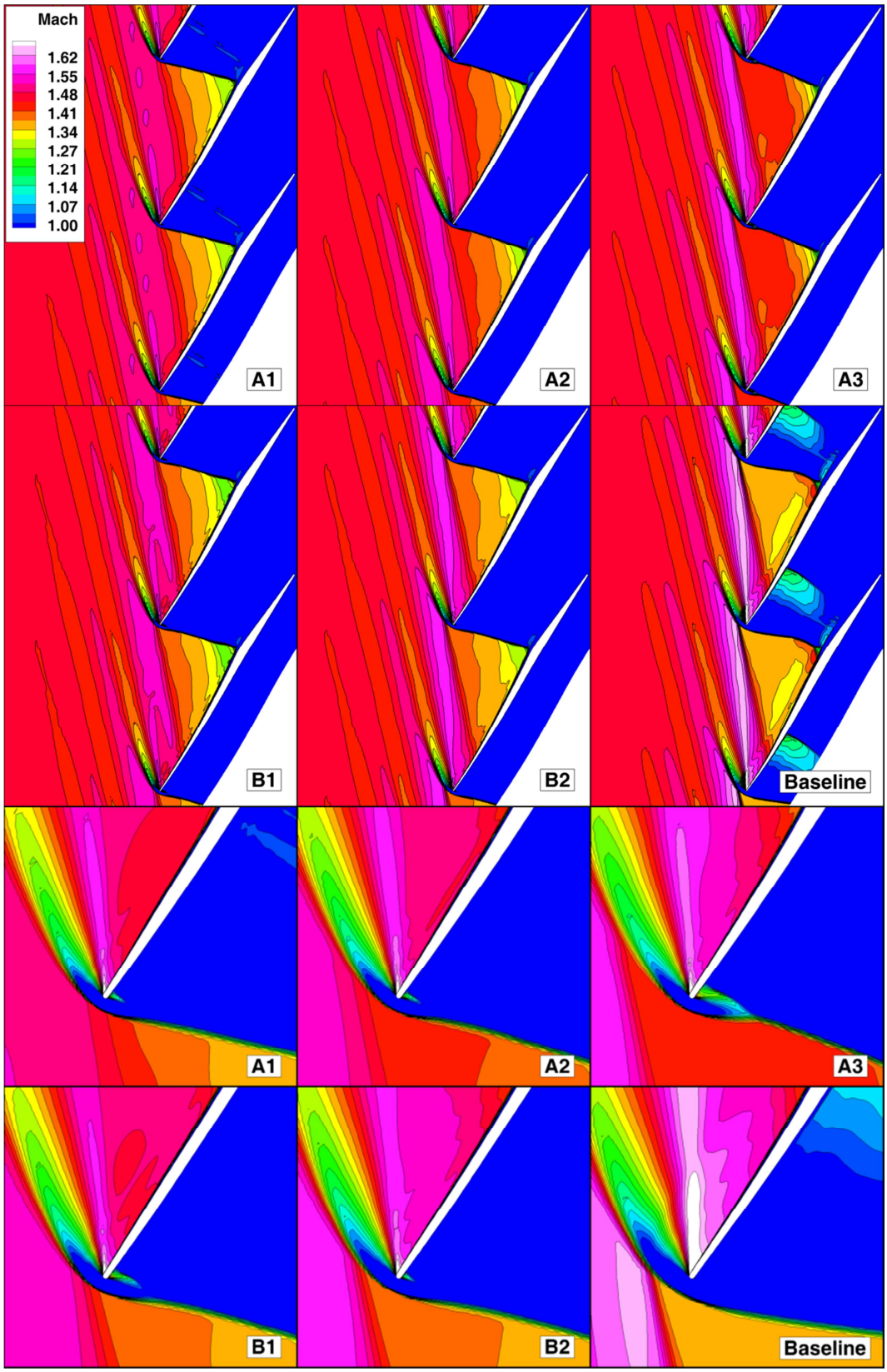


Figure 4.26: Mach number contours of the OPT1 optimal designs. Passage global view (top) and a close-up on the leading edge (bottom).

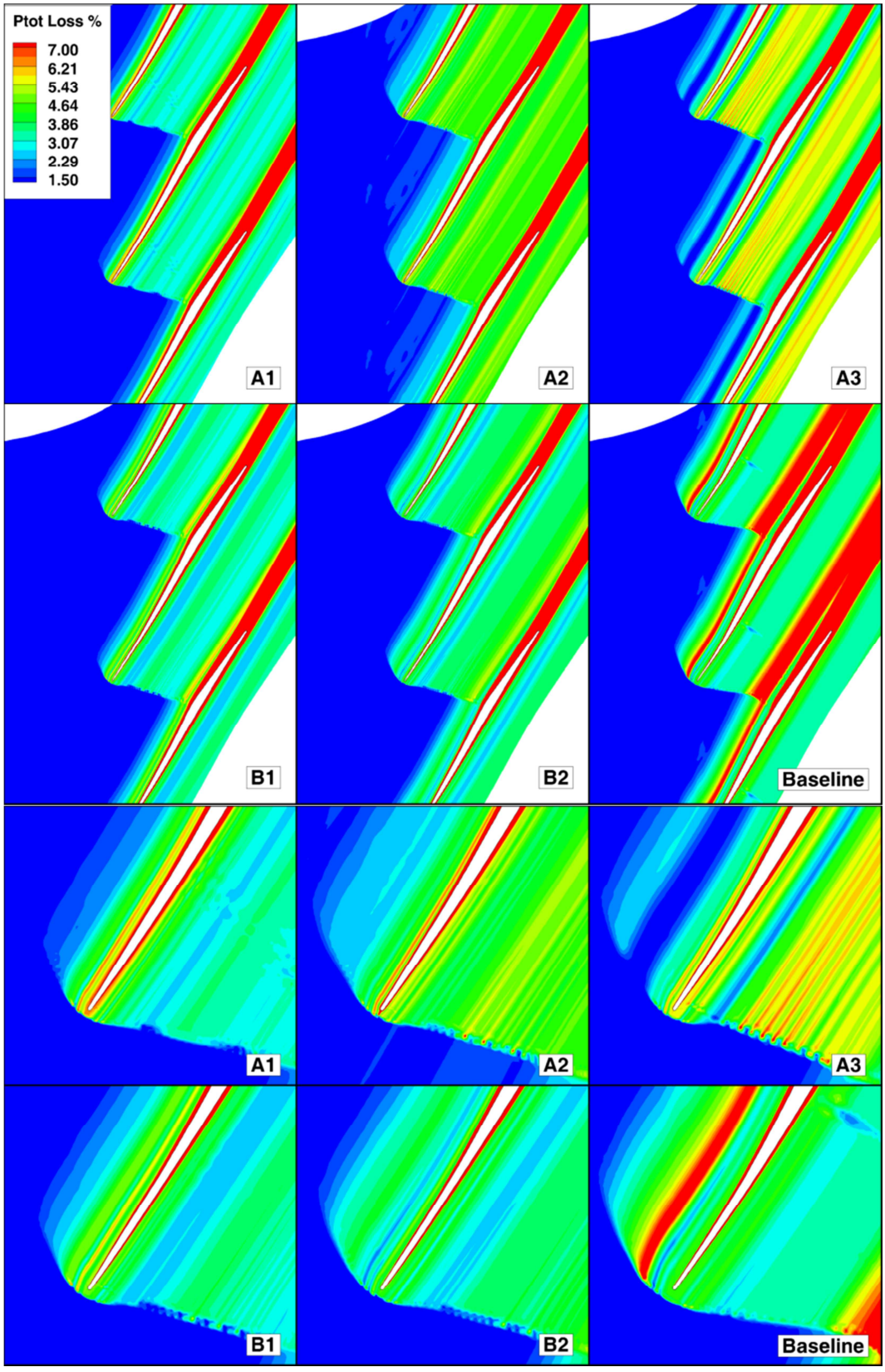


Figure 4.27: Total pressure loss contours of the OPT1 optimal designs. Passage global view (top) and a close-up on the leading edge (bottom).

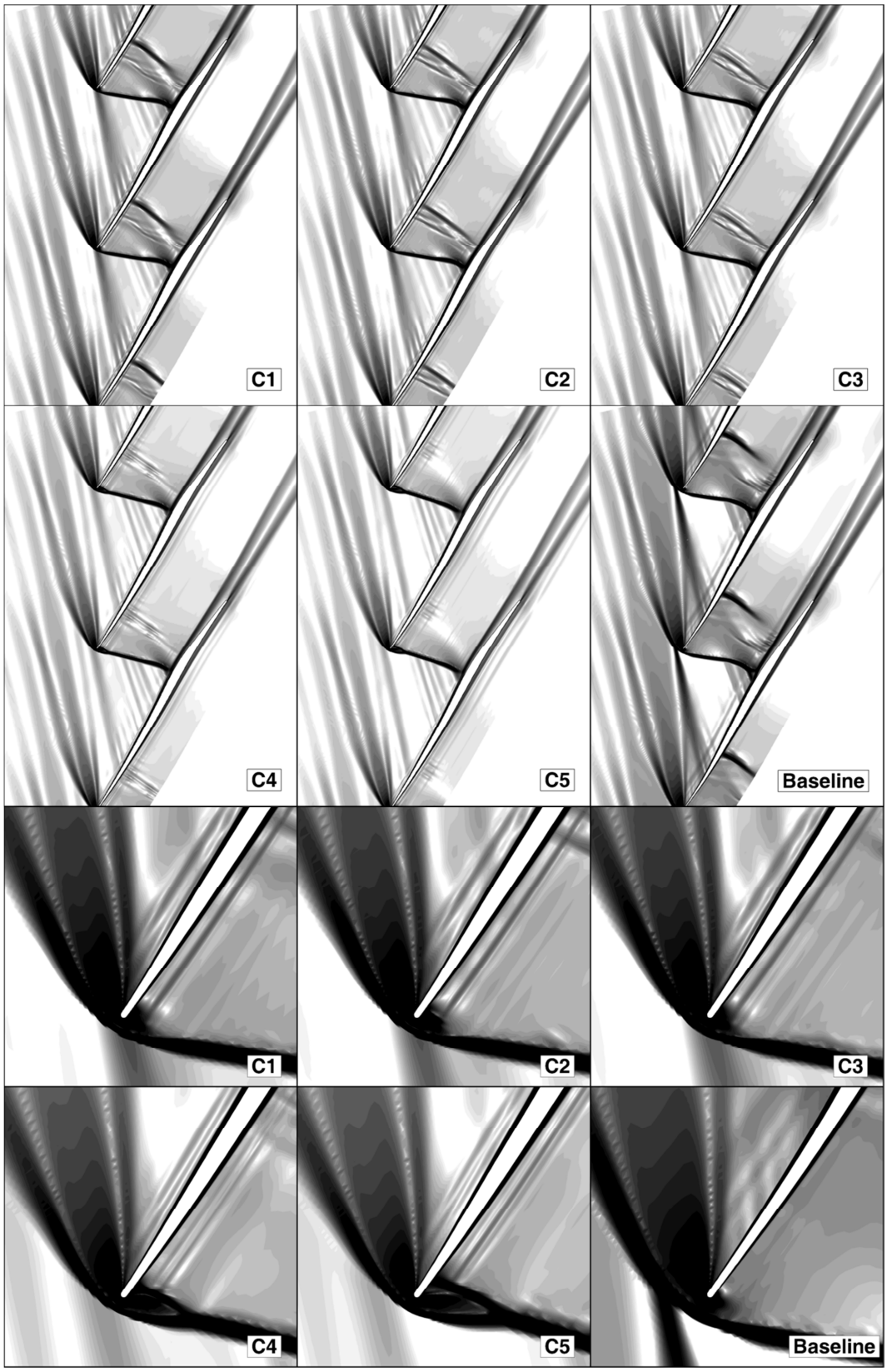


Figure 4.28: Density gradient magnitude of the OPT2 optimal designs. Passage global view (top) and a close-up on the leading edge (bottom).

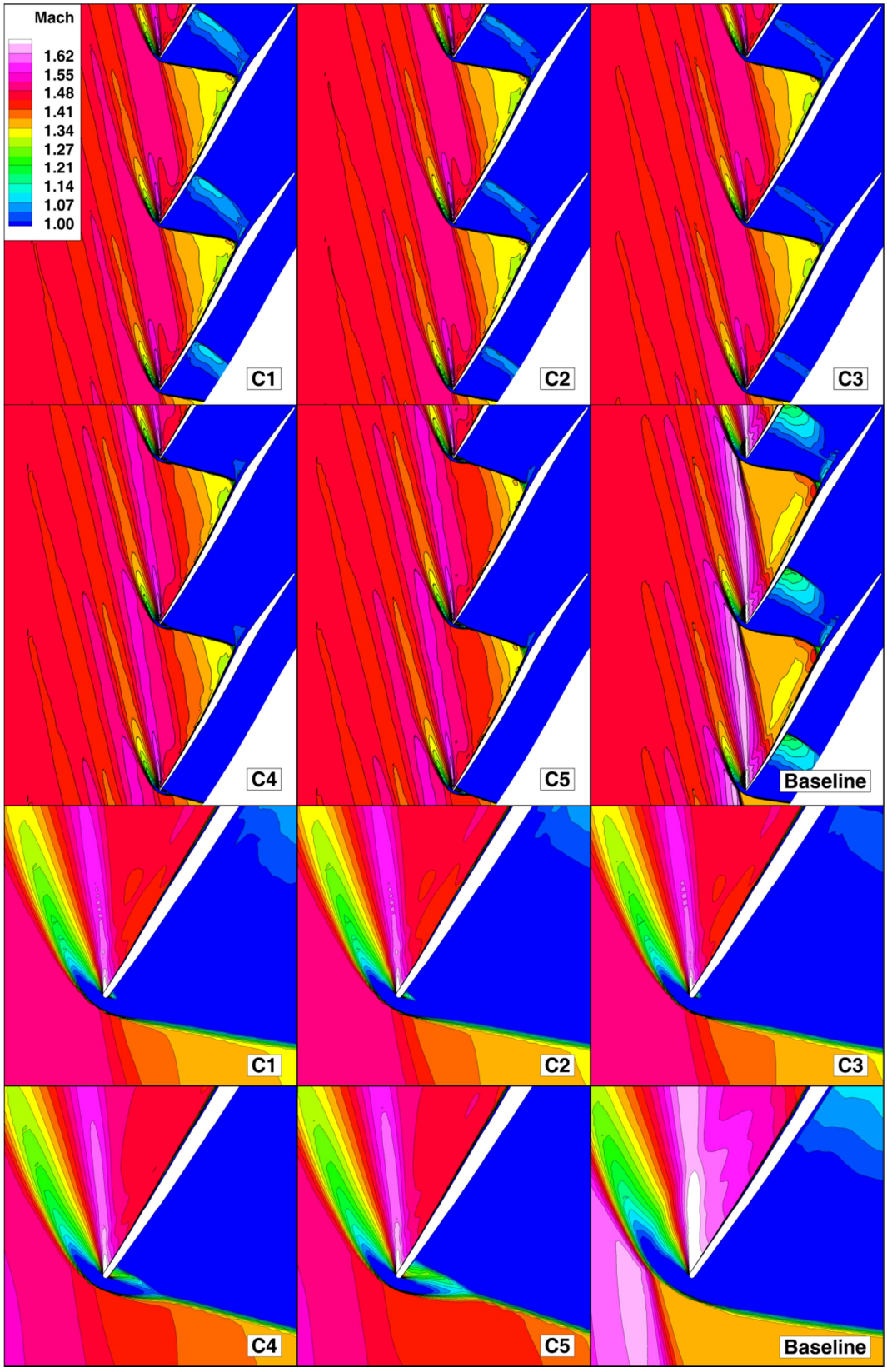


Figure 4.29: Mach number contours of the OPT2 optimal designs. Passage global view (top) and a close-up on the leading edge (bottom).

