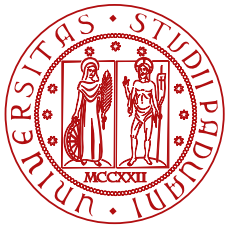
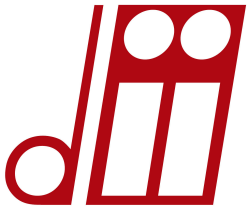


Università
degli Studi
di Padova



Master Thesis in Mechanical Engineering

Reduced order models for small-scale turbopumps

Master Candidate

Marco Miolo

Student ID 2054695

Supervisor

Prof. Ernesto Benini

University of Padova

Co-supervisor

Prof. Jens Honoré Walther

Technical University of Denmark

Academic Year
2023/2024

Preface

This thesis contains the work done over five months at the Laboratory of Applied Mechanical Design of the École Polytechnique Fédérale de Lausanne (EPFL), in partial fulfillment for the Master of Science on Mechanical Engineering at the University of Padova. The research presented in this thesis was carried out from January 2024 to June 2024.

The project was carried out under the primary supervision of Professor Jürg Alexander Schiffmann and Postdoc Sajjad Zakeralhoseini from the Laboratory of Applied Mechanical Design of EPFL, Rue de la Maladière 71b, Neuchâtel, Switzerland. And with the supervision of Professor Ernesto Benini at University of Padua. And the co-supervision of Professor Jens Honoré Walther.

This thesis explores the application of Proper Orthogonal Decomposition (POD) in constructing reduced order models for turbopumps, particularly focusing on geometrical variations. The research aims to enhance the design process by offering a computationally efficient alternative to traditional full-order models.

Marco Miolo - 2054695

.....

Signature

8th June 2024

Abstract

In the field of turbomachinery, flow field and performance analyses are typically conducted using full-order models like Computational Fluid Dynamics (CFD). While these methods provide high-quality results, they demand significant time and computational resources. This study aims to develop faster methods by inferring Reduced Order Models (ROMs) that maintain high accuracy.

Principal component analysis techniques, such as Proper Orthogonal Decomposition (POD), have been used in fluid mechanics to identify flow structures and develop ROMs for various classical problems. Despite extensive research in turbulence and aerodynamic applications, no prior work has focused on creating a ROM capable of predicting the performance impact of design changes in centrifugal turbomachinery.

This research bridges this gap by integrating POD with commercial CFD software to extract valuable information and demonstrate the model's predictive capabilities. Specifically, the study focuses on a turbopump designed for an 8 kW Organic Rankine Cycle (ORC). The results show that the developed ROM can accurately predict the performance of the turbopump, achieving significant reductions in computational time compared to traditional CFD methods. Additionally, the study identifies and discusses the limitations of using POD in this context.

By offering a faster, computationally efficient alternative to full-order CFD simulations, this work contributes to the optimization of turbomachinery design, with potential applications in enhancing the efficiency of ORC systems.

Acknowledgements

Marco Miolo, MSc Mechanical Engineering, DTU
Thesis author

Ernesto Benini, Professor, Industrial Engineering Department at University of Padova
Thesis co-supervisor

Jens Honoré Walther, Professor, Department of Civil and Mechanical Engineering Fluid Mechanics, Coastal and Maritime Engineering at DTU
Thesis supervisor

Jürg Alexander Schiffmann, Associate Professor, Laboratory of Applied Mechanical Design at EPFL
Thesis co-supervisor

Sajjad Zakeralhoseini, Postdoc, Laboratory of Applied Mechanical Design at EPFL
Thesis co-supervisor

Abbreviations

CFD Computational Fluid Dynamics
DES Detached Eddy Simulations
DNS Direct Numerical Simulations
DOF Degrees Of Freedom
FOM Full Order Model
LES Large Eddy Simulations
ORC Organic Rankine cycle
OSF Optimal-Space Filling
PCA Principal Component Analysis
POD Proper Orthogonal Decomposition
PSO Particle Swarm Optimization
RANS Reynolds-averaged Navier–Stokes
RBF Radial Basis Function
ROM Reduced Order Model
TKE Total Fluctuating Kinetic Energy
URANS Unsteady Reynolds-averaged Navier–Stokes

Nomenclature

β_{out}	Blade angle, deg
ψ_{sp}	Pitch of splitter blades
θ_{mb}	Angle between adjacent blades, deg
θ_{sp}	Angle between splitter blade and pressure side of main blade, deg
m'_{le}	Relative meridional location of main blades
n	Rotational speed, krpm
RE_v	Relative error for the velocity field
S_N	Normalized span
Z	Number of blades
$m'_{le,sp}$	Relative meridional location of splitter blades
b	Blade height, m
c	Clearance gap, m
pt	Point

Contents

Preface . . . iii
Abstract . . . iv
Acknowledgements . . . v
Abbreviations . . . vi

1 Introduction **1**
1.1 Context . . . 1
1.2 Objectives . . . 2
1.3 State of the art . . . 3
1.3.1 Reduced Order Models . . . 3
1.3.2 Proper Orthogonal Decomposition . . . 4
1.4 Organization of the thesis . . . 5

2 Methods **7**
2.1 Turbopump for Organic Rankine Cycle . . . 7
2.2 Proper Orthogonal Decomposition . . . 9
2.2.1 2-D case . . . 10
2.2.2 n-dimensional case . . . 13
2.2.3 Energy considerations . . . 14
2.2.4 Snapshot POD . . . 14
2.3 CFD results extraction . . . 15
2.4 Mesh interpolation . . . 16
2.4.1 Span . . . 16
2.4.2 Inlet and outlet . . . 19
2.4.3 Blades . . . 21
2.4.4 Data format . . . 23
2.5 POD reconstruction . . . 23
2.6 POD model construction . . . 26
2.7 Pump performance evaluation through POD . . . 27

3 Results **31**
3.1 Surface flow field POD . . . 31
3.1.1 Blade to blade surface . . . 32
Error analysis . . . 35
3.1.2 Inlet and outlet surfaces . . . 36
Error analysis . . . 38
3.1.3 Main and splitter blade surfaces . . . 39
Error analysis . . . 41

3.2 Modal visualization 43
3.3 First POD model for pump performance prediction 47
3.4 POD model tuning three parameters 51
3.4.1 Four blades turbopump 51
3.4.2 Five blades turbopump 58
3.4.3 Six blades turbopump 62

4 Discussion **67**
4.1 Mesh interpolation and data extraction 67
4.2 POD model construction 69
4.3 POD reconstruction 70
4.4 POD capabilities 70

5 Conclusions **71**

A Extra plots **77**
A.1 Surface mesh interpolation 77
A.2 Blade to blade plane initial POD 79
A.3 Outlet and inlet initial POD 80
A.4 Main and splitter blades 83
A.5 Modal visualization 84
A.6 Four blades turbopump 86
A.7 Five blades turbopump 88
A.8 Six blades turbopump 93

Chapter 1

Introduction

1.1 Context

The transportation sector accounts for around one-fifth of the global carbon dioxide (CO_2) emissions. This can be translated into the production of 7.9 billion tonnes of CO_2 [1]. Emerging technologies such as electrification and hydrogen fuel are expected to decarbonize this sector within the coming decades. However, several challenges, including infrastructure development and the electrification of certain ground vehicles, such as trucks, must be addressed before widespread adoption. These challenges include the need for infrastructure development and the complexities associated with electrifying certain ground vehicles [2]. Facing these challenges with the goal of decreasing the impact of the transportation sector, a validated and viable existing technology such as Organic Rankine Cycle (ORC) can be used. Particularly, ORCs can recover energy from high-temperature exhaust gases, which waste 40-50 % of the fuel energy [3]. This trend opens up potential markets, including applications in long-haul trucks and internal combustion engines, as highlighted by Shu et al. [4].

The Organic Rankine Cycle distinguishes itself from the conventional Rankine cycle by employing a working fluid different from water. The employed working fluids are characterized by high molecular weights and low boiling pressure/temperature compared to water. The utilization of a different working fluid enables ORC applications at lower temperatures, typically less than 500°C. Evaporators and condensers are well-established in both technological and manufacturing aspects. Ongoing studies focus on their performance in supercritical regions and the development of compact heat exchangers. The most technically advanced component of an ORC is the expander. Turbines and expanders have been studied extensively. Pumps, although traditionally adapted from other applications, are now undergoing investigation for novel types suitable for micro- and small-scale ORC applications.

Poor pump performance can deeply affect the performance of ORC systems, due to the reduction in thermal efficiency and underlined by Meng et al. [5] by investigating an industrial multistage centrifugal pump for a specific ORC system for waste heat recovery. The efficiency of the pump is reported to vary between 15 % and 65.7 %, suggesting that pump design should be considered in the design process of an ORC system. The challenges presented by multistage centrifugal pumps are reducing their weight and size, in order for them to fit in an engine and to affect the least its performance. This could be done by designing high-speed single-stage turbopumps in ORC systems allowing for a compact design of the energy recovery systems. The design

and behaviour of such complicated devices have been investigated by Zakeralhoseini through the design, computational analysis, and experimental investigations of a high-speed small-scale turbopump for mobile waste heat recovery applications based on Organic Rankine Cycles [6]. Moreover, Zakeralhoseini investigated the influence of tip leakage, splitter blades and meridional profiles on the performance of these machines [7, 8].

The design process for these pumps typically begins with empirical correlations derived from efficiency considerations and a deep technical understanding of the main design challenges. Engineers start by analyzing requirements and proposing solutions expected to deliver optimal performance. These empirical correlations are refined and strengthened using experimental data, particularly from existing machines analyzed applying similarity laws. However, in the case of small-scale turbopumps, the accuracy of similarity laws diminishes due to the extreme working conditions. To achieve an optimal design, advanced optimization procedures are necessary, which bear the risk of being time-consuming. Optimization plays a crucial role in maximizing energy output and reducing emissions, making it an area of ongoing research. Many strategies rely heavily on costly Computational Fluid Dynamics (CFD) simulations to obtain reliable results across various geometrical and operational conditions. Substantial effort has been dedicated to developing software capable of efficiently simulating turbomachinery.

Given the vast amount of data available in every field, data analysis and mining have become integral to research efforts. Strategies developed in this context have shown promise in the fluid mechanics domain, particularly in constructing Reduced Order Models (ROMs). These ROMs can provide relatively accurate results with minimal computational resources. This could allow to investigate new design spaces with less computational demand [9].

Given the significance of ORC in improving energy efficiency, this study focuses on the design and optimization of turbopumps, a critical component of ORC systems. The innovative aspect of this work lies in using Proper Orthogonal Decomposition (POD) to create reduced order models for predicting the performance impact of design changes in centrifugal turbomachinery. This approach aims to significantly reduce computational time while maintaining high accuracy, offering a more efficient alternative to traditional CFD simulations.

1.2 Objectives

The primary objectives of this thesis are:

- To develop a reduced order model using POD for turbopump design and optimization.
- To integrate POD with commercial software such as `ANSYS` for efficient results extraction and data adaptation.
- To evaluate the performance of the reduced order model in predicting the impact of design changes on turbopump performance.
- To optimize the turbopump design by tuning specific parameters and assessing the improvements in energy efficiency and performance.
- To assess the limitations and accuracy of the POD-based reduced order model in comparison to traditional CFD simulations.

The questions arise: can POD be leveraged to optimize turbomachinery? Could data mining offer a faster, more efficient approach to either replace or complement traditional CFD full-order model (FOM) simulations?

1.3 State of the art

In recent years, dramatic increases in computing power and advances in CFD methods have significantly enhanced the complexity and fidelity of computational models for turbomachinery performance prediction. Today, the compressible and fully viscous Reynolds-averaged Navier–Stokes (RANS) equations are a well established technique in CFD modelling. The only simplification being the time averaging of turbulence quantities using one to six equation models. A more detailed approach is the Unsteady Reynolds-averaged Navier-Stokes (URANS) model, which accounts for temporal variations in turbulence. Thanks to the enhancement of computational power of the working stations, this technique is now frequently adopted to model unsteady 3D flows in turbomachinery, with multiple million node meshes. More accurate and computationally expensive techniques have been developed, such as: Large Eddy Simulations (LES) or Detached Eddy Simulations (DES), where the larger eddies are directly modelled, and the smaller treated with RANS. While the most accurate and computationally demanding CFD modelling is done through Direct Numerical Simulation (DNS). The 'less accurate' models, RANS, are routinely used for design aid and evaluation. An example is provided by Zakeralhoseini in [6]. 'More accurate' models, LES, are used to deeply investigate losses such as it was done in centrifugal pumps from Sorguven et al. in [10], but could not be applied routinely given the required computational time and the challenges that could be faced when setting up an LES simulation.

In order to take advantage of the accuracy in the results and strong descriptiveness of the flow field offered by LES, DES and DNS, machine learning techniques, such as neural networks, have been integrated into CFD solvers to accelerate computations and enhance prediction accuracy, as detailed by Hammond et al. [11]. This trend is also retrieved in commercial software offers such as `ANSYS`. In the latest version, the software offer includes `SimAI`, a cloud-enabled generative AI platform which uses user's simulation results to reliably assess the performance of a new design within minutes. In the work from Hammond et al. most of the offered solution can be referred to as intrusive techniques, requiring the manipulation of the original CFD solver, which may not be available or desirable.

CFD results serve various applications, including the development of ROMs through non-intrusive data analysis techniques. This approach involves analyzing key performance indicators of the studied application. For instance, Zakeralhoseini [7] demonstrated the application of feed-forward Neural Networks trained on CFD simulations of small-scale turbopumps. By selecting relevant input and output parameters for training, these models can accurately predict system behavior.

Alternatively, data analysis can focus on fundamental components of CFD results, such as spatial discretization of the domain. By investigating how changes in operational or geometrical conditions affect these components, one can infer the impact on overall performance. Luo et al. [12] utilized POD to analyze CFD data and 'simulate' new geometries for the blades of an axial turbine. This approach enables the prediction of performance under varying conditions, facilitating design optimization and efficiency improvement.