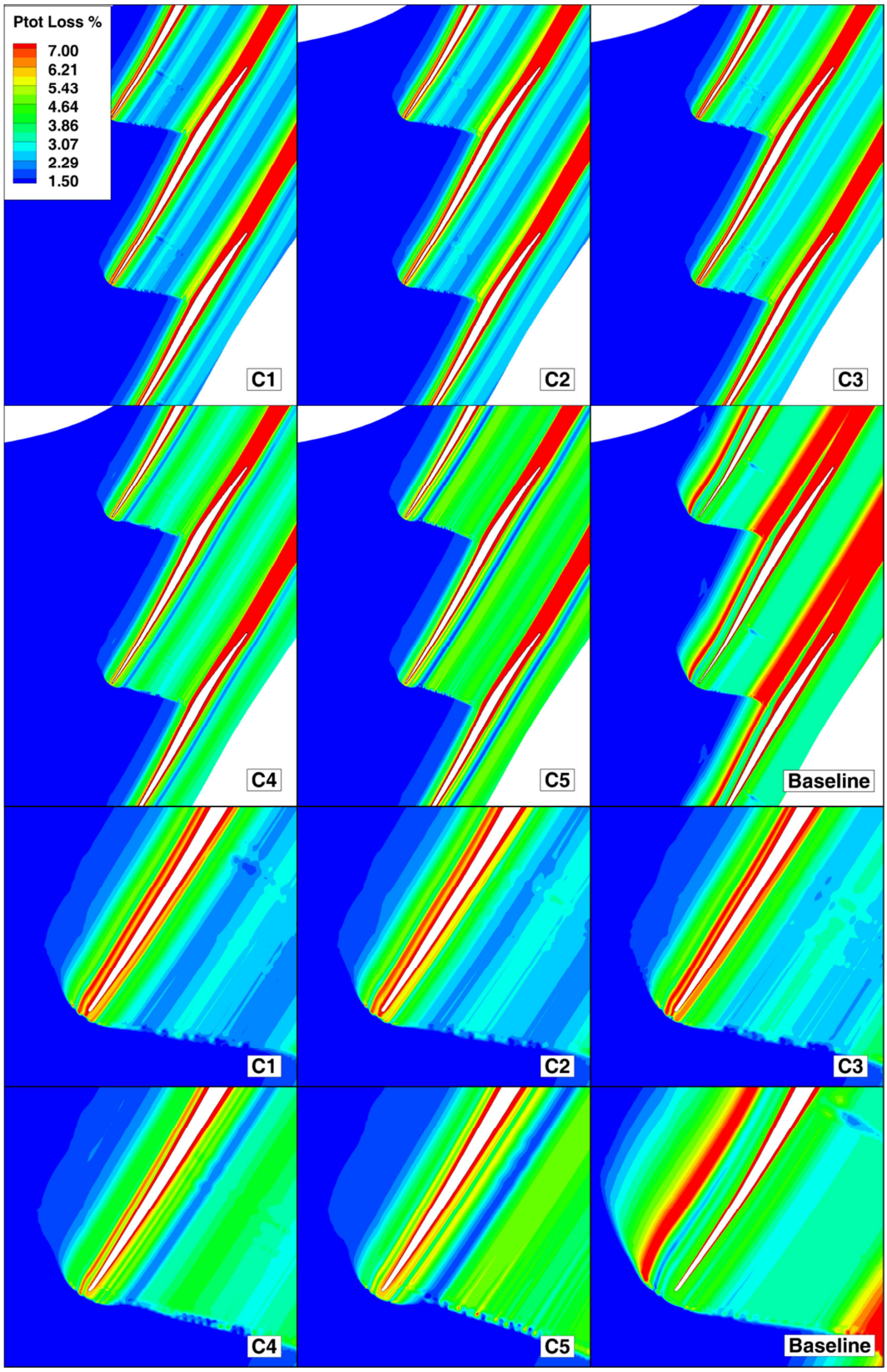


Figure 4.30: Total pressure loss contours of the OPT2 optimal designs. Passage global view (top) and a close-up on the leading edge (bottom).

4.7 Conclusions

The DLR-PAV-1.5 supersonic compressor cascade was optimized following a multi-objective approach where loss coefficient and static pressure ratio were the two objectives to be optimized. The optimization was carried out by means of the novel Kriging-assisted genetic algorithm GeDEA-II-K (§Chapter 2) coupled with CFD solver. The interesting peculiarity is that the cascade was optimized at its unique incidence condition. In particular, two optimizations were performed: the first one took into account only the reshaping of the forepart of the cascade airfoil geometry, while the second one comprises the entire geometry. Among the Pareto designs, the loss coefficient was reduced by 25% and the static pressure ratio by 6.5%. It was demonstrated that the forepart geometry all alone was responsible for the 95% of the loss reduction, thanks to the remarkable decrease of the pre-shock Mach number by means of an *extended* pre-compression mechanism that involves the entire suction surface until the passage shock.

The unique incidence condition introduced a tradeoff between the two objectives. Higher compressions were achieved diminishing the obliquity of the passage shock wave, which tended to a more “normal” layout, however with an increase on the shock wave losses. Furthermore, it was identified a peculiar mechanism that established on what extend the passage shock can rotate before the unique incidence is met.

Chapter 5

Shape Optimization of a Transonic Compressor

In this chapter, the shape optimization of a transonic rotor is carried out. The NASA Rotor 67 (Figure 5.1) is taken as reference and a 3D parameterization involving lean, sweep and airfoil reshaping is applied. The novel Kriging-based genetic algorithm GeDEA-II-K is employed for a two-objective optimization, in which total pressure ratio and polytropic efficiency have to be optimized at a given operating condition.

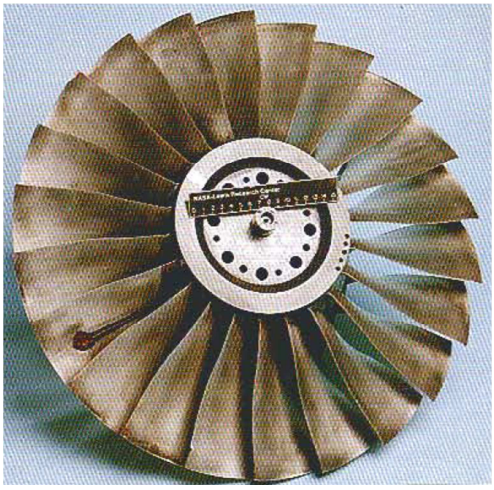


Figure 5.1: Nasa Rotor 67 transonic fan.

5.1 NASA Rotor 67

Designed by NASA Lewis Research Center, the Rotor 67 (Figure 5.1) is a low-aspect-ratio transonic rotor and is the first rotor of a two-stage fan. According to NASA report [**108**], the rotor has 22 blades and an aspect ratio of 1.56 (based on average span/root axial chord); the solidity varies from 3.11 at the hub to 1.29

at the tip, and the inlet and exit hub/tip radius ratios are 0.375 and 0.478, respectively. The running tip clearance is 1.016 mm. The design pressure ratio is 1.63 at a mass flow of 33.25 kg/sec and the design wheel speed is 16043 rpm, which yields a tip speed of 429 m/sec.

5.1.1 Literature Review

In literature there are several works focused on the shape optimization of Rotor 67. The following review has the purpose to highlight which part of the blade was parameterized and how the optimization strategies, in terms of objectives and constraints, affect the outcomes of the optimization process; a snapshot of blade parameterization and optimization strategies are reported in Table 5.1 and Table 5.2.

Authors	**Parameterization**	
	Type	Design variables
Oyama *et al.* [**109**]	Airfoils shape: 4 spanwise sections; camber line + thickness distribution	53
Pierret [**110**]	Airfoils shape: 4 spanwise sections	35
Lian and Liou [**111**]	Airfoils shape: 4 spanwise sections (perturbation approach)	32
Lian and Kim [**112**]	Airfoils shape: 4 spanwise sections (perturbation approach)	32
Okui *et al.* [**113**]	Airfoils shape: 5 spanwise sect.; camber + spanwise chord distribution Sweep	14
Luo *et al.* [**114**]	Airfoils shape: 32 Hicks-Henne shape functions (perturbation approach)	238
Hu *et al.* [**115**]	Blade loading (inverse design strategy) Only airfoil shape can change	-

Table 5.1: Previous studies on NASA Rotor 67: parameterization.

	Authors	Optimization Strategy		
		Method	Objectives	CFD runs
No Surrogates	Oyama *et al.* [**109**]	GA+CFD	-Min. Entropy generation @peak efficiency -Constr.: mass flow + total pressure ratio	12800
Off-line	Lian and Liou [**111**]	RSM+CFD offline	-Min. Entropy generation + Max. total pressure ratio @peak eff. - Constr.: mass flow	1023
In-line	Lian and Kim [**112**]	RSM+CFD inline	- Max. total pressure ratio + Min. Blade Weight @peak eff. - Constr.: mass flow + probability of failure	1678
	Okui *et al.* [**113**]	DE+ANN +CFD	-Max. Isentropic eff.+ throttle margin @peak eff. -Const. mass flow @choke + mass flow @stall	-
	Pierret [**110**]	RBF+GA+CFD	-Max. Isentropic Eff.+Structural obj @Multipoint (near-stall, near-peak efficiency and choke) →Weights to perform single obj.	-
	Luo *et al.* [**114**]	Adjoint method	-Min. Entropy generation @ Multipoint (near-stall, near-peak efficiency and choke) -Constr.: mass flow + total pressure ratio	-
	Hu *et al.* [**115**]	Inverse method	-Max. Isentropic Eff. @Design point - Constr.: mass flow @Choke	-

Table 5.2: Previous studies on NASA Rotor 67: optimization strategies.

Starting from direct optimization methods, Oyama *et al.* [**109**] minimized flow loss manifested of Rotor 67 via entropy generation with an adaptive range genetic algorithm. According to the authors' experience, maximizing isentropic efficiency would endorse maximization of total pressure ratio rather than minimization of flow loss. Mass flow and total pressure ratio were constrained, being the variation compared to the baseline design of 0.5% and 1%, respectively. After 200 generations with 64 individuals for population (i.e., 12800 CFD

evaluations on the whole!), isentropic efficiency was improved by 1.78% at design point. The entropy production was reduced in regions between hub and mid-span, and near tip. The former effect was due to a diminishing of incidence angle (Figure 5.2), while the latter was determined by the maximum camber position that moves toward aft, resulting in a weaker and shifted aft shock on the blade suction side. Although the optimization was performed at design point, the optimized design still maintained higher isentropic efficiency over the entire range of operating conditions, from the choke to stall, and the stall margin remained unchanged.

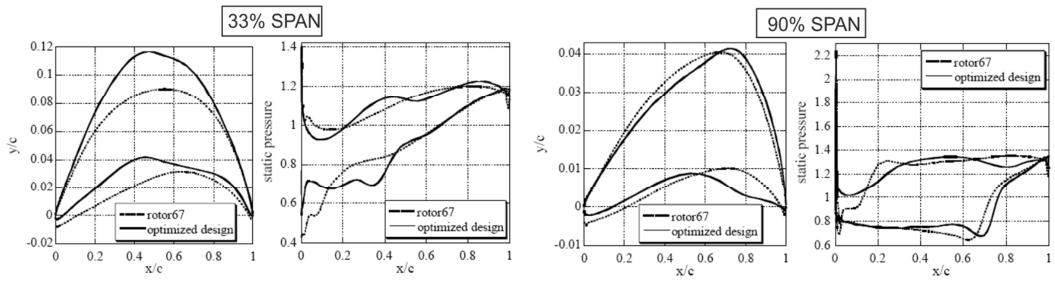


Figure 5.2: NASA Rotor 67 optimized foils (Oyama *et al.* [**109**])

When only aerodynamic objectives are taken into account, it is possible that poor structural shapes, at least in the other part of the blade, are achieved. In fact, small thicknesses in the leading edge region positively affect shock losses, and thus it reflects on global efficiency, but lead to a blade shape which could not satisfy mechanical constraints. Accounting for structural objectives has been demonstrated to be a feasible way to address this problem (Pierret [**110**]). A single objective formulation was defined by weighting aero-structural objectives and constraints at three operating points (e.g., near-stall, near-peak efficiency and choke), and the optimization involved a genetic algorithm assisted by RBF interpolation technique. As expected, higher values of isentropic efficiency were achieved for the aerodynamic optimization compared to the aero-structural one, for all the speed-line.

Lian and Liou [**111**] adopted a different parameterization approach, in which modified rotor blades were defined superimposing perturbation on the original baseline Rotor 67's geometry. One of the advantages of perturbation approach relies on the fact that the baseline geometry is always enclosed in the search space; this is done by setting design variables to null perturbation. The multi-objective optimization was carried out at peak efficiency. The total pressure ratio was

maximized and the entropy generation minimized, while mass flow was constrained within a 0.5% variation with respect to the baseline. Among all the Pareto points, the total pressure ratio increased as much as 1.8% and the entropy production decreased by 6.2%, which implied an isentropic efficiency enhancement of 0.7%.

It is worth nothing that the two aforementioned works did not take into account neither lean nor sweep deformations.

Although this survey is strictly focused on aerodynamic optimization, we could not overlook the work of Lian and Kim [**112**] in which both structural and aerodynamic performance of Rotor 67 were contextually optimized. Total pressure ratio and rotor mass were taken as objectives, while mass flow rate and probability of failure were treated as constrains. It turns out that most of the optimal designs broke the safety constraint, at least with a safety factor a little bit lower than one (above one safety constrain is respected).

Okui *et al.* [**113**] proposed a more flexible 3D parameterization. Three main geometric deformations were permitted: variation of the mean camber line, the spanwise distribution of chord, and sweep. The thickness distribution was maintained the same as the baseline to avoid mechanical issues. A multi-objective evolutionary algorithm coupled with neural networks was adopted as optimizer. The objectives to be maximized were peak isentropic efficiency and throttle margin. In order to accelerate the optimization process and respect the baseline choked mass flow two constraints related with the choked and stall mass flow were imposed, for which the off-design performance curve was needed and evaluated via CFD. Two optimizations were performed. First, only chord and sweep variation were activated in the optimization process; then, mean camber design variables were introduced, and a new optimized rotor was identified. At the end, a complete stage optimization was performed keeping fixed the stator and increasing degrees of freedom of camber line parameterization. As the chord and sweep deformations were activated, the optimal blade had a forward sweep shape and a +0.3% gain in isentropic efficiency. Such shape involved the outward span region and cut the bow shock into a very weak compression shock on the suction side and a weaker passage shock. On the contrary, when the camber line deformation was added, the optimal blade had a backward sweep shape and a higher isentropic efficiency gain equal to +0.6%. The explanation of this result is due to the S-shape camber line and a barreling chord length, which are both able to compensate the negative effects of backward sweep on stability and efficiency.

The last work related to the direct optimization approach is the one of Luo *et al.* [**114**], in which an adjoint method was used. The main advantage of such approach is that computational effort is insensitive to the number of design variables. In this case, the rotor was described with 288 design variables. A single-point and a multipoint optimization were performed in order to minimize the entropy production, while keeping to a minimum the variation of mass flow rate and total pressure ratio compared to the baseline. The single objective multipoint optimization took into account three operating conditions (i.e., near choke, near peak efficiency and near stall), which were aggregated by means of weights. In the single-point optimization, the isentropic efficiency had an increment of about 1.10%, but a noticeable decrease in total pressure ratio for all the operating conditions between stall and peak efficiency. On the other hand, in the multipoint optimization gains of isentropic efficiency were achieved at all three operating conditions, with increments of about 1.24% near choke, 0.84 near peak efficiency, 0.54% near stall; the total pressure ratio was almost unchanged compared to the baseline all over the range.

As far as the inverse optimization method is concerned, we can enumerate three studies related with Rotor 67. The first two works, i.e., the one of Tiow and Zangeneh (2002) [**116**] and Watanabe and Zangeneh (2003) [**117**], are focused on the inverse method itself and its ability to reconstruct the geometry given a target loading, rather than realize a complete optimization. Two different pressure-loading distributions, which were derived from the real loading distribution of Rotor 67 but arbitrary shaped, were tested in order to analyze the repercussions on the isentropic efficiency. The first distribution had a completely aft-loaded characteristic, while the second one was chosen to give a fore-loaded characteristic at the hub of the blade and a middle-loaded characteristic form the mid-span to the tip location. As for the adiabatic efficiency, the second distribution performed better with an improvement of about 0.6% over most of the working range. Moreover, the stall margin was increased by 3%.