1.3.1 Reduced Order Models

ROMs are used to perform numerical calculations with small numbers of Degrees Of Freedom (DOFs). Reduced order models in fluid mechanics have multiple aims. They provide quantitatively accurate descriptions of the dynamic behavior of a system at a computational cost much lower than the original numerical model. Additionally, they offer means by which system

dynamics can be readily interpreted, as demonstrated by Wei et al. in [13]. These models are also useful in shape optimization, where multiple simulations may be necessary to evaluate a certain design space, as presented by Marques et al. in [14].

1.3.2 Proper Orthogonal Decomposition

Proper Orthogonal Decomposition is a Principal Component Analysis (PCA) technique used to identify dominant patterns in data sets. This technique has been widely used for different applications: in low-dimensional description of turbulent flows [15], structure vibration [16], image processing [17], the particle image velocimetry technology, and aerodynamic airfoil shape optimization [18].

POD was introduced to the fluid-dynamics community by Lumley in 1967 [15] as an attempt to decompose the random vector field representing turbulent fluid motion into a set of deterministic functions, each capturing some portion of the total fluctuating kinetic energy in the flow. The goal was to use a limited number of these deterministic functions, the POD modes, to provide insight into the organization of the flow.

A variant of POD, the Gappy POD has been proven to be successful by Bui-Thanh and Damodaran [18]. Both, for its capability of flow reconstruction in the presence of missing data in the flow around an airfoil and its application for the inverse design starting from the imposing of a desired pressure distribution along an airfoil.

Despite its advantages, using POD in a non-intrusive way for design optimization presents limitations, particularly when dealing with transonic or supersonic flows. These flows exhibit strong non-linearity in their characteristics, as investigated by Iuliano et al. [19]. Their study demonstrated that a zonal strategy, applying POD only in the domain far from the airfoil, was more accurate than a completely non-intrusive strategy. However, it was also highlighted that using more POD modes yields more precise results, though still significantly different from the true solution.

The strong linearity of the POD formulation is one reason for these limitations. To address this issue, non-linear interpolation of the POD coefficients can be implemented. Ripepi et al. [20] at DLR enhanced POD capabilities using Radial Basis Functions, Neural Networks, or Physics-Based Constraints. Their research showed that accuracy could be improved without sacrificing the computational efficiency of POD.

In the context of aerodynamic shape optimization, Marques et al. [14] successfully applied a non-intrusive Aerodynamic Shape Optimization with a POD-DEIM Based Trust Region Method. Their results showed good agreement with CFD-based adjoint optimization and significant time savings.

When considering turbomachinery, there are fewer applications of reduced order models and non-intrusive POD in the literature. Wei et al. [13] used POD for fast prediction of centrifugal pump performance. They constructed the POD model using 80 samples, tuning mass flow and rotational speed while keeping the geometry unchanged. Their method achieved average errors of 7 % for the velocity field, 1 % for the pressure field, and 3 % for the pump head prediction, with POD simulations being 2000 times faster than CFD.

Duan et al. [21] focused on performance improvement through POD-based data mining of the NASA Rotor 37. After an initial optimization using Particle Swarm Optimization (PSO), they selected CFD simulations providing higher efficiency and applied POD to study the geometrical changes impacting compressor performance. Similarly, Luo et al. [22] worked on the last stage

of a low-speed 4.5-stage compressor, using a POD hybrid model to modify the spanwise stagger angle and the camber of the stator blade while maintaining constant mass flow rate. Luo's work includes a comprehensive comparison of different hybrid-based POD models, utilizing various interpolation methods to find the best coefficients for reconstructing solutions for new geometries.

Regarding centrifugal compressors, Zhang et al. [23] parameterized the pump blade shape using quartic Bezier curves. They generated a substantial amount of data to build the POD model, specifically through Gappy POD, to obtain the geometry that would provide the desired pressure distribution on the blade. Their results showed good accuracy.

Additional research has explored intrusive POD methods applied to turbomachinery [24, 25, 26], further demonstrating the versatility and potential of POD in various fluid dynamics applications.

1.4 Organization of the thesis

This thesis begins with a brief overview of the work conducted by Zakeralhoseini during his Ph.D., as the CFD simulations from his research provided the foundational data. Following this, a mathematical description of Proper Orthogonal Decomposition is presented, illustrating its application to CFD data and its adaptation to develop a reduced order model for a turbopump.

The methodology section details the steps required to integrate POD with commercial software like `ANSYS`, covering both results extraction and data adaptation for applying POD. It provides an in-depth explanation of the POD model construction, focusing on how data is reconstructed to predict the performance of new geometrical configurations. This section concludes with an explanation of how pump performance is evaluated.

The results section demonstrates the extraction of information from specific regions of the pump's flow field and the overall prediction of pump performance. It highlights the success of POD in optimizing the turbopump by tuning relevant parameters. Finally, the discussion addresses the results and limitations of POD, offering a comprehensive view of the work conducted and the effectiveness of POD in inferring a reduced order model capable of accurately predicting the impact of design changes on a centrifugal pump's flow field.

Chapter 2

Methods

2.1 Turbopump for Organic Rankine Cycle

Zakeralhoseini [6] designed and analyzed a small-scale turbopump intended for powering an 8 kW ORC system aimed at waste heat recovery in truck engines. The design specifications of the turbopump are presented in Table 2.1.

Mass flow rate $\dot{m}_r$ (kg s^{-1})	Pressure rise Δp (Pa)	Rotational speed n (rpm)	Specific speed n_q
0.280	20.0	25000	8.3

Table 2.1: Design specification of the turbopump.

Zakeralhoseini developed a novel and fully parameterized design model, which was utilized to generate various turbopump geometries. These geometries were employed for conducting three-dimensional computations across the impeller stage. Comprehensive details about this model can be found in [8]. Additionally, a flowchart illustrating the structure of the model is presented in the following figure.

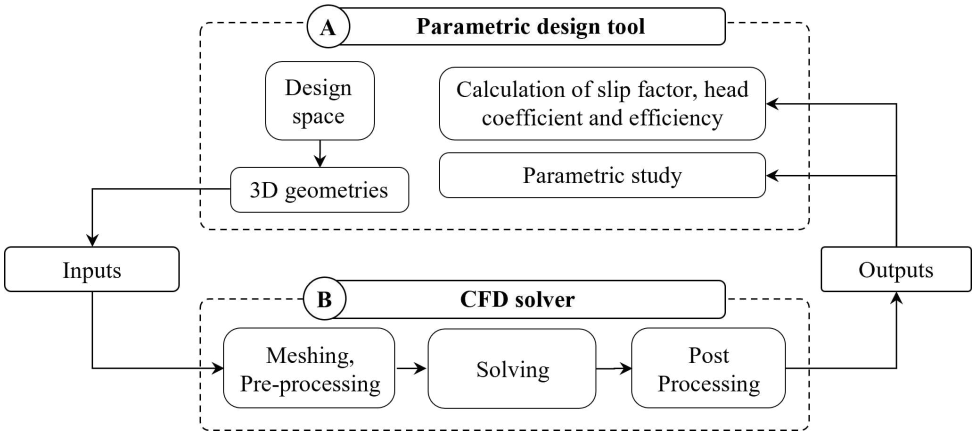


Figure 2.1: Methodology flowchart. Figure retrieved from [8].