In the third work, Hu *et al*. [**115**] optimized the blade loading pressure distribution. For this purpose, the target loading pressure was parameterized at various span wise sections by B-spline control points. The blade geometry deformations affected only the 2D spanwise sections, while lean and sweep modifications were not included. Since the optimization of the isentropic efficiency was performed at a given operating point, the optimized blade behavior throughout the entire operating range was a mere consequence of the outcomes of

the optimization process. The main concern was to control choke flow rate, keeping its value as near as possible to the baseline. Two methods were proposed. First, it was observed that changing the operating point while keeping the same target pressure loading, could lead to a horizontal shift of the characteristic curve, which indeed modify the choke flow rate. Second, the choke flow was affected by the peak position of the optimal pressure loading. An aft-loading resulted in smaller throat areas and hence smaller choke flow, while fore-loading increased the choke flow. The isentropic efficiency of NASA Rotor 67 was increased by 1.26% at design point compared to the baseline, while the choke flow rate was reduced by -0.8%. Most of the improvement comes from 25-100% of span. The normal passage shock in the original rotor at section 0.8 was changed to an oblique shock in the optimized blade.

5.2 Lean And Sweep Deformations

The 3D shock wave established in the blade passage is responsible not only for losses related with both the shock itself and for the shock-boundary layer interaction but also for the stability of the compressor in terms of stall margin. A learned reshaping of the stacking line could positively affect the shape of the 3D shock and, in turn, the compressor performance. The investigations conducted so far have demonstrated that there are three compelling reshaping groups: blade airfoils deformation ([**118**], [**109**]), sweep and lean.

Sweep and lean involve 3D deformations of the blade stacking line and they are defined as the translation of the foil, seen in the conformal plane, along the rotation axis direction (sweep) or tangential direction (lean) (see Figure 5.12). Both lean and sweep can be detailed with *forward* and *backward* adjectives. Different definitions are used in the literature. As for the lean, forward and backward refer to the agreement or not of the stacking line displacement with the sense of rotation of the blade. On the other hand, we refer to forward or backward sweep depending on the displacement direction, which could point in the upstream (forward sweep) or downstream (backward sweep) direction.

As far as sweep is concerned, backward sweep was first investigated by Hah and Wennerstrom [**119**] on Rotor 6, an aft-swept rotor developed to strengthen the sweep effect, which demonstrated a remarkable improvement in peak efficiency but a detrimental reduction in stall margin. Such penalization was caused (Wadia *et al.* [**120**]) by the local increase of loading at the tip section resulting in a stronger bow shock, in addition with a more intense migration of fluid particles on

the suction side boundary layer (a secondary flow that follows the imbalance between the centrifugal force and the pressure gradient). On the other hand, Hah *et al.* [**121**] showed the advantages of a forward-swept rotor in terms of higher margin stall and higher peak efficiency compared to the baseline unswept rotor. Denton and Xu confirmed similar results via CFD [**122**].

Despite blade sweep is known to be an effective technique to redistribute the radial loading [**123**], sweep also affects the meridional position of the shock, which tends to resemble the shape of the sweep. Moreover, it has been observed the so-called "endwall effect", this is the shock approaches the casing at right angles (Hah *et al.* [**121**]). Because of the superimposing of the sweep and the "endwall effect", near the casing the position of the shock is moved downstream in the case of forward sweep whereas upstream in the backward sweep, as depicted in Figure 5.3. Usually, a more backwards position of the shock reflects a better stability, and this partially justifies the improvement in stall margin observed with forward-swept rotors.

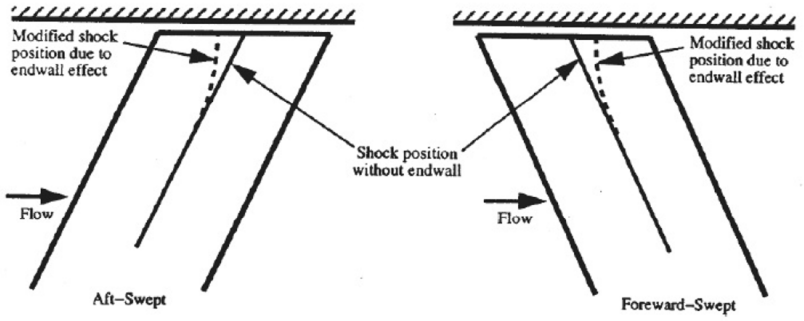


Figure 5.3: "Enwall effect" on shock structure near casing [**124**].

Investigations on the lean have highlighted that its use can produce favorable effects thanks to the change on the 3D shock structure (Bergner *et al*. [6]). In particular, forward lean (i.e., towards the direction of rotation) resulted to have positive influence on the overall rotor efficiency (Ahn and Kim [7] and Benini [8]). Moreover, an important investigation on the aerodynamics of swept and leaned transonic compressor rotors conducted by Biollo and Benini [**125**].

5.3 Blade Geometry

The geometry provided in NASA report [**108**] comprise 14 blade spanwise sections, which were formatted in a proper way to be given to ANSYS®

TurboGrid in order to reconstruct the baseline geometry. The result of such operation is illustrated in Figure 5.4.

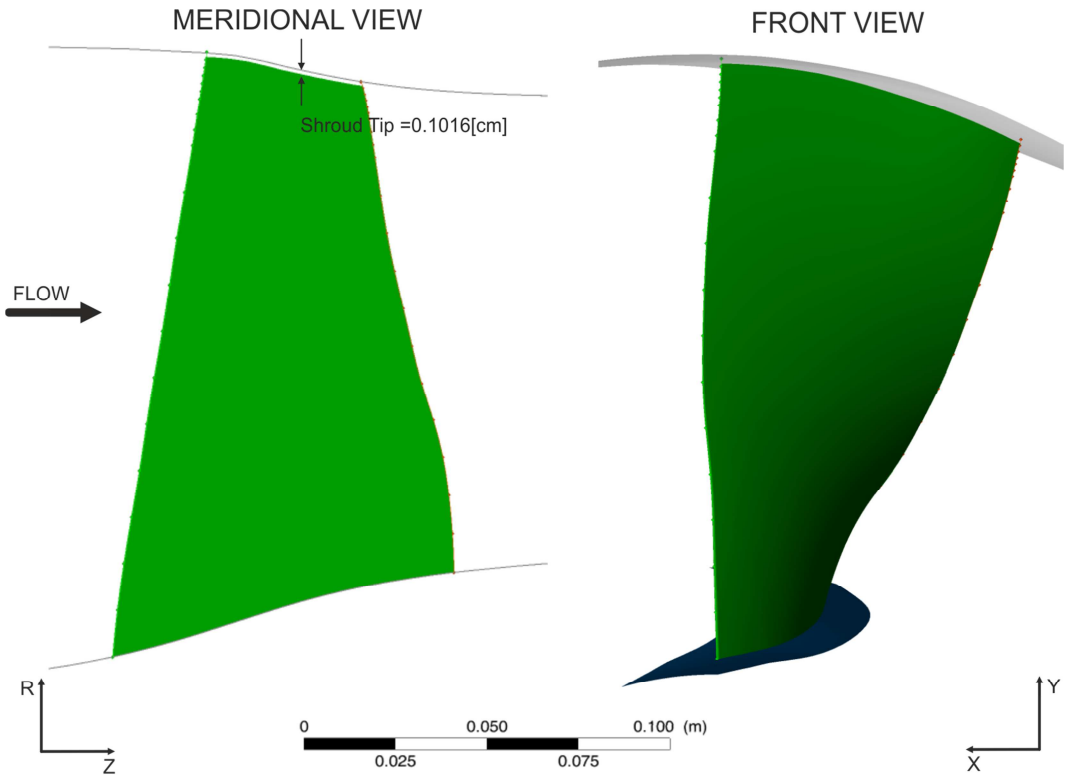


Figure 5.4: R67 geometry.

5.4 Flow Solver And Computational Domain

In this section, the numerical model setup is described, which comprise the computational grid and the CFD solver setup.

5.4.1 Grid

As far as the numerical grid is concerned, the Automatic Topology and Meshing tool (ATM optimized) within ANSYS® TurboGrid v14 was used to generate a multi-block structured grid. As depicted in Figure 5.5, there are three blocks: inlet, passage and outlet. The passage block is critical because grid quality is strongly dependent on how the grid topology is anchored to the blade geometry, being this match enforced by moving the topology control points. In an optimization context, where blade geometry could have important deformations, the meshing tool should be enough robust to guarantee high mesh quality also for geometries far from the baseline one. Such tool should be able to adapt the control

points' position to the new geometry. As deformations increase, the passage block mesh could quickly deteriorate, at least keeping unchanged the topology control points. It turns out that ATM could handle this problem by optimizing control points position and by smoothing the mesh, achieving a high grid quality even for distorted geometries.

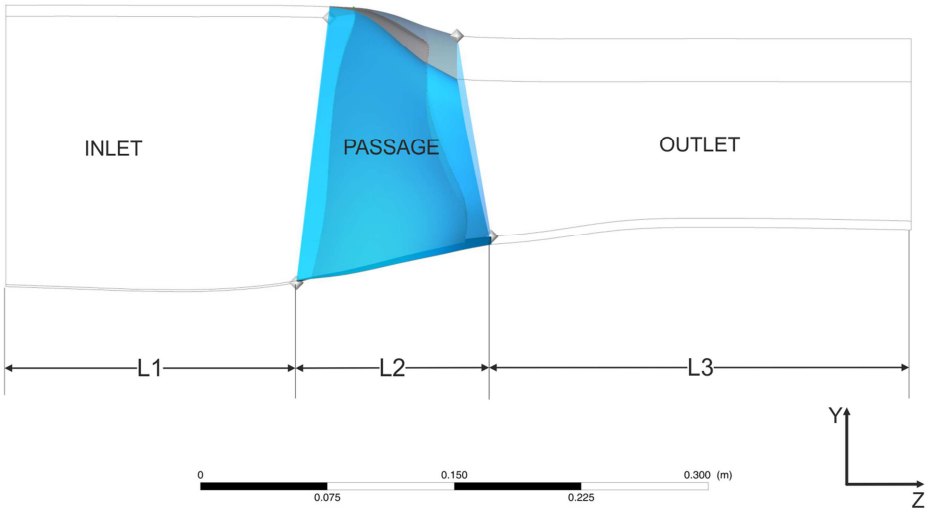


Figure 5.5: Computational domain.

The main grid parameters were deduced from those adopted in the validation analysis carried out in [**126**]. In particular, the structured grid of the passage block were created interpolating 10 spanwise layers, as those depicted in Figure 5.6, which are in such high number to reduce the mesh distortion when the baseline geometry is modified. The topology is ATM based, which indeed does not match with any of the standard topologies. The *target passage mesh size* method with a target value of 1.7 M elements were prescribed. The first layer height were selected in order to achieve a y+ lower than unity all over the wall surfaces imposing 2e+06 as reference Reynolds number .In the spanwise direction 110 elements were adopted, in addition 42 more nodes in the shroud tip region. As for the inlet and outlet blocks, H-grid topology with 28 and 34 streamwise elements respectively was selected. At the end, the whole mesh was composed by 2.2M elements.

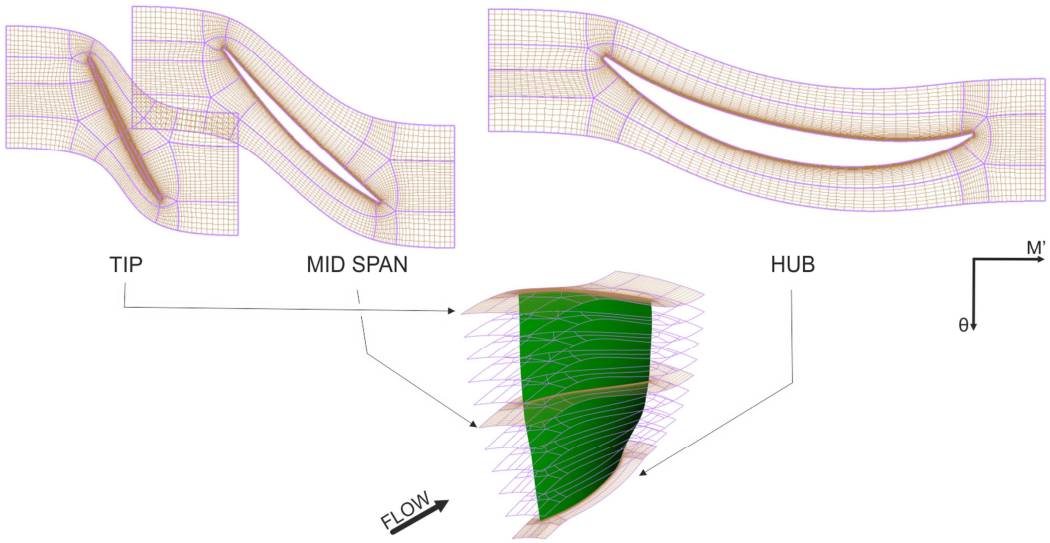


Figure 5.6: ANSYS® TurboGrid topology: layers of hub, midspan and tip section.

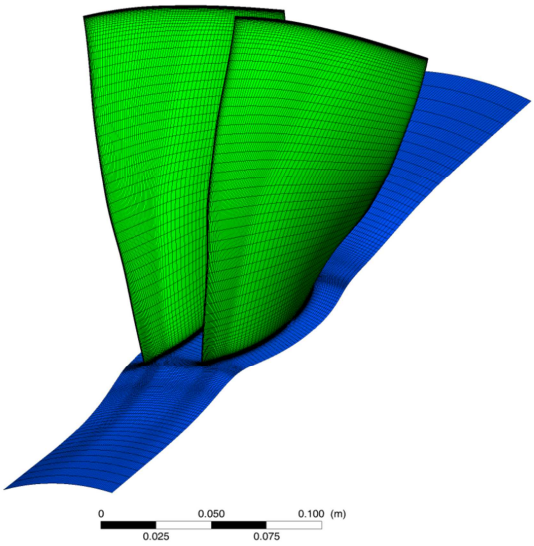


Figure 5.7: Computational grid.

5.4.2 Flow solver Setup

The steady state 3D flow field around the blade were computed by means of the commercial CFD code ANSYS® CFX v14, in which the Reynolds-averaged form of the Navier–Stokes equations are solved using a finite-element based finite-volume method.

The computational analysis comprised one blade passage and a periodic condition was applied on lateral passage surfaces. The flow was fully turbulent

and the k-ω SST [**104**] turbulence model was adopted. All the computational domain rotated at -16043 [rpm] along the z axis. The setup of the CFX-Pre is summarized in Table 5.3.

It is worth nothing that convergence was critical throughout all the calculation, so a *user function* was implemented with the purpose of adapting the maximum timescale, which indeed is of utmost importance for calculation stability. In particular, at the beginning a value of 1e-007 [s] was set, then it was increased up to 1e-004 [s], for then return to lower values around 1e-005 [s] before calculation was stopped.

The convergence was established when the RMS maximum residue were lower than 1e-005 and the variables of interest described in the next paragraph had an asymptotic behavior.

ANSYS© CFX solver setup		
Formulation	Steady State	
Domain	Material	Air Ideal Gas
	Domain Motion	-16043 @z-axis
	Reference Pressure	0 [atm]
	Heat Transfer	Total Energy
	Turbulence Model	k-ω SST + high speed
BC		-Inlet + frame stationary
	Inlet	-Subsonic
		-Stationary Tot. Pressure 101325 [Pa]
		-Direction normal to boundary
		-Stationary Tot. Temp. 288.2 [K]
		-Turbulence intensity 5%
	Outlet	-Outlet + frame stationary
		-Subsonic
		-a)Average Static Pressure 50000[Pa]
		-b)Mass Flow Rate
	Blade	-Wall + no slip
		-Adiabatic
	Periodic	-Periodic
		-Conservative Interface Flux
	Hub	-Wall + no slip
		-Rotating frame
		-Adiabatic
	Shroud	-Wall + no slip
		-Rotating frame
		-Counter Rotating wall
		-Adiabatic
Solver Control	Advection Scheme	High resolution
	Turbulence Numerics	High resolution
	Timescale Control	Auto timescale

Table 5.3: ANSYS® CFX solver setup.

5.4.3 Calculation Of The Variables Of Interest

Performance was calculated at specific upstream and downstream survey stations, which are shown in Figure 5.8. These stations were the same of those adopted in the NASA Report [**108**], for which experimental data are available.

Total pressure ratio and polytropic efficiency were evaluated taken into account quantities calculated with a *mass-weighted average* surface integral at the appropriate survey stations. Specifically, the total pressure ratio was defines as

$$\pi_c = \frac{p_{02}}{p_{01}} \qquad \text{Eq. 5.1}$$

where p_{02} and p_{01} are the total pressure in the stationary frame at station 2 and 1, respectively, extracted with the aforementioned mass-weighted integral. On the other hand, the polytropic efficiency was defined as

$$\eta_{pol} = \frac{\left(\frac{\gamma-1}{\gamma}\right)\ln\left(\pi_c\right)}{\ln\left(\frac{T_{02}}{T_{01}}\right)} \qquad \text{Eq. 5.2}$$

in which T_{02} and T_{01} are the total temperatures in stationary frame at station 2 and 1, respectively. For completeness, the isentropic efficiency was defined as:

$$\eta_{is} = \frac{\pi_c^{\left(\frac{\gamma-1}{\gamma}\right)}-1}{\frac{T_{02}}{T_{01}}-1} \qquad \text{Eq. 5.3}$$

In addition to global variables, spanwise distributed performance is of utmost importance because it can give a better understanding on which part of the blade affects in greater amount global performance. Given a spanwise location, the value of a generic spanwise variable was representative of the pitchwise circular arc. Similarly to the global variables, first total pressure and total temperature were evaluated with the mass-weighted integral on the surface associated with the pitchwise circular arc, and then total pressure ratio and polytropic efficiency were calculated for that specific spanwise location. In such process, the quantities belonging to section 1 were considered constant along the entire span.

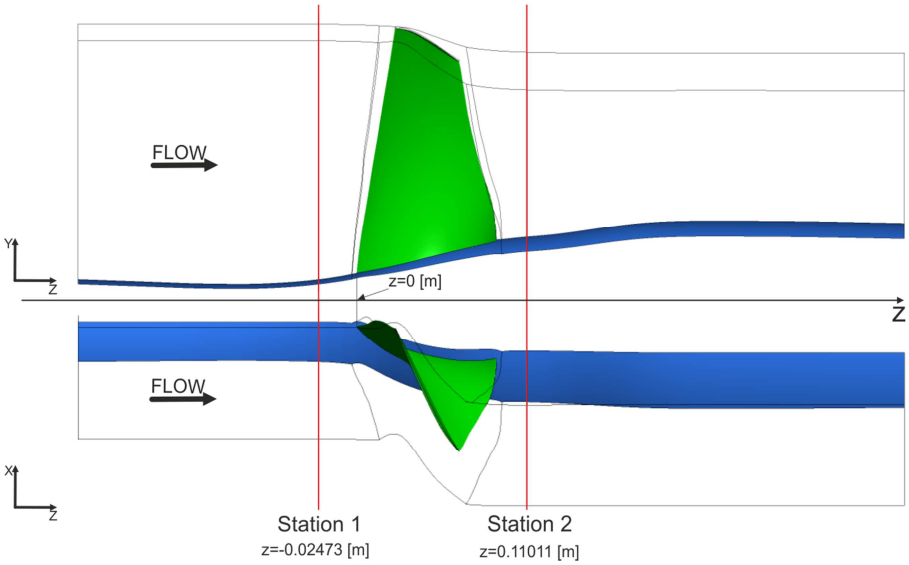


Figure 5.8: Survey stations for performance calculation.

5.4.4 Validation

The sensitivity analysis and the validation of the CFD model were previously carried out in [**126**]. Hereafter the main outcomes of the reconstructed CFD model employed in the optimization loop are reported. Figure 5.9 shows the isentropic efficiency and the total pressure ratio against the normalized mass flow. It is common practice to normalize the mass flow rate with the choking mass flow. At station 2 the spanwise distribution of the exit flow angle is reported in Figure 5.10, while in Figure 5.11 the blade-to-blade Mach contour at 90% of span is depicted.

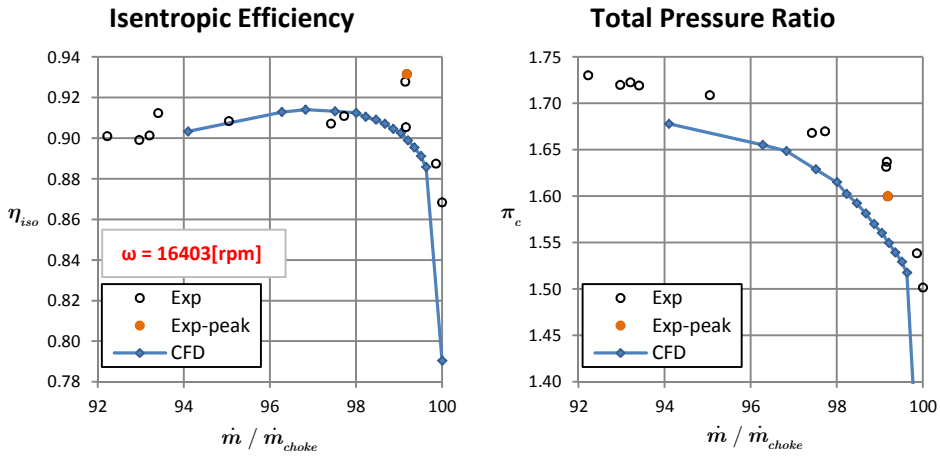


Figure 5.9: Main performance of the NASA Rotor 67.

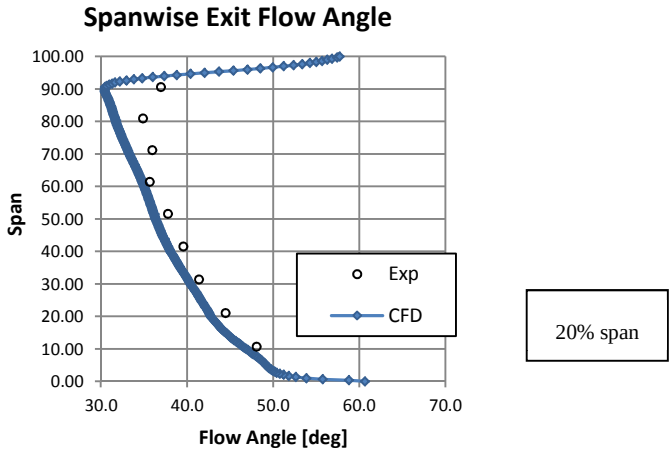


Figure 5.10: Exit flow angle of the NASAS Rotor 67 at peak efficiency.

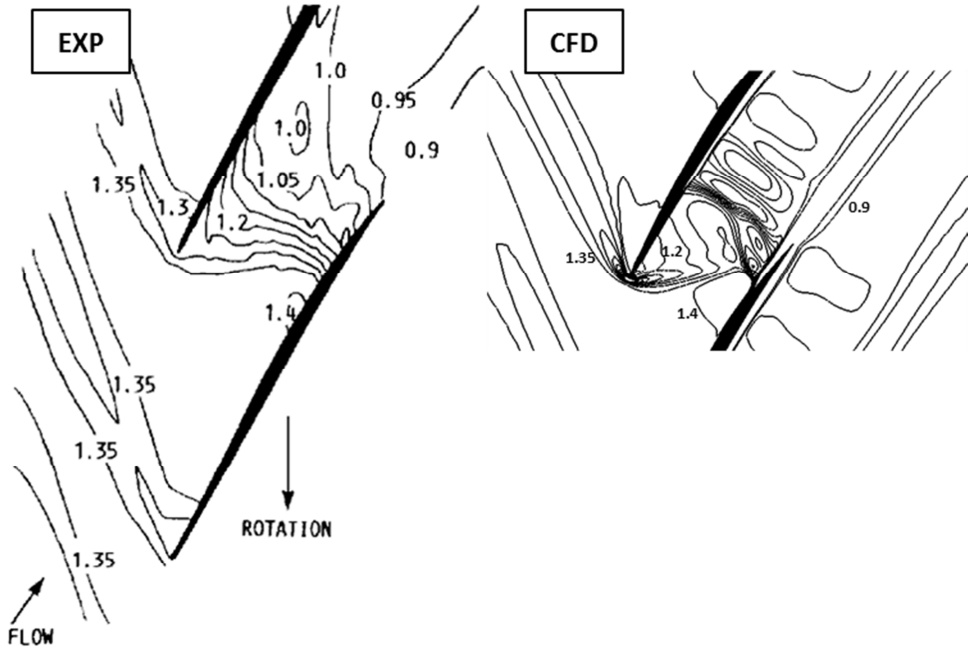


Figure 5.11: Mach number contour at 90% span of NASA Rotor 67.

5.5 Parameterization

Shape deformation was accomplished by superimposing a displacement field to the baseline geometry points. The reader is referred to § 4.4 for a detailed explanation of the method. Briefly, the displacement field is described by a B-spline surface that provides the displacements in the three directions, i.e., *dx*, *dy*, *dz*, for each point of the baseline geometry.

5.5.1 B-Spline Surfaces: Brief Review

A B-spline surface is the extension of the B-spline concept to 3D space. The reader is referred to the work of Mortenson [**106**] for a detailed explanation of B-spline surfaces and their numeric implementation. The surface equation is defined as the tensor product:

$$\boldsymbol{p}(u,w)=\sum_{i=0}^{m}\sum_{j=0}^{n}P_{i,j}B_{i,K}(u)N_{j,L}(w) \qquad \text{Eq. 5.4}$$

, which is parametric in the natural parameters *u* and *w;* the curve $\boldsymbol{p}(u,w)$ is defined by the polyhedron control points $P_{i,j}$ and the basis functions polynomials $B_{i,K}(u)$ and $N_{j,L}(w)$, which depends on the specific i-th or j-th control point, but their degrees *K* and *L* are independent on the number of control points. The basis function are defined in the same way as for the two dimensional case (see § 4.4.1).

5.5.2 Parameterization Framework

In this work, blade parameterization is the result of superimposing two main deformations: 2D profile deformation and 3D stacking line deformation, each of which is treated independently with B-spline curves; afterwards, they are joined together forming the final B-spline surface displacement field.

For this purpose, six spanwise sections located at 0%, 30%, 61%, 92%, 97%, 100% of span were selected as *control sections*. Each section is obtained by intersecting the baseline blade with the streamline surface at the given span, being such surface generated as revolution of the geometric streamline curve.

5.5.3 Sweep And Lean Deformations

2D translations of control sections affect the entire 3D shape of the staking line (Figure 5.12). A pure sweep would produce modification only on the meridional plane, whereas a pure lean would reshape the stacking line only on the tangential plane. Generally, modified blades involve a mix of the two.

Two B-spline curves were built interpolating the displacement values at the six control sections. These two curves are sufficient to describe the displacement field spanwise. In fact, at a given span, they act on the control section foil, which is forced to lie in the conformal plane while translation imposed by the displacement filed is applied.

Such a displacement field involves 12 control points, 6 for lean and 6 for sweep. However, in order to achieve an effective staking line deformation, it was decided to use the control points to define *shapes*, each of which imposes relationships between control points that became no more independent one another. As shown in Figure 5.13, lean and sweep shapes are still independent, and they can affect the blade in a local (near tip) or global manner. In the whole, 9 shapes were defined and adopted as design variable in the optimization process.

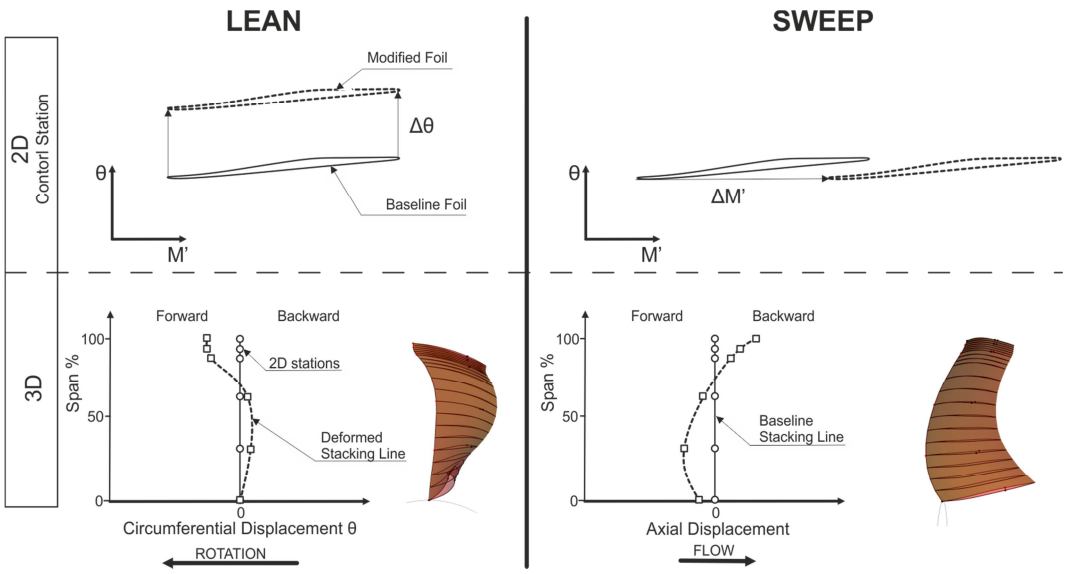


Figure 5.12: Lean and sweep definition.

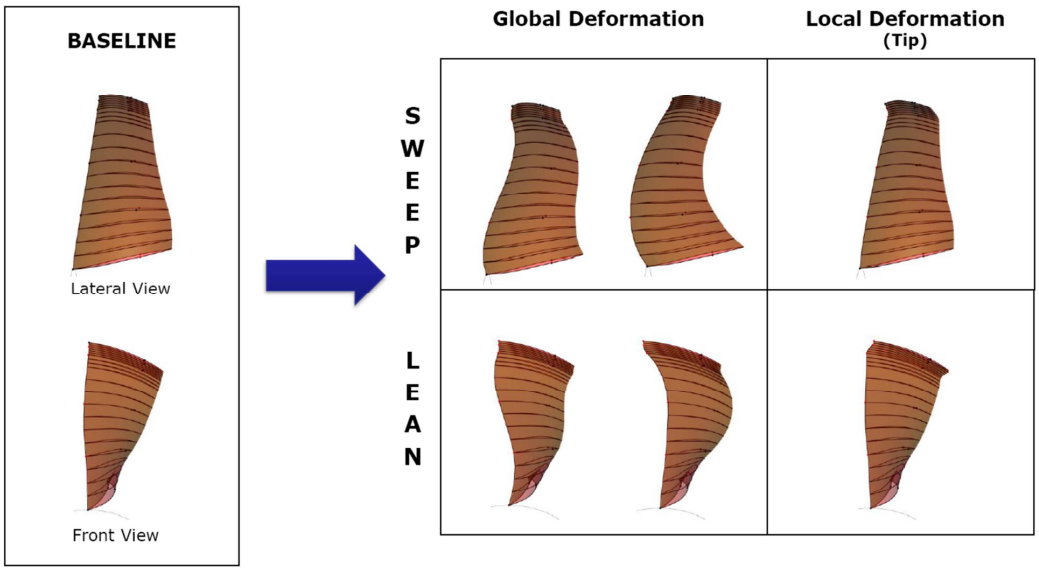


Figure 5.13: Blade parameterization: global and local deformations of sweep and lean.

5.5.4 Foil Deformations

Each of the control sections foil was parameterized with a displacement field approach having several control points. Similarly, 2D foils were deformed using the *shape* approach, which allows reducing the number of design variables while keeping promising deformations within the search space. Of course, this was intended to be a trade-off between computational resources and width of the search space.

The shapes were defined with the aim to decouple the maximum thickness from the camber line shape as best as possible. In order to keep to a minimum the number of design variables, only two shapes were adopted. Figure 5.14 shows how the hub section could be reshaped applying high and low values of its two shape control points. Such high displacements are just to highlight the capability of the parameterization approach, and are not representative of deformations adopted in the optimization process. Moreover, the chord value was kept unchanged.

In the whole, 2 shapes x 6 control sections add to 12 design variables, plus the 9 of the stacking line parameterization leads to 21 design variables in total.