The design of the turbopump is carried out in `MATLAB`, employing empirical correlations and adhering to established design practices within the field of pumps. Subsequently, a file detailing the main characteristics of the pump is generated and transferred to `BladeDesign`, a specialized program within `ANSYS` dedicated to generating 3D CAD models of turbomachinery. The resulting geometry is then meshed using `Turbogrid`, and CFD simulations are performed using `CFX`.

All simulations have been instrumental in identifying trends and general principles applicable to small-scale turbopumps. This category of turbomachinery may present challenges in adhering to classical similitude theory due to its extreme working conditions and compact size. Consequently, a substantial amount of CFD results were generated, providing an excellent opportunity to evaluate the effectiveness of Proper Orthogonal Decomposition (POD) in identifying a reduced order model. Such a model could significantly aid in future designs of these machines. Therefore, the extensive database of simulations served as the foundation of constructing a reduced order model.

In his work, Zakeralhoseini investigated various design variables, including the outlet diameter, blade angle at the outlet (β_{out}), number of blades (Z), tip clearance ratio (c/b), rotational speed (n), pitch of the splitter blade (ψ_{sp}), and relative meridional location of the splitter blade ($m'_{le,sp}$). The range of parameters explored is presented in Table 2.2. Further information regarding the parameters that are being tuned and their visual explanation can be found in [7].

Parameter	Symbol	Values	Unit
Blade angle	β_{out}	$22.5, 50, 90$	(deg)
Number of blades	Z	$4, 5, 6$	(-)
Tip clearance ratio	c/b	shrouded, $0.05, 0.1$	(%)
Rotational speed	n	$40, 50, 75$	(krpm)
Pitch of splitter blades	ψ_{sp}	$25, 37, 50, 62, 75$	(%)
Relative meridional location of splitter blades	$m'_{le,sp}$	$12, 25, 40, 55$	(%)

Table 2.2: Parameter values of the investigated design for an ORC turbopump.

The geometry of the turbopump designed by Zakeralhoseini is illustrated in Figure 2.2. In the figure, a six-blade geometry is depicted, comprising six main blades and six splitter blades. The longer blades are the main blades, while the shorter ones are the splitter blades. On the left side of the figure, the leading edge meridional position of the main blades is set at a value of 0, while the outlet position is set to 1. In this work, β_{2B} is indicated with β_{out}. Additionally, ψ_{sp} is defined as:

$$\psi_{sp} = \frac{\theta_{sp}}{\theta_{mb}}. \tag{2.1}$$

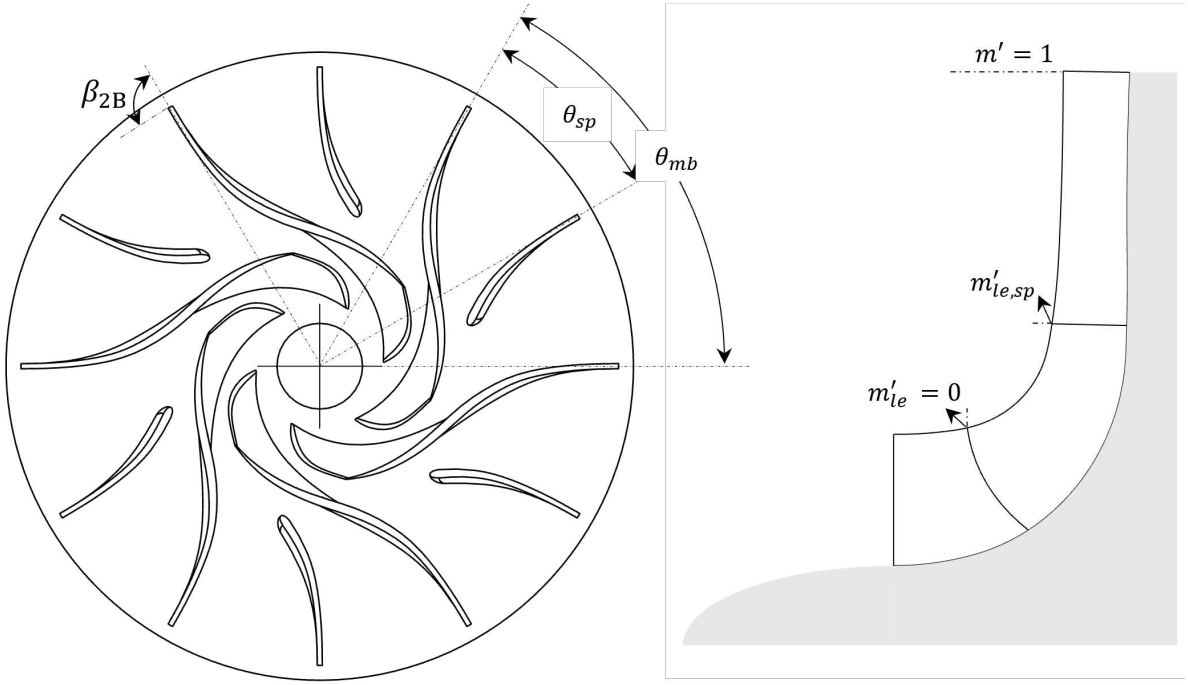


Figure 2.2: Turbopump geometry and parameters. Figure retrieved from [8].

2.2 Proper Orthogonal Decomposition

Proper Orthogonal Decomposition is a principal component analysis technique. Consider a vector field $u'(x,t)$ which can be expressed as the inner product of a deterministic set of spatial functions $\Phi_k(x)$ modulated by time coefficients $a_k(t)$:

$$u'(x,t) = \sum_{k=1}^{\infty} a_k(t)\Phi_k(x). \tag{2.2}$$

In a more general framework, $a_k(t)$ is also referred to as modal coefficient. The name Proper Orthogonal Decomposition comes from the fact that, 'proper' is a synonym of optimal, since the sequence $\sum_{k=1}^{\infty} a_k(t)\Phi_k(x)$ maximizes the kinetic energy that can be captured by the first n spatial modes. While orthogonal comes from the fact that the modes are orthonormal, which means that in a suitable function space one can write:

$$\iiint_x \Phi_{k_1}(x)\Phi_{k_2}(x)\mathrm{d}x = \begin{cases} 1 \text{ if } k_1 = k_2 \\ 0 \text{ if } k_1 \neq k_2. \end{cases} \tag{2.3}$$

This property implies that each time coefficient $a_k(t)$ only depends on its spatial mode $\Phi_k(x)$. This is shown also by multiplying Equation 2.2 with $\Phi_k(x)$ and integrating over the volume, which yields:

$$a_k(t) = \iiint_x u'(x,t)\Phi_k(x)\mathrm{d}x. \tag{2.4}$$

In order to align with the purpose of this work, the vector field $u'(x,t)$ will not be considered

time- and space-dependant. Instead, the time variable will be replaced by a different parameter defined as μ. Thus, the modified representation is:

$$u'(x,\mu) = \sum_{k=1}^{\infty} a_k(\mu)\Phi_k(x). \tag{2.5}$$

Throughout this thesis, the term 'snapshot' will refer to a complete vector field for a specific condition μ, such as a particular geometrical configuration of a 2-dimensional channel. To analyze this field using POD, it is necessary to have consistent information on the variable of interest, particularly in terms of spatial discretization.

In Figure 2.3, a flow field for a 2D channel is presented. Assuming there are m simulations for different geometrical configurations (μ_i with $i = 1, 2, ..., m$), where the height and length of the channel vary. Each configuration is discretized using a grid with the same number of cells, specifically $N_1 = 10$ and $N_2 = 6$.

For this example, the POD snapshot would be the flow field for one specific configuration, represented as $U(x)$, where $x = (x_1, x_2)$. Here, x_1 and x_2 denote the vertical and horizontal directions, respectively.

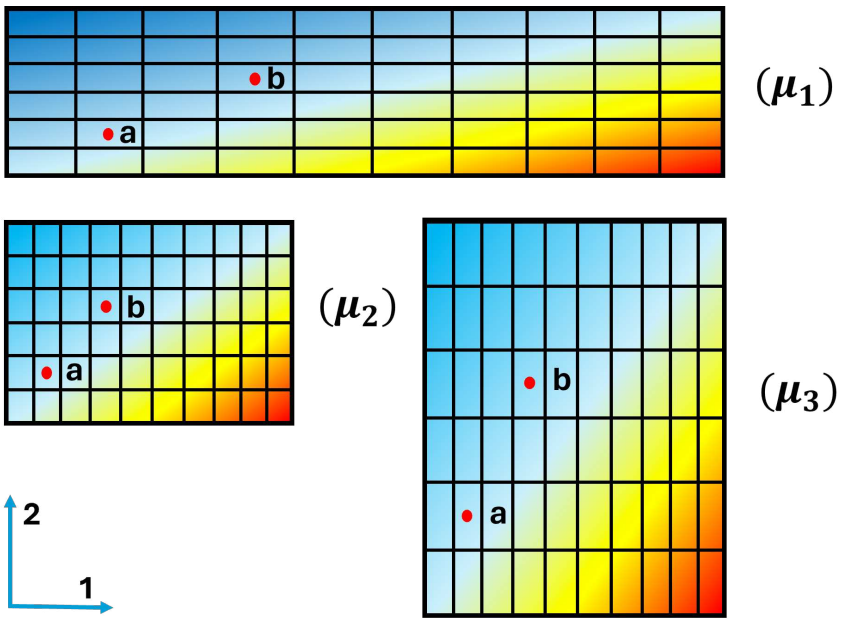


Figure 2.3: 2D flow field example for different geometrical configurations and constant discretization.

2.2.1 2-D case

To analyze two points in the flow field, denoted as a and b, across all m configurations, the data for POD computation would consist of two arrays of m velocity values ($U_a(\mu_i)$ and $U_b(\mu_i)$ with $i = 1, 2, ..., m$). These arrays can be concatenated into a single $m \times 2$ matrix S, known as the matrix of snapshots:

$$S = \begin{pmatrix} U_a(\mu_1) & U_b(\mu_1) \\ U_a(\mu_2) & U_b(\mu_2) \\ \vdots & \vdots \\ U_a(\mu_m) & U_b(\mu_m) \end{pmatrix}. \tag{2.6}$$

Since the main interest lays in flow dynamics it is possible to subtract the average velocities $\overline{U}_a$ and $\overline{U}_b$ from the respective columns and obtain a new snapshot matrix consisting only of the velocity fluctuations: $u'_a(\mu) = U_a(\mu) - \overline{U}_a$ and $u'_b(\mu) = U_b(\mu) - \overline{U}_b$. This new matrix is denoted as U:

$$U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ \vdots & \vdots \\ u_{m1} & u_{m2} \end{pmatrix} = \begin{pmatrix} u'_a(\mu_1) & u'_b(\mu_1) \\ u'_a(\mu_2) & u'_b(\mu_2) \\ \vdots & \vdots \\ u'_a(\mu_m) & u'_b(\mu_m) \end{pmatrix}. \tag{2.7}$$

To identify the modes characterizing the flow field, the covariance matrix, C, is first calculated as shown in Equation 2.8. For the 2D case, this calculation results in a 2×2 matrix:

$$C = \frac{1}{m-1} U^T U = \frac{1}{m-1} \begin{pmatrix} \sum_{i=1}^{m} u'^2_a(\mu_i) & \sum_{i=1}^{m} u'_a(\mu_i)\, u'_b(\mu_i) \\ \sum_{i=1}^{m} u'_b(\mu_i)\, u'_a(\mu_i) & \sum_{i=1}^{m} u'^2_b(\mu_i) \end{pmatrix}. \tag{2.8}$$

The diagonal elements of the covariance matrix are the respective variances of u'_a and u'_b, while each off-diagonal element represents the covariance between u'_a and u'_b. Consequently, the covariance matrix is necessarily symmetric. If the off-diagonal terms are non-zero, it implies a statistical correlation between u'_a and u'_b. A diagonal covariance matrix would indicate statistical independence between u'_a and u'_b. Given that C, defined in Equation 2.8, is symmetric, the matrix can be diagonalized and thus can be expressed as eigenvalues and eigenvectors. If ordered by magnitude, the eigenvalues with the largest value represent the most dominant modes whereas the smallest represent the least dominant ones. Formally, the covariance matrix C is diagonalized as:

$$\begin{aligned} C &= \Phi \Lambda \Phi^{-1} = \Phi \Lambda \Phi^T \\ &= \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{pmatrix}, \end{aligned} \tag{2.9}$$

where the two columns of Φ are the eigenvectors of C. The second equality in Equation 2.9 comes from the fact that C is symmetric, which implies that its eigenvectors are orthonormal or, equivalently, that Φ is an orthogonal matrix ($\Phi^{-1} = \Phi^T$).