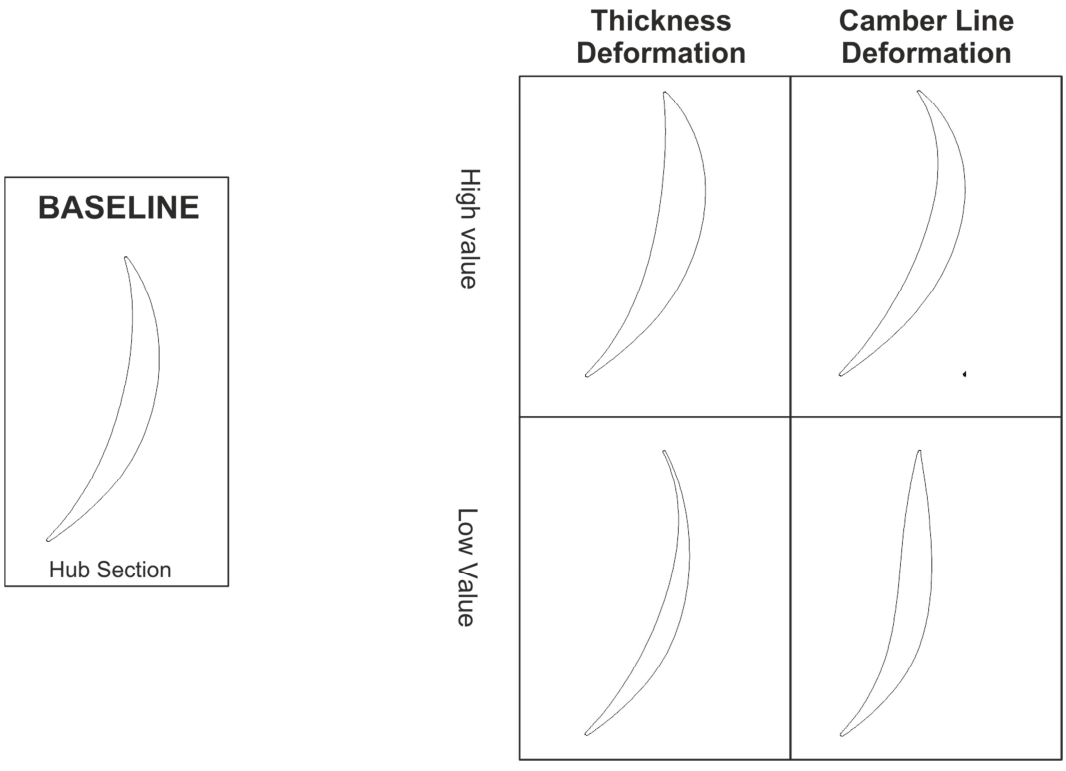


Figure 5.14: 2D foils deformations.

5.5.5 Spline Model Of The Entire Blade

A brief description on how the B-spline surface comprising the two deformation categories, i.e. staking line and 2D foil, is presented. The modified geometry is the result of the baseline geometry perturbed by the displacement field:

$$\boldsymbol{p}_{\mathrm{mod}}(u,w) = \boldsymbol{p}_{base}(u,w) + \boldsymbol{p}_{displ}(u,w) \qquad \text{Eq. 5.5}$$

where the u and w are the natural parameters of the B-spline surface, the former goes around the foil, being 0 at trailing edge, 0.5 at leading edge and 1 at trailing edge; while the latter spans from hub to shroud, being 0 at hub and 1 at shroud. In particular, $\boldsymbol{p}_{base}(u,w)$ represents the high fidelity approximation (i.e., the number of control points are of the same order of magnitude of the points) of the baseline blade, which is defined by 14 spanwise sections.

On the other hand, $\boldsymbol{p}_{displ}(u,w)$ is defined as:

$$\boldsymbol{p}_{displ}(u,w) = \boldsymbol{p}_{lean}(w) + \boldsymbol{p}_{sweep}(w) + \boldsymbol{p}_{foil}(u,w) \qquad \text{Eq. 5.6}$$

in which both the lean and sweep displacement fields are function of the only spanwise location, whereas the foil displacement obviously depends on the foil we are looking at.

5.6 Formulation Of The Optimization Problem

The aim of the multi-objective optimization problem was to minimize the two-objective function at a specific design mass flow rate:

$$\boldsymbol{F}(\boldsymbol{P}) = (f_1, f_2) = (-\eta_{pol}, -\pi_c) \qquad \text{Eq. 5.7}$$

where $\boldsymbol{P}$ is the vector of the optimization parameters or decision variables of the design (i.e., the control parameter of *shapes* describing the displacement field), η_{pol} is the polytropic efficiency calculated at the outlet station in Eq 5.2 and π_c the total pressure ratio. The optimization was run at the 98% of the choking mass flow. A generic new design was first simulated at choking condition imposing a very low static pressure outlet and then, starting from this solution domain, an outlet mass flow BC was adopted to reach the 98% of its choking mass flow.

The constraints complete the problem formulation. A penalty function approach was adopted to handle the violation on the choking mass flow, for which a variation of 0.5% with respect to the baseline was acceptable. The penalty term ***Q*** depends linearly on the level of violation of the constraints:

$$\boldsymbol{Q}=(\mathrm{Q}_{1,}Q_2)=(r_1,r_2)\cdot G$$

$$G=\max\left\{\frac{\left|\dot{m}_{choke\_bl}-\dot{m}_{choke\_new}\right|}{\dot{m}_{choke\_bl}}-0.005,0\right\} \qquad \text{Eq. 5.8}$$

No geometric constraints were imposed, since the range of the decision variables was set in such a way that unfeasible geometries were avoided. It is worth noting, however, that the range was intentionally set very wide with the aim at reaching the most disparate configurations.

5.7 Discussion Of Results

The optimization was carried out by means of ASEMOO (see §2.2.2). In the whole, 210 CFD direct evaluations were performed. The wall-clock time for each CFD computation was of 36 hours on an 8 core INTEL E5-2650 2.0GHz.

The total pressure ratio (TPR) and the polytropic efficiency (E) were the two objectives to be maximized at the operating point $\dot{m}/\dot{m}_{choke}=98\%$. In the Pareto-optimal front, the optimal designs were much clustered in a narrow region characterized by high TPR and moderate E. Benini [**46**] also recorded this clustered behavior due to the constraint imposed on the choked mass flow; in other words, it seems that for this optimization problem the Pareto front is inherently narrow. Moreover, Oyama *et al.* [**109**] stated that taking into account the efficiency as objective inherently pushes the search towards high total pressure ratios designs. Thereby, considering the total pressure ratio as second objective could be in some extent repetitive, but not wrong from a theoretical point of view, and could introduce a further difficulty for the optimizer to spread designs over the just narrow Pareto-optimal front. In addition, we believe that the causes behind the clustering are to be also searched in the extensively width of the search space against a low number of CFD direct evaluations. In fact, the parameterization was deliberately tailored in such a way to guarantee a huge search capability, aware of the complexity introduced for the optimization algorithm. From this point of view, it is not surprising that the Pareto-optimal

front encloses designs that have similar shapes and thus similar performance. Still, the improvements in TPR are remarkable and it is very likely that a further exploration could arrive at substantial results. If we include designs that slightly violate the constraint on the chocking mass flow, the Pareto-optimal front is enlarged by higher E designs.

The results presented hereafter involve two optimal designs. The first one is TPR optimized (TPR-O), which belongs to the original Pareto-optimal front; it increases TPR by 3%. The second design is E optimized (E-O) with a choking mass flow 1% higher than the baseline rotor. Its E is 0.9% higher than Rotor 67 (see Table 5.4). In Figure 5.15, the speed lines of the baseline and the optimized geometries are reported. Both blades keep their performance enhancement over the entire operating range. However, the stall margin of both designs had a severe drop compared to the baseline one. In Figure 5.15 the points further to the left of each speed line are representative of those CFD computation beyond which convergence was not achieved due to numerical instabilities. Such instabilities of the flow domain may be associated with the near-stall condition.

The 3D blade and the blade-to-blade sections at hub, midspan, and tip of the baseline, TPR-O and E-O configurations are shown in Figure 5.17. The TPR-O blade had a strong modification in both camber and maximum thickness along the entire span, while the stacking line did not differ so much from the baseline. In particular, the blade was characterized by higher cambered profiles toward rear in the span region from hub to midspan; the opposite behavior was recorded in the outer span region. This reshaping was responsible for the increased loading at hub and midspan locations and a reduction in the other region, as can be observed in Figure 5.16.

As for the E-O design, the reshaping involved mainly the stacking line and only lightly the airfoils. The lean conferred an s-shape profile to the blade leading edge. Starting from the hub, the lean was first forward until midspan and then backwards; in the proximity of the tip, a strong lean was observed in the direction of rotation, creating a sort of “horn”. As for the sweep, the blade had a forward sweep from hub to midspan and then went backward. As it is well known, the sweep is the major responsible for the increase in choking mass flow and this was the main reason why the E-O design had 1% increase in choking mass flow compared to the baseline. Finally, at the hub section a higher camber and slightly lower thickness compared to the original design was observed.

Designs	T.P.R.	Polytropic Eff.	$\dot{m}_{choke} / \dot{m}_{choke\_bl}$
Baseline	1.615	0.9181	1
TPR-O	1.664	0.9244	1.0032
E-O	1.645	0.9266	1.0104

Table 5.4: Performance at optimization condition.

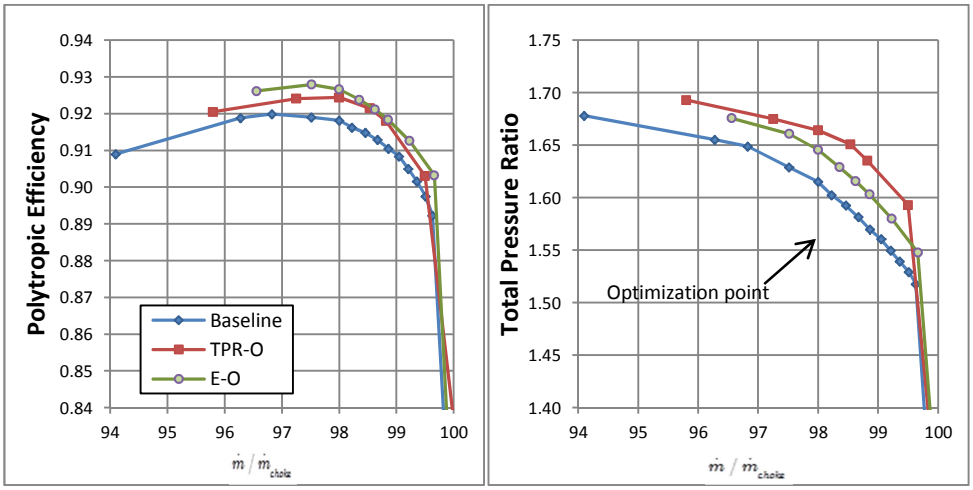


Figure 5.15: Performance maps of the baseline and optimized configurations.

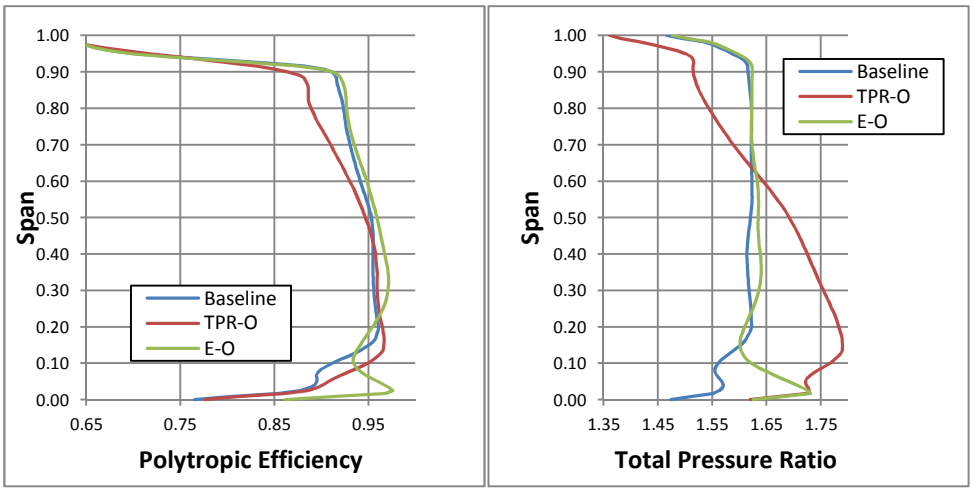


Figure 5.16: Spanwise distribution of polytropic efficiency and total pressure ratio of the baseline and optimized configurations.

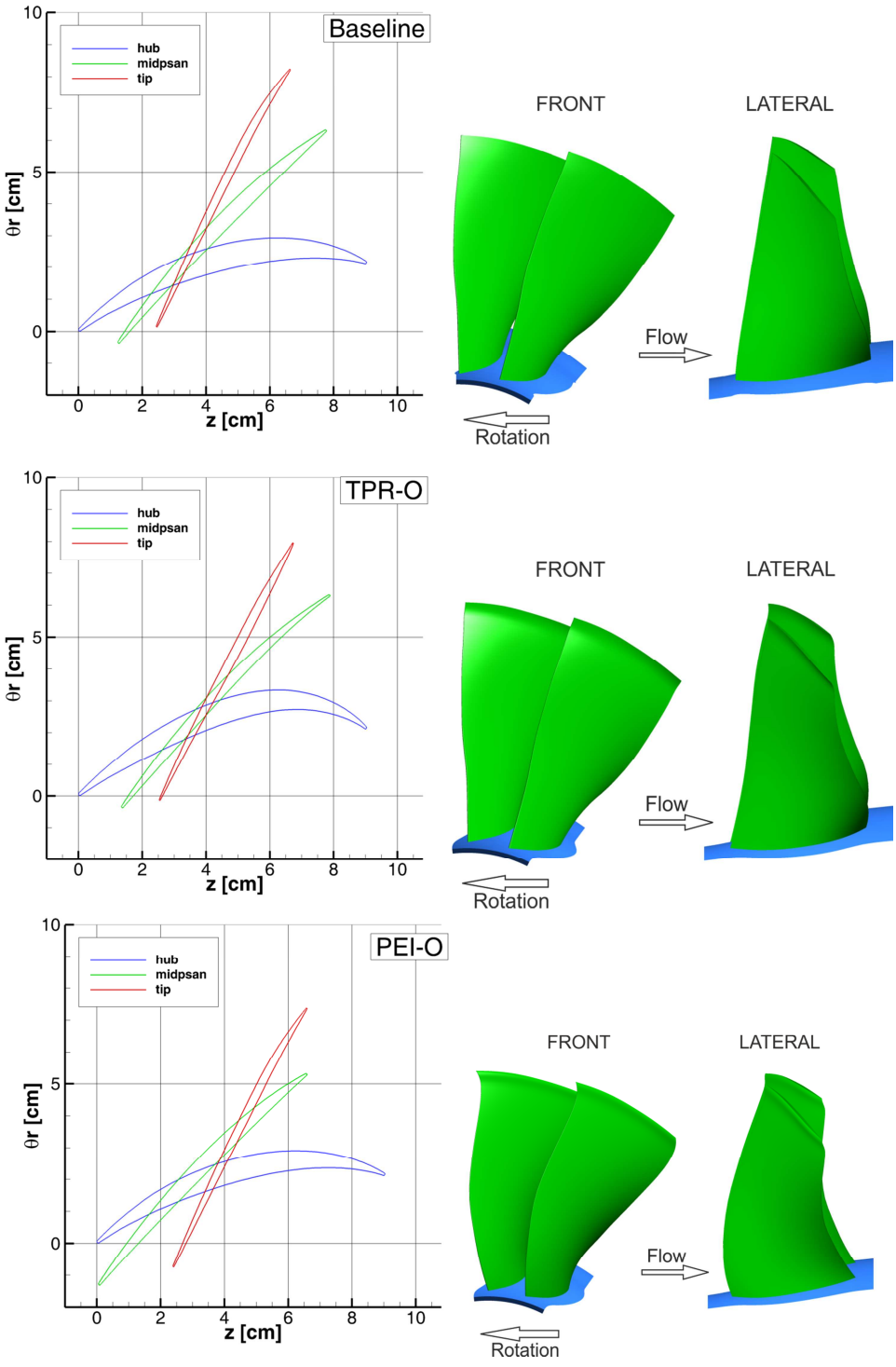


Figure 5.17: Blade geometries of the baseline and optimized configurations.