The variance on the major axis (defined by the first mode) is given by the first (largest) eigenvalue λ_1 and the variance on the minor axis (defined by the second mode) is given by the second (smallest) eigenvalue λ_2. Since the covariance matrix C' of the data projected on the proper orthogonal basis is the diagonal matrix Λ:

$$C' = \frac{1}{m-1} A^T A = \frac{1}{m-1} (U\Phi)^T (U\Phi) = \frac{1}{m-1} \left(\Phi^T U^T U \Phi \right) = \Phi^T C \Phi = \Phi^T \Phi \Lambda \Phi^T \Phi = \Lambda. \tag{2.10}$$

Finally, having identified the main eigenvectors of the covariance matrix it is possible to project the matrix U onto each eigenvector and obtain the matrix of the modal coefficient A as:

$$A = U\Phi = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ \vdots & \vdots \\ u_{m1} & u_{m2} \end{pmatrix} \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}. \tag{2.11}$$

Given the orthogonality of Φ, where $\Phi^{-1} = \Phi^T$. Therefore, the original matrix U can be easily expressed in terms of A:

$$A = U\Phi \Rightarrow U = A\Phi^{-1} = A\Phi^T \tag{2.12}$$

or, written explicitly,

$$\begin{aligned} U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ \vdots & \vdots \\ u_{m1} & u_{m2} \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{pmatrix} \begin{pmatrix} \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}\phi_{11} + a_{12}\phi_{12} & a_{11}\phi_{21} + a_{12}\phi_{22} \\ a_{21}\phi_{11} + a_{22}\phi_{12} & a_{21}\phi_{21} + a_{22}\phi_{22} \\ \vdots & \vdots \\ a_{m1}\phi_{11} + a_{m2}\phi_{12} & a_{m1}\phi_{21} + a_{m2}\phi_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}\phi_{11} & a_{11}\phi_{21} \\ a_{21}\phi_{11} & a_{21}\phi_{21} \\ \vdots & \vdots \\ a_{m1}\phi_{11} & a_{m1}\phi_{21} \end{pmatrix} + \begin{pmatrix} a_{12}\phi_{12} & a_{12}\phi_{22} \\ a_{22}\phi_{12} & a_{22}\phi_{22} \\ \vdots & \vdots \\ a_{m2}\phi_{12} & a_{m2}\phi_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \begin{pmatrix} \phi_{12} & \phi_{21} \end{pmatrix} + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} \begin{pmatrix} \phi_{12} & \phi_{22} \end{pmatrix} \\ &= \tilde{U}^1 + \tilde{U}^2, \end{aligned} \tag{2.13}$$

where, by definition, for $k = 1, 2$:

$$\tilde{U}^k = \begin{pmatrix} \tilde{u}^k_{11} & \tilde{u}^k_{12} \\ \tilde{u}^k_{21} & \tilde{u}^k_{22} \\ \vdots & \vdots \\ \tilde{u}^k_{m1} & \tilde{u}^k_{m2} \end{pmatrix} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix} \begin{pmatrix} \phi_{1k} & \phi_{2k} \end{pmatrix}. \tag{2.14}$$

The original velocity field fluctuations have been decomposed into two contributions: one from the first mode $\tilde{U}^1$ and another from the second mode $\tilde{U}^2$. Reconstruction of the data using solely the dominant mode could yield reasonably accurate results, as demonstrated by [27].

2.2.2 n-dimensional case

This can be extended to the n-dimensional case, where n is the total amount of point values or cells on the flow field, starting from re-writing the complete U matrix as:

$$U = \begin{pmatrix} u_{11} & \dots & u_{12} \\ u_{21} & \dots & u_{22} \\ \vdots & & \vdots \\ u_{m1} & \dots & u_{m2} \end{pmatrix} = \begin{pmatrix} u'_a(x_1, y_1, \mu_1) & \dots & u'_b(x_{N_1}, y_{N_2}, \mu_1) \\ u'_a(x_1, y_1, \mu_2) & \dots & u'_b(x_{N_1}, y_{N_2}, \mu_2) \\ \vdots & & \vdots \\ u'_a(x_1, y_1, \mu_m) & \dots & u'_b(x_{N_1}, y_{N_2}, \mu_m) \end{pmatrix}. \tag{2.15}$$

The covariance matrix C is now of the size $n \times n$, which could require some consistent computational effort to calculate its eigenvalues and eigenvectors. Solving the `MATLAB` function `eig` will provide n eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ and a set of n eigenvectors arranged as columns in an $n \times n$ matrix Φ:

$$\Phi = \begin{pmatrix} \phi_{11} & \dots & \phi_{1n} \\ \phi_{21} & \dots & \phi_{2n} \\ \vdots & & \vdots \\ \phi_{n1} & \dots & \phi_{nn} \end{pmatrix}. \tag{2.16}$$

The n eigenvectors (the n columns of Φ), are the proper orthogonal modes of the dataset. They can be seen as the axes of an n-dimensional ellipsoid that fits the total dataset in a n-dimensional space. The modes are arranged in descending order based on the variance of the data along the directions they define.

Similar to the 2-dimensional case, the original dataset can be projected onto each of the n modes by expressing $A = U\Phi$:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} u_{11} & \dots & u_{1n} \\ u_{21} & \dots & u_{2n} \\ \vdots & & \vdots \\ u_{m1} & \dots & u_{mn} \end{pmatrix} \begin{pmatrix} \phi_{11} & \dots & \phi_{1n} \\ \phi_{21} & \dots & \phi_{2n} \\ \vdots & & \vdots \\ \phi_{n1} & \dots & \phi_{nn} \end{pmatrix}. \tag{2.17}$$

After applying some linear algebra, it is possible to decompose this as:

$$\begin{aligned} \begin{pmatrix} u_{11} & \dots & u_{1n} \\ u_{21} & \dots & u_{2n} \\ \vdots & & \vdots \\ u_{m1} & \dots & u_{mn} \end{pmatrix} &= \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \begin{pmatrix} \phi_{11} & \dots & \phi_{n1} \end{pmatrix} + \dots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \begin{pmatrix} \phi_{1n} & \dots & \phi_{nn} \end{pmatrix} \\ &= \begin{pmatrix} \tilde{u}^1_{11} & \dots & \tilde{u}^1_{1n} \\ \tilde{u}^1_{21} & \dots & \tilde{u}^1_{2n} \\ \vdots & & \vdots \\ \tilde{u}^1_{m1} & \dots & \tilde{u}^1_{mn} \end{pmatrix} + \dots + \begin{pmatrix} \tilde{u}^n_{11} & \dots & \tilde{u}^n_{1n} \\ \tilde{u}^n_{21} & \dots & \tilde{u}^n_{2n} \\ \vdots & & \vdots \\ \tilde{u}^n_{m1} & \dots & \tilde{u}^n_{mn} \end{pmatrix}, \end{aligned} \tag{2.18}$$

with,

$$\begin{pmatrix} \tilde{u}^k_{11} & \dots & \tilde{u}^k_{1n} \\ \tilde{u}^k_{21} & \dots & \tilde{u}^k_{2n} \\ \vdots & & \vdots \\ \tilde{u}^k_{m1} & \dots & \tilde{u}^k_{mn} \end{pmatrix} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix} \begin{pmatrix} \phi_{1k} & \dots & \phi_{nk} \end{pmatrix} = \tilde{U}^k. \tag{2.19}$$

It can be noted that each matrix $\tilde{U}^k$ has the same dimension as U and that:

$$U = \sum_{k=1}^{n} \tilde{U}^k. \tag{2.20}$$

This implies that the initial flow field has been decomposed into a sum of n contributions from n proper orthogonal modes. This is equivalent to the initial definition of the POD as stated in Equation 2.5. In this study, $a_k(\mu)$, the modal coefficient of the POD mode Φ_k, is represented as the column vector $(a_{1k}, a_{2k}, \dots, a_{mk})^T$, and Φ_k is represented as the row vector $(\Phi_{1k}, \Phi_{2k}, \dots, \Phi_{nk})$.

2.2.3 Energy considerations

The POD ranks each modes according to their contribution to the total variance, which can be evaluated from the eigenvalues of the covariance matrix. To express in which amount each mode contributes to the description of the full data, it is possible to use the total fluctuating kinetic energy (TKE) which can be calculated as:

$$\text{TKE} = \frac{\lambda_i}{\Sigma_k \lambda_k}. \tag{2.21}$$

This definition is generally used to rank the modes, identify the dominant ones decide how many modes to use in reconstructing the data and truncate the combination of the modes as expressed in Equation 2.20 where n-modes have been used.

2.2.4 Snapshot POD

Sirovich [17] introduced an alternative approach to traditional Proper Orthogonal Decomposition (POD), known as Snapshot POD. This method leverages the inherent symmetry of POD, as illustrated in Equation 2.5, where there is no fundamental distinction between the geometrical variable μ and the spatial variable x. Consequently, interchanging μ and x in the computations is feasible and can enhance computational performance. In this approach, the correlation matrix is constructed as $C_s = \frac{1}{m-1} UU^T$, which is of size $m \times m$ rather than $n \times n$ as in the direct method. Here, m represents the number of available snapshots or simulations where the geometry has been altered, while n denotes the number of points discretizing the flow field in each snapshot. The eigenvalues and eigenvectors of C_s are then computed in the same manner as the direct method. However, the decomposition is now performed in m-dimensional space, resulting in a set of m eigenvalues and eigenvectors. These m eigenvectors of C_s serve as geometrical modes, analogous to the m modal coefficients in the direct method. To obtain the spatial coefficients, it is possible to calculate: $\Phi_s = U^T A_s$. Consequently, U can be reconstructed as follows:

$$\Phi_s = U^T A_s \Rightarrow U^T = \Phi_s {A_s}^{-1} = \Phi_s {A_s}^T \Rightarrow U = A_s {\Phi_s}^T. \tag{2.22}$$

As for the direct method it is possible to write:

$$U = \sum_{k=1}^{m} \tilde{U}_s^k, \tag{2.23}$$

with

$$\tilde{U}_s^k = \begin{pmatrix} (a_s)_{1k} \\ (a_s)_{2k} \\ \vdots \\ (a_s)_{mk} \end{pmatrix} \begin{pmatrix} (\phi_s)_{1k} & \dots & (\phi_s)_{nk} \end{pmatrix}. \tag{2.24}$$

The snapshot POD method is less computationally demanding due to the size of the correlation matrix C_s. When dealing with CFD data, this matrix is most likely smaller than the correlation matrix C used in the direct method. This is because the number n of spatial measurement points is often larger than the number m of snapshots. Finally, as discussed by Weiss [27], the results are essentially the same: when $m < n$, the last $n - m$ eigenvalues of the direct method are zero and the last $n - m$ modes have no effect. Thus, whenever $m \neq n$, either procedure returns $\min(m, n)$ modes.

Along the thesis it will be referred to the Snapshot POD as POD, since this is the only method that was used and represents the most suitable solution for the current application.

2.3 CFD results extraction

The information that can be extracted from a CFD simulation of a pump is extensive, encompassing parameters such as pressure, velocities, temperature, and energy, among others. This data is typically obtained at the central points of all cells constituting the mesh of the pump. When the geometry and mesh of the pump remain unchanged, as in the work of Wei et al. [13], and only operational conditions are varied, the location of the flow information remains constant. This stability facilitates the application of POD models and enables straightforward data analysis.

Challenges arise when the geometry of the pump is modified, as in the case of Zakeralhoseini's design optimization. Changes in geometry lead to alterations in the mesh, complicating the application of POD, which relies on consistent results to identify main flow modes and correlations between points under different conditions. One potential solution is mesh morphing, which aims to maintain a constant mesh topology and preserve the relative positioning of cells with respect to the domain boundary. However, mesh morphing is most effective for minor geometric changes; substantial alterations may result in poor quality CFD results due to excessive stretching or contraction of the mesh.

In the developed methodology, mesh properties were configured according to specific settings chosen after conducting a thorough grid independence analysis. The settings used for the meshing are detailed in Table 2.3.

Parameter	Setting
Global size factor	1.20
Inlet growth ratio	1.80
Maximum expansion rate	1.30
Boundary layer specification	Reynolds number and first element y^+
Reynolds number	Set based on blade chord
Vertex offset specification mode	y^+
First element y^+	< 2

Table 2.3: Specification of the computational grid settings.

Having common settings for mesh, results in different mesh sizes and topologies due to significant changes in geometrical features, such as adjustments to splitter pitch position and length. While this strategy guarantees high-quality meshes, it complicates the application of POD decomposition. This is because the positions of cells vary between geometries, making it difficult to assume that the n^{th} cell will correspond to a certain cell on to the main blade leading edge across different configurations.

`Python` was used to extract all the information in a consistent and efficient manner, particularly the `re` package for modifying text files used to create a journal file[1]. This journal file can be read by `ANSYS`, specifically `CFX`, to automate actions such as defining a surface or polyline, specifying parameters to extract, and saving to `CSV` files.

Using the `Python` package `subprocess`, `CFX` is run through `Python` and the journal file executed, automating the process and eliminating the need for manual intervention.

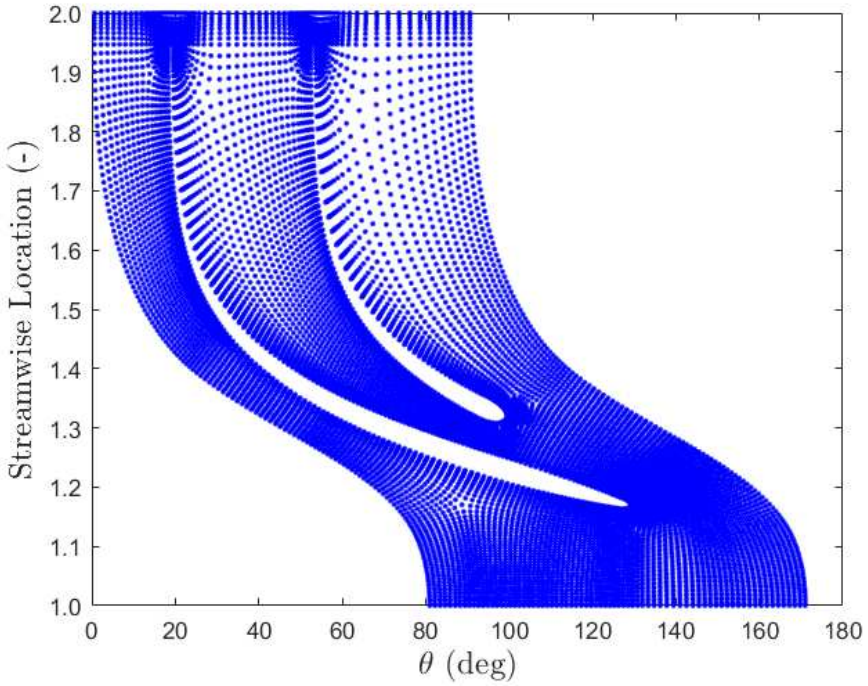
2.4 Mesh interpolation

To address the challenge posed by the necessity of having a constant mesh, a method was developed for projecting the flow field information onto locations within the domain that maintain a constant topology. This requirement distinguishes this work from existing literature and underscores the need for innovative approaches to adapt POD techniques to varying geometries.