In order to gain insight on the improvements of the two optimal designs, Mach contours on blade-to-blade sections and on the blade surface are reported in

Figure 5.18 and Figure 5.19, respectively. For all three configurations, the passage shock wave approached the case with normal angle, as expected.

Compared to the baseline, the TPR-O design had a stronger passage shock wave near the tip, shock that was characterized by a “quasi-normal” inclination with respect to the incoming flow. The interaction between the severe shock and the suction side boundary layer gave rise to a remarkable boundary layer detachment resulting in impressive wave losses. Furthermore, TPR-O did not accomplish further diffusion of the subsonic flow after the “quasi normal” shock. Both the compression mechanisms as the second passage shock, which was visible in the baseline configuration, and the subsonic flow turning due to airfoil camber, were not implemented. Therefore, the flow was reaccelerated in the passage (look at the Mach contour on the pressure side Figure 5.19) and the compression effect of the first passage shock vanished. As a result, detrimental performance was monitored in terms of both the spanwise distributions of total pressure ratio and polytropic efficient. Moving toward hub, the deficit of polytropic efficiency between 80% and 45% span was mainly due to the passage shock that remained very strong compared to the baseline till 50% span. From 65% span to hub, TPR-O design performed a higher total pressure ratio mainly thanks to the highly curved rear camber that imposed a higher turning on the subsonic flow. Furthermore, the hub corner stall featured by the baseline was here partially absorbed by the change in the airfoil shape, both maximum thickness and camber, but still a conspicuous wake was release.

As far as the E-O design is concerned, the passage oblique shock near tip was similar to the baseline, except for the slightly lower obliqueness featured by E-O that induced higher shock losses. Nevertheless, a weaker second passage shock (look at the Mach contours on pressure side in Figure 5.19) mitigated the effect of first shock in such a way that the efficiency of the outer span was slightly higher compared to the baseline. In this case, the shock/boundary layer interaction was not as severe as the one of TPR-O and was comparable with the one of the baseline. The sweep introduced by E-O design affected the location of the first passage shock, which resembled the sweep deformation on the blade suction side and, consequently, increased its obliquity with respect to the incoming flow, thus reducing the shock losses. As confirmed by the Mach contours on the suction side and the spanwise distribution of the polytropic efficiency, the passage shock was forked around 80% span in to branches of lower intensity that advantaged the efficiency, while contextually supporting a significant compression effect. As

reported in many studies, the substitution of a strong shock wave with a pair of double weaker shocks is an optimum compromise between efficiency and compression effect. Finally, at the inner region near the hub, the airfoil reshaping makes the difference: for E-O design the corner stall was almost entirely absorbed and, compared to the TPR-O design, the rear camber was more soft, thus reducing the losses in the wake (see span 10% in Figure 5.18) and achieving a higher efficiency.

As far as the stall margin is concerned, the high freedom permitted to the parameterization may lead the optimization process to find very impressive configurations at design conditions, but less appealing in off-designs operations. Unfortunately, the computational efforts needed to reach the near-stall condition (keep it simple with steady state simulation) was not affordable in this work. However, as suggested by Okui *et al.* [**113**], a multipoint optimization for sure gives something extra and should be adopted when a deep redesign of the blade is accomplished.

It must be very clear that the results presented corroborate the fact that the optimization process needs more CFD computations to reach high-level results. Nevertheless, the optimal designs that were found so far gave a snapshot over some of the compelling mechanisms involved in transonic compressors.

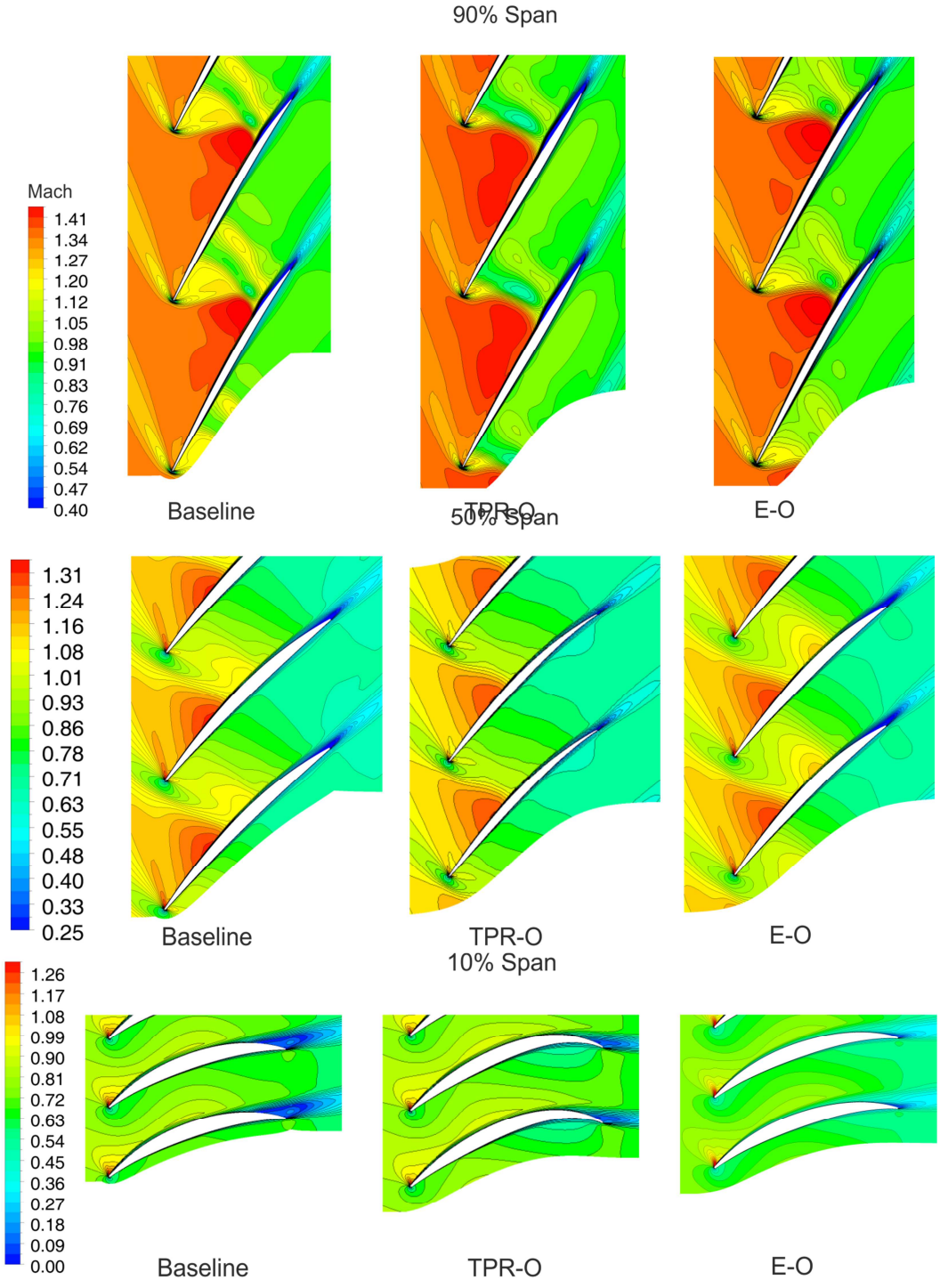


Figure 5.18: Mach number contours at 90, 50 and 10% span of baseline and optimized geometries.

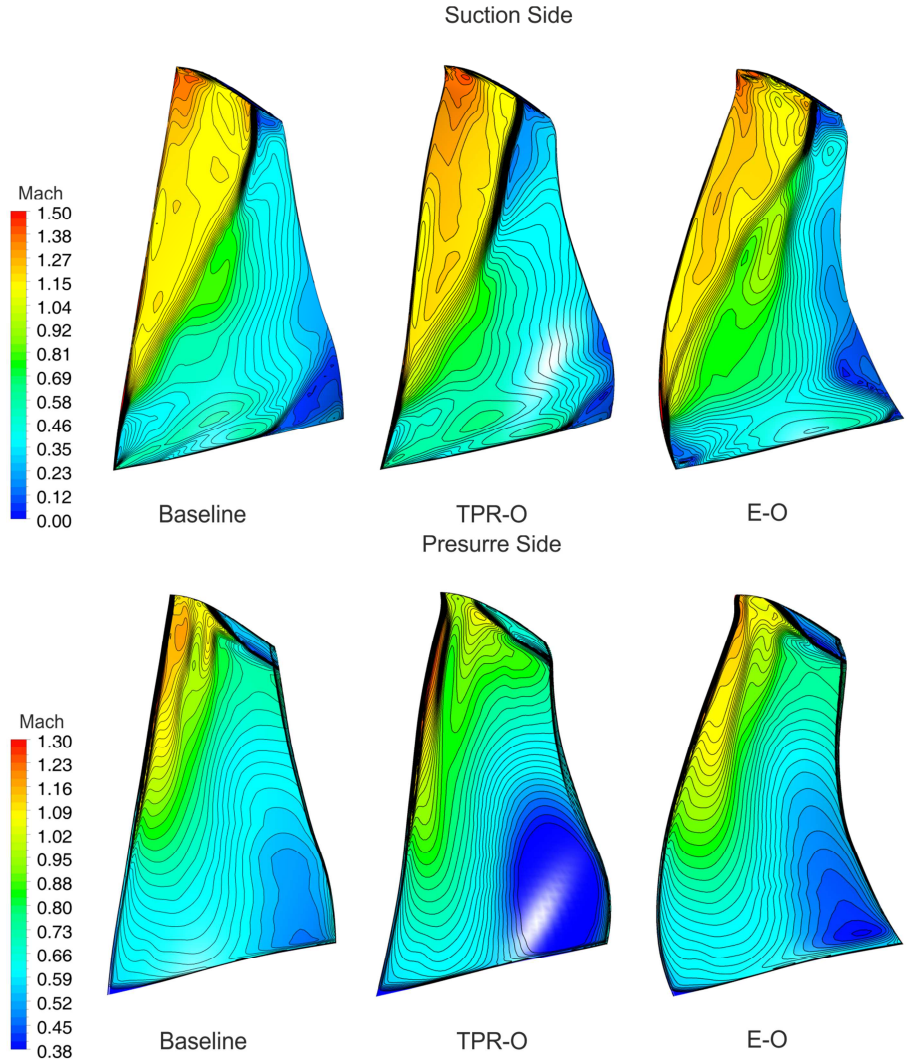


Figure 5.19: Mach number contours on the surface of the baseline and optimized blades.

5.8 Conclusions

The multi-objective shape optimization of the NASA Rotor 67 transonic compressor was carried out by means of a Kriging-based optimizer, ASEMOO, coupled with a CFD solver. The polytropic efficiency and the total pressure ratio were maximized at a specific mass flow rate condition, while the choking mass flow was constrained. A complete and deep reshaping of the rotor geometry was accomplished by means of a generalized parameterization framework that involved sweep and lean deformations, in addition to the modification of six spanwise sections. In order to reduce the complexity of the problem, but still allowing a huge exploration capability, 21 shapes were cleverly defined and were accounted as optimization variables.

Due to the limited number of direct CFD evaluations joined with a wide range for the decision variables, the Pareto-optimal front was clustered in a narrow region characterized by high values of total pressure ratios and little efficiency improvement. Two Pareto-optimal designs were analyzed, being representative for the maximization of each objective. The first optimal design increased the overall polytropic efficiency by 0.9% with a contextual increment of total pressure ratio of 1.8 by giving the blade a proper back sweep and by slightly changing the airfoil camber at the hub. The second optimal design featured a strong recambering of the entire blade, giving more deflection at the inner regions and unloading the outer span. Both the optimal design had a severe reduction of the stall margin compared to the baseline.

The results corroborate the fact that the optimization process needs more CFD computations to reach high-level results. Unfortunately, computational power was in small amount for this work. Nevertheless, the optimal designs found so far give a snapshot over some of the compelling mechanism involved in transonic compressors, demonstrating the effectiveness of the entire optimization strategy.

Conclusions And Future Work

In this Thesis, the shape optimizations of a supersonic compressor cascade and the transonic compressor NASA Rotor 67 were conducted by means of an automatic CFD-based optimization loop.

In order to carry out high speed turbomachinery optimizations, a novel Kriging-assisted genetic algorithm for multi-objective optimization problems, namely the GeDEA-II-K, was specifically developed, with the aim at speeding up the optimization process by taking advantage of the surrogate model. The cooperation between the GeDEA-II genetic algorithm and the ASEMOO Kriging-based algorithm is realized by means of the Kriging filter featured in the GeDEA-II-K algorithm. The comparison over two- and three-objective test functions revealed the effectiveness of the GeDEA-II-K, which is able to unite GeDEA-II and ASEMOO in such a way to avoid destructive interactions between the two and to synergistically intensify the strengths of both. The comparison made use of very few direct evaluations, which resembles the real-life application where high costly models as CFD are used.

When optimization is employed in research, the optimal designs are the primary outcomes of the optimization process, for sure. However, besides the designs themselves, it is certainly more important what such optimal configurations can tell us about the physics behind their improvements. An impressive example came from the multi-objective optimization of a supersonic compressor cascade. In this investigation, the pressure ratio and the cascade loss coefficient were taken as objectives. A substantial improvement of both objectives was achieved: among the Pareto designs, the loss coefficient was reduced by 25% and the static pressure ratio was raised by 6.5%. It was demonstrated that the forepart geometry all alone was responsible for the 95% of the overall loss reduction, thanks to the remarkable decrease of the pre-shock Mach number by means of an *extended* pre-compression mechanism, which involves the entire suction surface until the passage shock. Moreover, it was discovered that the *unique incidence* operating condition, at which the optimization was run thanks to

the unique incidence control loop (UICL), introduces a strong tradeoff between the two objectives. An explanation of the flow mechanism accountable for such behavior was given.

On the other hand, the multi-objective shape optimization of the transonic compressor NASA Rotor 67 was carried out by means of a Kriging-based optimizer, ASEMOO, coupled with a CFD solver. The polytropic efficiency and the total pressure ratio were maximized at a specific mass flow rate condition, while the choking mass flow was constrained. A complete and deep reshaping of the rotor geometry was accomplished by means of a generalized parameterization framework that involved sweep and lean deformations, in addition to the modification of six spanwise sections. Although the outcomes were not as incisive as those of the previous optimization due to the limited number of direct CFD evaluations joined with a wide range of decision variables, the optimal designs found gave a snapshot over some of the compelling mechanisms involved in transonic compressors, demonstrating the effectiveness of the entire optimization strategy. The analysis of the Pareto optimal designs corroborates the fact that the optimization process needs more CFD computations to reach high-level results. Unfortunately, computational power was not available in much amount.

As far as future work is concerned, various aspects of the optimization loop may be investigated further. In particular, the optimization engine can be improved by introducing a tuning parameter that allows the optimization engine to switch between GeDEA-II-K and ASEMOO. In fact, when the Kriging gets insight on the function landscape (i.e., this happens after a certain number of direct evaluations depending on the problem complexity), its capabilities are simply superior. Another strong limitation that was encountered during this Ph.D. regards the employment of CFD commercial packages inside the optimization loop. The limitation lays mainly on the required licenses and the software customizability. Comparable open-source software is available but needs a steep learning curve with no guaranteed results. However, efforts in this direction would be useful.

As for the cascades, it would be very appealing to perform a multi-point optimization, not just at the unique incidence. This requires a new control loop on the CFD boundary conditions in order to force the cascade to work at a specific operating point. Finding new families of airfoils with specific behaviors at given design Mach numbers could be the first step for a completely redesign of a 3D transonic compressor rotor based on these new families. Moreover, such

investigation could give further understanding on supersonic flow throughout compressor cascades, simplifying this subject once for all to its lowest terms.

References

[1] J. Durillo et al., "A Study of Multiobjective Metaheuristics When Solving Parameter Scalable Problems," *IEEE Transactions on Evolutionary Computation*, vol. 14, pp. 618–635, 2010.

[2] J. Sacks, W. J. Welch, T.J. Mitchell, and H. P. Wynn, "Design and Analysis of Computer Experiments," *Statistical Science*, vol. 4, pp. 409-435, 1989.

[3] V.V. Toporov, "Modelling and Approximation Strategies in Global Optimization and Mid-Range Metamodels, Response Surface Methods, Genetic Programming, and Low/High Fidelity Models," in *Emerging Methods for Multidisciplinary Optimization*. New York: Blachut J., Eschenauer H.A., 2001, pp. 205-256.

[4] T. Hill and P. Lewicki, *Statistics Methods and Applications*. Tulsa,OK: StatSoft, 2007.

[5] E.P. Box and N.R. Draper, *Empirical Model Building and Response Surfaces*. New York: Wiley, 1998.

[6] A. A. Giunta and L. T. Watson, "A Comparison of Approximation Modeling Techniques: Polynomial Versus Interpolating Models," in *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis & Optimization*, St. Louis, MO, September 2-4, 1998, pp. 392-404.

[7] V.V. Toporov, U. Schramm, A. Sahai, R.D. Jones, and T. Zeguer, "Design Optimization and Stochastic Analysis Based On the Moving Least Squares Method," in *6th World congress of structural and multidisciplinary optimization*, Rio de Janeiro, 2005.

[8] S.W. Cleveland, "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, vol. 74, pp. 829-836, 1979.

[9] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications.*: Chapman & Hall, 1996.

[10] D. Ruppert, S. J. Sheather, and M. P. Wand, "An Effective Bandwidth Selector For Local Least Squares Regression," *Journal of the American Statistical Association*, vol. 90, no. 432, 1995.

[11] D.S. Broomhead and D. Lowe, "Multivariate Functional Interpolation and Adaptive Networks," *Complex Systems*, vol. 2, pp. 321-355, 1988.

[12] M. Hassoun, *Fundamentals of Artificial Neural Networks.*: MIT, 1995.

[13] M. Orr, *Introduction to Radial Basis Function Networks*, available at http://www.anc.ed.ac.uk/rbf/rbf.html#pprs. ed., University of Edinburgh, Ed., 1996.

[14] M. Orr, "Regularisation in the Selection of RBF Centres," *Neural Computation*, vol. 7(3), pp. 606-23, 1995.

[15] S. Chen, E.S. Chng, and Alkadhimi, "Regularized Orthogonal Least Squares Algorithm for Constructing Radial Basis Function Networks," *International Journal of Control*, vol. 64, no. 5, pp. 829-837, 1996.

[16] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[17] M. Orr, "Optimizing the Widths of Radial Basis Functions," *available from http://www.anc.ed.ac.uk/~mjo/rbf.html.*

[18] D.R. Jones, "A Taxonomy of Global Optimization Methods Based On Response Surfaces," *Journal of Global Optimisation*, vol. 21, pp. 345-83, 2001.

[19] N.A.C. Cressie, *Statistics for Spatial Data, Probability and Mathematical Statistics. Revised Ed.*, Wiley, Ed. New York, 1993.

[20] V.R. Joseph, Y. Hung, and A. Sudjianto, "Blind Kriging: A New Method For Developing Metamodels," *ASME Journal of Mechanical Design*, p. 130, 2008.

[21] VR. Joseph, "A Bayesian Approach To the Design and Analysis of Fractional Experiments," *Technometrics*, vol. 48, pp. 219-29, 2006.

[22] A.I.J. Forrester, B. So, and A.J. Keane, "Multi-Fidelity Optimization via Surrogate Modelling," in *Proceedings of the Royal Society*, 2007, pp. 463(2088):3251–69.

[23] A.I.J. Forrester, A.J. Keane, and N.W. Bressloff, "Design and Analysis of 'Noisy' Computer Experiments," *AIAA Journal*, vol. 44(10), pp. 2331–9, 2006.

[24] A.J. Smola and B. Schölkopf, "A Tutorial On Support Vector Regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199-222, 2004.

[25] A.L.J. Forrester and A.J. Keane, "Recent Advances in Surrogate-Based Optimization," *Progress in Aerospace Science*, vol. 45, pp. 59-79, 2009.

[26] R. Bellmann, *Adaptive Control Processes: A Guided Tour*.: Princeton University Press, 1961.

[27] W.K. Chen, *Linear Networks and Systems*. Belmont, CA: Wadsworth, 1993, pp. 123-135.

[28] R. Meyers and D. Montgomery, *Response Surface Methodology*. 1995: John Wiley&Sons, Inc., New York.

[29] R.T. Haftka, E.P. Scott, and J.R. Cruz, "Optimization and Experiments: a Survey," *Applied Mechanic Review*, vol. 51, no. 7, pp. 435-448, 1998.

[30] L. Breiman, "Bagging Predictors," *Machine Learning*, vol. 24, pp. 123-140, 1996.

[31] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*.: Chapman and Hall, 1993.

[32] A.I.J. Forrester, A. Sòbester, and A.J. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*. Chichester, UK: John Wiley and Sons Ltd., 2008.

[33] A.I.J. Forrester, "Global Optimization of Deceptive Functions With Sparse Sampling," in *12th AIAA/ISSSMO Multidisciplinary Analysis and Optimization Conference*, Victoria, British Columbia Canada, 10-12 September 2008.

[34] S. Koziel, W.J. Bandler, and K. Madsen, "Quality Assessment of Coarse Models and Surrogates for Space Mapping Optimization," *Optimization and Engineering*, vol. 9, no. 4, pp. 375-391, 2008.

[35] J.W. Bandler et al., "Space Mapping: The State of the Art," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 1, pp. 337-361, 2004.

[36] Y. Mack, T. Goel, W. Shyy, T.R. Haftka, and V.N. Queipo, "Multiple Surrogates for the Shape Optimization of Bluff Body- Facilitated Mixing," in *43rd AIAA Aerospace Sciences Meeting and Exhibit, AIAA Paper*, 2005, pp. 2005-0333.

[37] F.A.C. Viana and R.T. Haftka, "Using Multiple Surrogates for Metamodeling," in *7th ASMO-UK/ISSMO Internationa Conference on Engineering Design Optimization, International Society for Structural and Multidisciplinary Optimization*, Bath, England, U.K., 2008.

[38] P. Stoica and Y. Selen, "Model-Order Selection: A Review of Information Criterion Rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36-47, 2004.

[39] Y. Yang, "Regression with Multiple Candidate Models: Selecting or Mixing?," *Statistica Sinica*, vol. 13, no. 387, pp. 783-809, 2003.

[40] F.A.C. Viana, R.T. Haftka, and V.Jr. Steffen, "Multiple Surrogates: How Cross-Validation Errors Can Help Us to Obtain the Best Predictor," *Structural and Multidisciplinary Optimization*, vol. 39, no. 4, pp. 439-457, 2009.

[41] A. Samad, K. Kim, T. Goel, T.R. Haftka, and W. Shyy, "Multiple Surrogate Modeling for Axial Compressor Blade Shape Optimization," *Journal of Propulsion and Power*, vol. 24, no. 2, pp. 302-310, 2008.

[42] R. Jin, W. Chen, and T.W. Simpson, "Comparative Studies of Metamodeling Techniques Under Multiple Modeling Criteria," *Structural and Multidisciplinary Optimization*, vol. 22, no. 3, pp. 1-13, 2001.

[43] E. Benini and R. Ponza, "Nonparametric Fitting of Aerodynamic Data Using Smoothing Thin-PLate Splines," *AIAA Journal*, vol. 48, no. 7, pp. 1403-1419, July 2010.

[44] A.J. Kean, "Comparios of Several Optimization Strategies for Robust Turbine Blade Design," *Journal of Propulsion and Power*, vol. 25, no. 5, 2009.

[45] J. Peter and M. Marcelet, "Comparison of Surrogates Models for Turbomachinery Design," *Wseas Trans. On Fluid Mechanic*, vol. 3, no. 1, 2008.

[46] E. Benini, "Three-Dimensional Multi-Objective Design Optimization of a Transonic Compressor Rotor," *Journal of Propulsion and Power*, vol. 20, no. 3, pp. 559-565, May-June 2004.

[47] S. Burguburu and A. Pape, "Improved Aerodynamic Design of Turbomachinery Bladings by Numerical Optimization," *Aerospace Science and Technology*, vol. 7, no. 4, pp. 277-287, June 2003.

[48] C.M. Jang, P. Li, and K.Y. Kim, "Optimization of Blade Sweep in a Transonic Axial Compressor Rotor," *JSME International Journal, Series B*, vol. 48, no. 4, 2005.

[49] C.S. Ahn and K.Y. Kim, "Aerodynamic Design Optimization of a Compressor Rotor with Navier-Stokes Analysis," *Proc. Instn Mech. Engrs, Part A - J. Power and Energy*, vol. 217, no. 2, pp. 179-184, 2003.

[50] N. Chen, H. Zhang, Y. Xu, and W. Huang, "Blade Parameterization and Aerodynamic Design Optimization for a 3D Transonic Compressor Roto," *Journal of Thermal Science*, vol. 16, no. 2, March 2007.

[51] X.D. Wang, C. Hirsch, Sh. Kang, and C. Lacor, "Multi-Objective Optimization of Turbomachinery Using Improved NSGA-II and Approximation Model," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 9-12, pp. 883-895, 2011.

[52] A. Toffolo and E. Benini, "Genetic Diversity As An Objective In Multi-Objective Evolutionary Algorithms," *Evolutionary Computation*, vol. 11, no. 2, pp. 151–157, 2002.

[53] M.K. Karakasis, G.D. Koubogiannis, and C.K. Giannakoglou, "Hierarchical Distributed Metamodel-Assisted Evolutionary Algorithms in Shape Optimization," *Int. J. Numer. Meth. Fluids*, pp. 455-469, 2007.

[54] O. Oksuz and I.S. Akmandor, "Multi-Objective Aerodynamic Optimization of Axial Turbine Blades Using a Novel Multilevel Genetic Algorithm," *Journal of Turbomachinery*, vol. 132, October 2010.

[55] A. Keskin, P. M. Flassig, A. K. Dutta, and D. Bestle, "Accelerated Industrial Blade Design Based on Multi-Objective Optimization Using Surrogate Model Methodology," in *ASME Turbo Expo 2008: Power for Land, Sea, and Air*, Grmany, Berlin, June 9-13, 2008.

[56] K.C. Giannakoglou, D.I. Papadimitriou, and I.C. Kampolis, "Aerodynamic Shape Design Using Evolutionary Algorithms and New Gradient-Assisted Metamodels," *Comput. Methods Appl. Mech. Engrg*, vol. 195, 2006.

[57] M. Gendreau and Jean-Yves Potvin (Editors), *Handbook of Metaheuristics*, Second Edition ed., Springer Science+Business Media, Ed., LLC 2003, 2010, vol. 2010.

[58] I. Loshchilov, M. Schoenauer, and M. Sebag, "Comparison-Based Optimizers Need Comparison-Based Surrogates," in *Parallel Problem Solving from Nature (PPSN XI)*, Krakov,Poland, 2010, pp. 364–1373.

[59] C. Comis Da Ronco and E. Benini, "A Simplex Crossover Based Evolutionary Algorithm Including the Genetic Diversity as Objective," *Applied Soft Computing*, vol. 13, pp. 2104–2123, 2012.

[60] D.H. Wolpert and W.G. Macready, "No Free Lunch Theorems For Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67-82, 1997.

[61] D.R. Jones, M. Schonlau, and W.J. Welch, "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, vol. 13, no. 4, pp. 455-492, 1998.

[62] Jeong, S., M. Murayama, and K. Yamamoto, "Efficient Optimization Design Method Using Kriging Mode," *Journal of Aircraft*, vol. 42, no. 2, pp. 413-420, 2005.

[63] O. Roustant, D. Ginsbourger, and Y. Deville, "DiceKriging, DiceOptim: Two R Packages For the Analysis of Computer Experiments by Kriging-Based metamodeling and optimization," 2012.

[64] Ł. Łaniewski-Wołłk, "Expected Hypervolume Improvement – Criterion for Kriging Based Multiobjective Optimization," in *19th Polish National Fluid Dynamic Conference*, POZNAŃ, (available at http://kkmp2010.put.poznan.pl/registered/0608201026), 05-09 Sept 2010.

[65] K. Deb, "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II," *Evolutionary Computation, IEEE Transactions*, vol. 6, no. 2, pp. 182-197, 2002.

[66] G.E. Goldberg, *Genetic Algorithm in Search, Optimization, and Machine Learning*.: Addison-Wesley Professional, 1989.

[67] K. Deb, ""Multi-Objective Genetic Algorithms: Problem Difficulties and Construction of Test Problems," *Evolutionary Computation*, vol. 7, no. 3, pp. 205-230, 2000.

[68] E. Zitler, K. Ded, and L. Thiele, "Comparison of Multiobjective Evolutionary Algorithms: Empirical Results," *Evolutionary Computation*, vol. 8, no. 2, pp. 173-195, 2000.

[69] L. While, P. Hingston, L. Barone, and S. Huband, "A Faster Algorithm for Calculating Hypervolume," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 1, Febbruary 2006.

[70] A. Kantrowitz, "The Supersonic Axial Flow Compressor," NACA, Report 974 1950.

[71] H. Starken and H.J. Lichtfuss, "Supersonic Cascade Performance," AGARD LS 3970 1970.

[72] H.A. Shapiro, *The Dynamic and Thermodynamic of Compressible Fliud Flow*, 1st ed. New York: The Ronald Press Corp, 1953, vol. Vol I.

[73] S.B. Stratford and E.G. Sansome, "Thory and Tunnel Test of Rotor Blades for Supersonic Turbines," ARC R&M No. 3275 1962.

[74] W.W. Wilcox, "Investigation of Impulse-Type Supersonic Compressor with Hub-Tip Ratio of 0.6 and Turning to Axial Direction. II-Stage Performance with Three Different Sets of Stators," NACA, RM E 55 F 28 1955.

[75] J.M. Hartmann and R.E. Tysl, "Investigation of a Supersonic –Compressor Rotor with Turning to Axial Direction. II-Rotor Component Off-Design and Stage Performance," NACA, RM E 53L24 1953.

[76] F.J. Klapproth and G.N. Uliman, "Performance of an Impulse-Type Supersonic Compressor with Stators," NACA, RM E 52 B 22 1952.

[77] G.E. Johnson, E. von Ohain, O.M. Lawson, and R.K. Cramer, "A Blunt Trailing Edge Supersonic Compressor Blading," WADC TN 59-269 1959.

[78] S. Farokhi, "Aircraft Propulsion," in *Aircraft Propulsion*.: John Wiley & Sons, Inc., 2009, ch. 5, pp. 246-253.

[79] H.J. Lichtfuss and H. Starken, "Supersonic Cascade Flow," 1974.

[80] R.D. Seyler and H.L. Smith, "Single Stage Experimental Evaluation of High Mach Number Compressor Rotor Blading, Part 1, Design of Rotor Blading," NASA, CR-54581 1967.

[81] P.J. Gostelov, W.K. Krabacher, and H. L. Smith, "Performance Comparison of the High Mach Number Compressor Rotor Blading," NASA, CR-1256 1968.

[82] A.A. Mikolajczak, L. Morris A., and B.V. Johnson, "Comparison of Performance of Supersonic Blading in Cascade and in Compressor Rotors," *Trans. ASME, Journal of Engineering for Power*, vol. 1, no. Series A, pp. 42-48, 1971.

[83] C.J. Emery, R.J. Dunavant, and R.W. Westphal, "Experimental Investigation of a Transonic Compressor Cascade and Test Results for Four Blade Suction Sections," NASA, TM X 197 1960.

[84] M.R. Hearsey and J.A. Wennerström, "Axial Compressor Airfoils for Supersonic-Mach Numbers," ARL 70-0046 1970.

[85] K. Oswatitsch, "Pressure Recovery for Missiles with Reaction Propulsion at High Supersonic Speeds," NACA, TM 1140 1947.

[86] R.J.W. Creagh and F.J. Klapproth, "Utilization of External-Compression Diffusion Principle in Design of Shock-In-Rotor Supersonic compressor Blading," NACA, RM E 53F18 1953.

[87] J.J. Lawrence and H.J. Melvin, "Investigation of Supersonic Compressor Rotors Designed with External Compression," NACA, RM E 54G27a 1954.

[88] L.A. Morris, E.J. Halle, and E. Kennedy, "High-Loading, 1800 ft/sec Tip Speed Transonic Compressor Fan Stage. Aerodynamic and Mechanical Design," NASA, CR-120907 1972.

[89] H.A. Schreiber, "Experimental Investigation on Shock Losses of Transonic and Supersonic Compressor Cascades," DFVLR AD-P0005 513, 1986.

[90] H.A. Schreiber, "Shock-Wave Turbulent Boundary Layer Interaction in a Highly Loaded Transonic Fan Blade Cascade," *AGARD*, vol. CP571, pp. 17-1, 1996.

[91] T. Sonoda, T. Arima, M. Olhofer, and B. Sendhoff, "A New Concept of a Two-Dimensional Supersonic Relative Inlet Mach Number Compressor Cascade," in *Proceedings of ASME Turbo Expo 2009: Power for ALnd, Sea and Air*, Orlando,Florida,USA, 2009.

[92] H.A. Schreiber and H. Starken, "An Investigation of a Strong Shock-Wave Turbulent Boundary Layer Interaction in a Supersonic Compressor Cascade," *ASME Journal of Turbomachinery*, vol. 114, pp. 494-503, 1992.

[93] H.A. Schreiber, H. Starken, and W. Steinert, "Transonic and Supersonic Cascades," North Atlantic Trety Organization, AGARD AG-328 1993.

[94] A. Bölsc and P. Suter, *Transsonische Turbomachinen*. Kerlsruhe: Braun, 1986.

[95] P. Levine, "Two-Dimensional Inflow Conditions for a Supersonic Compressor with Curved Blades," *ASNE Journal of Applied Mechanics*, vol. 24, no. 2, pp. 165-169, 1957.

[96] O. Novak, "Flow in the Entrance Region of a Supersonic Cascade," in *Strojnicky Casposis XIX*, Slovenskej Akademic vied Bratislava, 1967, pp. 138-150.

[97] R.E. York and H. S. Woodard, "Supersonic Compressor Cascades-An Analysis of the Entrance Region Flow Field Containing Detached Shock Waves," *ASME Journal of Engineering for Power*, vol. 98, pp. 247-257, 1976.

[98] D.L. Tweedt, H-A. Schreiber, and H. Starken, "Experimental Investigation of the Performance of a Supersonic Compressor Cascade," *Transaction of ASME,* vol. 110, 1998.

[99] H. Starken, Y. Zhong, and H.A. Schreiber, "Mass Flow Limitation of Supersonic Blade Rows Due to Leading Edge Blockage," in *ASME Paper*, Amsterdam, 1984, pp. 84-GT-233.

[100] H.A. Schreiber and H. Starken, "Experimental Cascade Analysis of a Transonic Compressor Rotor Blade Section," *Journal of Engineering for Gas Turbines and Power*, vol. 106, no. 2, pp. 288-294, 1984.

[101] B. Küsters and H.A. Schreiber, "Compressor Cascade Flow with Strong Shock-Wave/Boundary Layer Interaction," *AIAA Journal*, vol. 36, no. 11, pp. 2072-2078, 1998.

[102] Inc. Ansys, "ICEM User's Guide," release 14.1 2013.

[103] Inc. Ansys, "FLUENT User's Guide," release 14.1 2012.

[104] R.F. Menter, "Two-Equation Eddy-Viscosity Turbulence Models for Engineering Applications," *AIAA Journal*, vol. 32, no. 8, pp. 1598–1605, 1994.

[105] R.P. Spalart and R.S. Allmaras, "A One-Equation Turbulence Model for Aerodynamic Flows," in *30th Aerospace Sciences Meeting & Exhibit*, Reno,NV, January January 6-9,1992.

[106] E.M. Mortenson, *Geometric Modeling*, Second edition ed., John Wiley & Sons, Ed., 1997.

[107] Inc. The MathWorks, "MATLAB," Natick, MA, release 2014a.

[108] A. J. Strazisar, J. R. Wood, M. D. Hathaway, and K. L. Suder, "Laser Anemometer Measurements in a Transonic Axial-Flow Fan Rotor," NASA, Technical Paper 2879 November 1989.

[109] A. Oyama, M.S. Liou, and S. Obayashi, "Transonic Axial-Flow Blade Shape Optimization Using Evolutionary Algorithm and Three-Dimensional Navier-Stokes Solver," *Journal of Propulsion and Power*, vol. 20, pp. 612-619, 2004.

[110] S. Pierret, "Multi-Objective and Multi-Disciplinary Optimization of Three-Dimensional Turbomachinery Blades," in *6th World Congress of Structural and Multidisciplinary Optimization*, Rio de Janeiro,Brazil, 30 May-03 June,2005.

[111] Y. Lian and M.S. Liou, "Multiobjective Optimization Using Coupled Response Surface Model and Evolutionary Allgorithm," *AIAA Journal*, vol. 43, no. 6, pp. 1316-1325, June 2005.

[112] Y. Lian and N.H. Kim, "Reliability-Based Design Optimization of a Transonic Compressor," *AIAA Journal*, vol. 44, no. 2, pp. 368-375, Febbruary 2006.

[113] H. Okui, T. Verstraete, R.A. Van De Braembussche, and Z. Alsalihi, "Three-Dimensional Design and Optmization of a Transonic Rotor in Axial Flow Compressors," *Journal of Turbomachinery*, vol. 135, May 2013.

[114] J. Luo, C. Zhou, and F. Liu, "Multipoint Design Optimization of a Transonic Compressor Blade by Using an Adjoint Method," *Journal of Turbomachinery*, vol. 13, May 2014.

[115] P. Hu, B. Choo, M. Zangheneh, and M. Rahmati, "On Design of Transonic Fan by 3D Inverse Design Method," in *Proceedings of GTS2006, ASME Turbo Expo 2006: Power for Land, Sea and Air*, Barcelona, Spain, May 8-11, 2006.

[116] W.T. Tiow and M. Zanageneh, "Application Of A Three-Dimensional Viscous Transonic Inverse Method To Nasa Rotor 67," *Proc Instn Engrs, Part A: Journal of Power and Propulsion*, vol. 216, pp. 243-255, 2002.

[117] H. Watanabe and M. Zangeneh, "Design of The Blade Geometry of Swept Transonic Fans by 3D Inverse Design," in *Proceedings of ASME Turbo Expo 2003 Power for Land, Sea, and Air*, Atlanta, Georgia, USA, 2003.

[118] C.H. Law and A.R. Wadia, "Low Aspect Ratio Transonic Rotors: Part2: Influence of Location of Maximum Thickness on Transoinc Compressor Performance," *ASME Journal of Turbomachinery*, vol. 115, pp. 226-239, 1993.

[119] C.A. Hah and Wennerstrom A.J., "Three-Dimensional Flow Field Inside a Transonic Compressor with Swept Blades," *ASME Journal of Turbomachinery*, vol. 113, pp. 241-251, 1991.

[120] A.R. Wadia, P.N. Szucs, and D.W. Crall, "Inner Workings of Aerodynamic Sweep," *ASME Journal of Turbomachinery*, vol. 120, pp. 671-682, 1998.

[121] C.A. Hah, S.L. Puterbaugh, and A.R. Wadia, "Control of Shock Structure and Secondary-Flow Field Inside Transonic Compressor Rotors Through Aerodynamic Sweep," in *ASME Paper 98-GT-561*, 1998.

[122] J.D. Denton and L. Xu, "The Effects of Lean and Sweep on Transonic Fan Performance," in *ASME Paper GT-2002-30327*, 2002.

[123] T. Sasaki and F. Breugelmans, "Comparison of Sweep and Dihedral Effects on Compressor Cascade Performance," *Transaction of the ASME*, vol. 120, pp. 454-463, July 1998.

[124] C.A. Hah, "Numeric Modeling of Endwall and Tip-Clearance Flow of an Isolated Compressor Rotor," *Jounral Eng. Gas Turbines Power*, vol. 108, no. 1, January 1986.

[125] E. Benini and R. Billo, "Aerodynamic Of Swept And Leaned Transonic Compressor-Rotors," *Applied Energy*, vol. 84, pp. 1012-1027, 2007.

[126] G. Abate, "Aerodynamic Optimization of a Transonic Axial Compressor Rotor," Università degli Studi di Padova, Padova, Tesi Magistrale in Ingegneria Aerospaziale 2012.

Appendix A

A.1 Optimization Framework

The automatic optimization framework is in charge to conduce the optimization process in an automatic and reliable way, preventing any kind of disturbance, *internal* (i.e., errors from programs inside the loop) and *external (*i.e., issues coming from the external environment, as shutdowns and so on), by means of error expectation handling. The entire framework was developed in MATLAB®, Bash scripting and C code.

In the scope of this research, the general conceptual strategy con be summarized in the flowchart reported in Figure A.1.

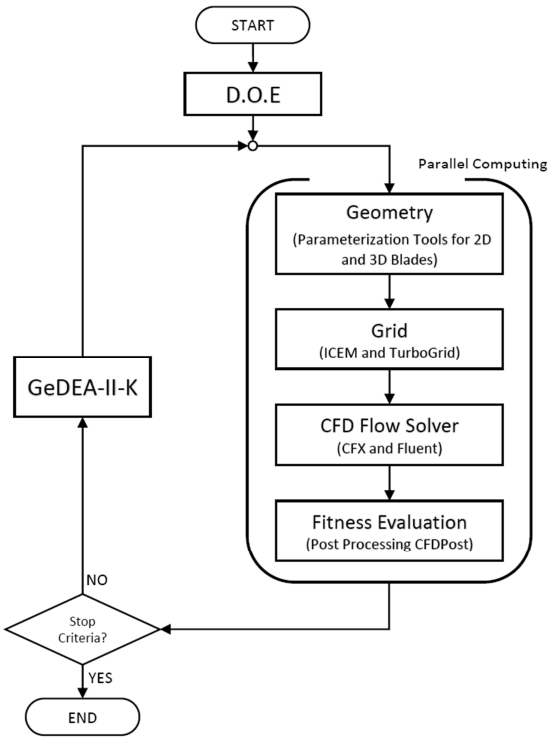


Figure A.1: Optimization conceptual strategy.

It comprise a starting procedure as the D.O.E., an optimization engine (i.e., the optimization algorithm, for instance the GeDEA-II-K), and the evaluation of the design fitness in a parallel environment.

The realization of such conceptual strategy in a Linux cluster is reported in the flow chart of Figure A.2. Each of the three levels is independent from each other, that is the crush of a level does not have any implication on the operation of the others. In particular, the features of the levels are as follow:

1. Level 0: refers to the optimization algorithm, which outputs new designs to be evaluated ("Population") and waits for their response ("Fitness").
2. Level 1: the "Exchange pool" is the link between the optimizer and the hard computing environment. It represents the locus where requests of the optimization algorithm are converted to design evaluation procedures, in order to calculate design fitness and, in turn, satisfy the algorithm's request.
3. Level 2: calculations are performed in parallel on a cluster machine, or in several clusters; there is no limitation from this point of view and we can take advantage of all the available computational power.

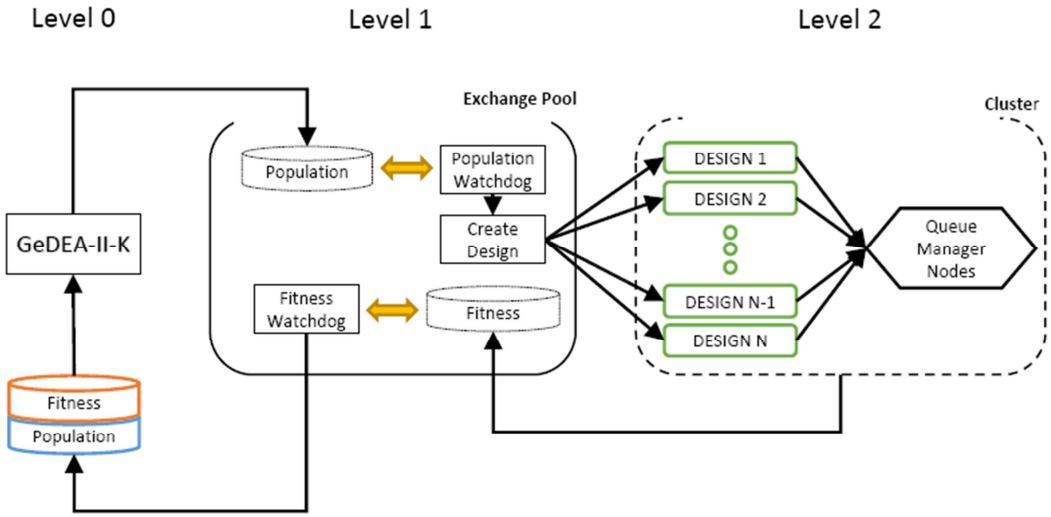


Figure A.2: Optimization framework layout.

In Figure A.3 is shown how design fitness evaluation is accomplished. The black-box named "watchdog" is a *looping process* that is linked (yellow arrow) to

a particular “critical process” that needs to be looked after; the watchdog endures until the critical process terminates. The aim of a watchdog is to control the evolution of the critical process and handle any kind of error by performing an appropriate action. Once each critical process has its own tailored watchdog, the optimization process can be started. Such approach is inherently general and can be adapted to any kind of program, commercial or free, involved in the optimization loop.

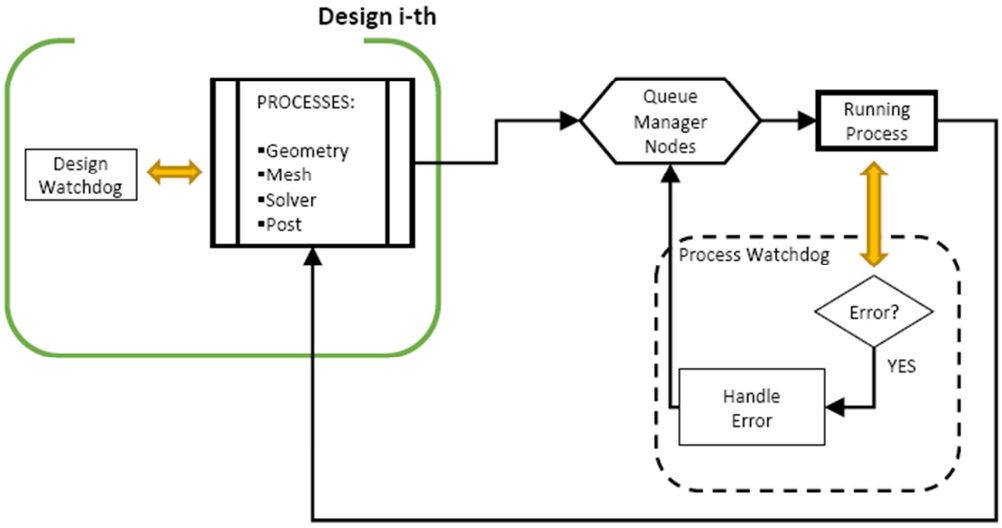


Figure A.3: Flowchart Detail of a single design.

Acknowledgements

Al mio supervisore, Prof. Ernesto Benini, un ringraziamento particolare per la passione che ha saputo trasmettermi negli anni, ancor prima di iniziare il percorso di Dottorato. Senza il Suo sostegno, sia professionale che umano, oggi non mi troverei a scrivere questa Tesi.

I would like to thank my Poland friend Łukasz Łaniewski-Wołłk, the only mathematician I now so far, for his grateful support during his stay here in Padova. Even now, we both have clear in mind how painful this transonic compressor has been for us!

Thanks to Tayeb Boulkera for introducing me to the field of high-speed propellers. The time spent working in team with you and Łukasz was simply amazing

Parte di questo lavoro non sarebbe stato possibile senza la disponibilità dell'Ing. Claudio Comis Da Ronco, che mi ha concesso di metter mano al suo prezioso algoritmo, il GeDEA-II.

Un pensiero particolare va ai colleghi del IV piano che hanno saputo rallegrare momenti di amara sconsolazione dettati dalla Divergenza CFD, in ogni sua forma. Un grazie di cuore a Eleonora, Stefano, Gabriele, Andrea, lo zio Enrico, Luca e Matteo (detto il Matteone) e, ovviamente, a "El can de Trieste".

Ultimi, ma no per importanza, ringrazio i miei colleghi di ufficio Davide e Giangi per tutto quello che hanno fatto, che stanno facendo, e che faranno per me (ancora non lo sanno!), rendendo indimenticabile l'ultimo anno di Dottorato.

Giovanni Venturelli
Padova, 1.2.2015

Table represents all transactions in progress in every domain, sending a graphic command to the scripting subsystem, all running tasks in system category, system spawned a process.