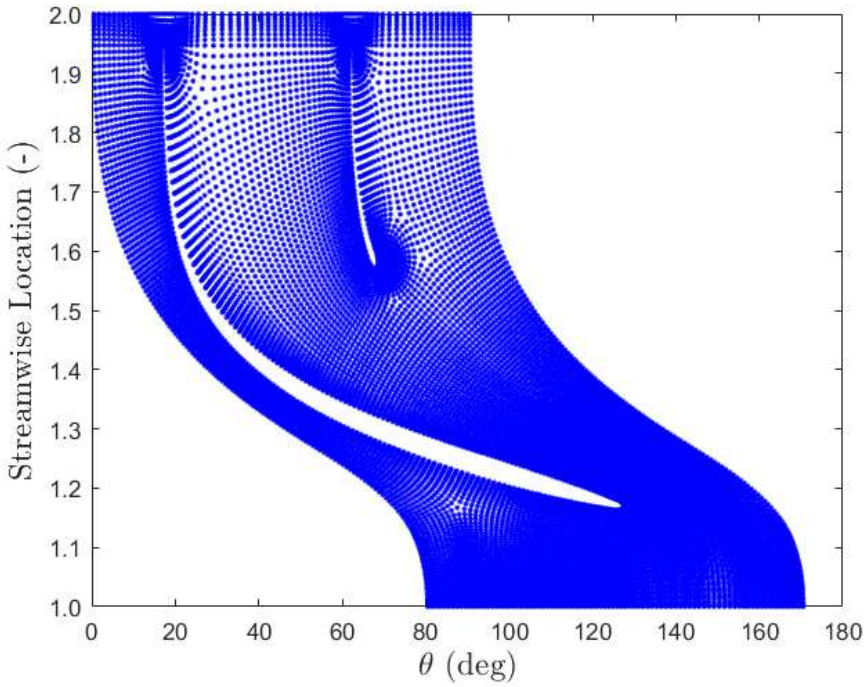
2.4.1 Span

Initially, the efficacy of POD was tested by analyzing its ability to provide results for the velocity and pressure fields of the blade to blade plane at various spans. An interpolation method for the mesh was developed. This method involved generating a grid with full control over the number of points using `MATLAB`, starting from the geometrical information of the turbopump configuration being analyzed.

[1] A journal file contains a sequence of `ANSYS` commands, arranged as they would be typed interactively into the program or entered through the Graphic User Interface or Text User Interface.



(a) $m'_{le,sp} = 0.25$, $\psi_{sp} = 0.25$



(b) $m'_{le,sp} = 0.50$, $\psi_{sp} = 0.50$

Figure 2.4: CFD mesh of the blade to blade surface at $S_N = 0.5$ for a four-blades shrouded turbopump.

Figure 2.4**(a)** and Figure 2.4**(b)** display two examples of the original mesh of the blade to blade plane at a normalized span (S_N) of 0.5. It should be evident how the mesh is clustered around the blade, particularly at the leading and trailing edges, to ensure proper resolution of flow in these intricate regions during the CFD calculations.

Moreover, it should be evident that there is no order in the positioning of the cells between the two presented meshes. To address these issues, the developed interpolation methods provides full control over the number of points in both the θ- and stream-wise directions. The domain has been divided into four areas in the streamwise direction: before the leading edge of the main blade, between the leading edge of the main blade and the leading edge of the splitter blade, between the leading edge of the splitter blade and its trailing edge, and after the trailing edge of the splitter blade (which is aligned with the trailing edge of the main blade).

In the θ-wise direction, the first and last sections of the streamwise direction are characterized by a single area each, while the second section (between the leading edges) is split into two areas, and the region around the splitter blade is divided into three areas. An explanatory diagram is presented in Figure 2.5.

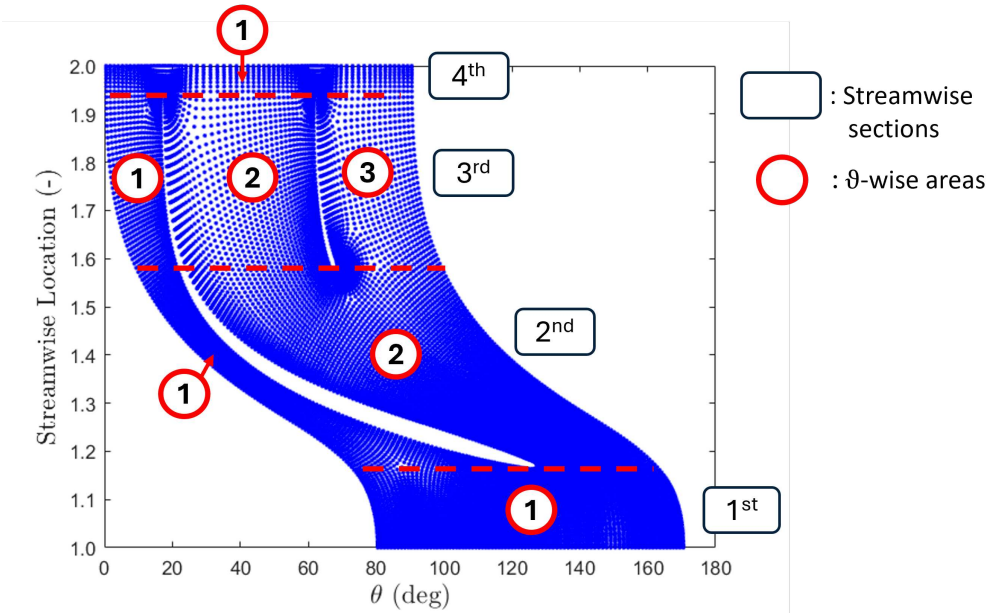


Figure 2.5: Graphical explanation of the discretized domain.

An example of the discretizations that can be performed with the developed function is presented in Figure 2.6, illustrating four different geometries. In the figure, the domains are under-discretized to facilitate visual inspection of the domain discretization. This example showcases the final outcome of the spatial discretization used to project the flow field results. Notably, in the second and third sections of the streamwise direction, the sum of the θ-wise points for each area is equal to the points present in the same direction in the first and last sections.

With the current discretization, it is possible to precisely determine the position of each point with respect to the domain boundaries. This precise positioning allows for the proper execution of POD and the construction of a covariance matrix that effectively serves as an analysis of the flow.

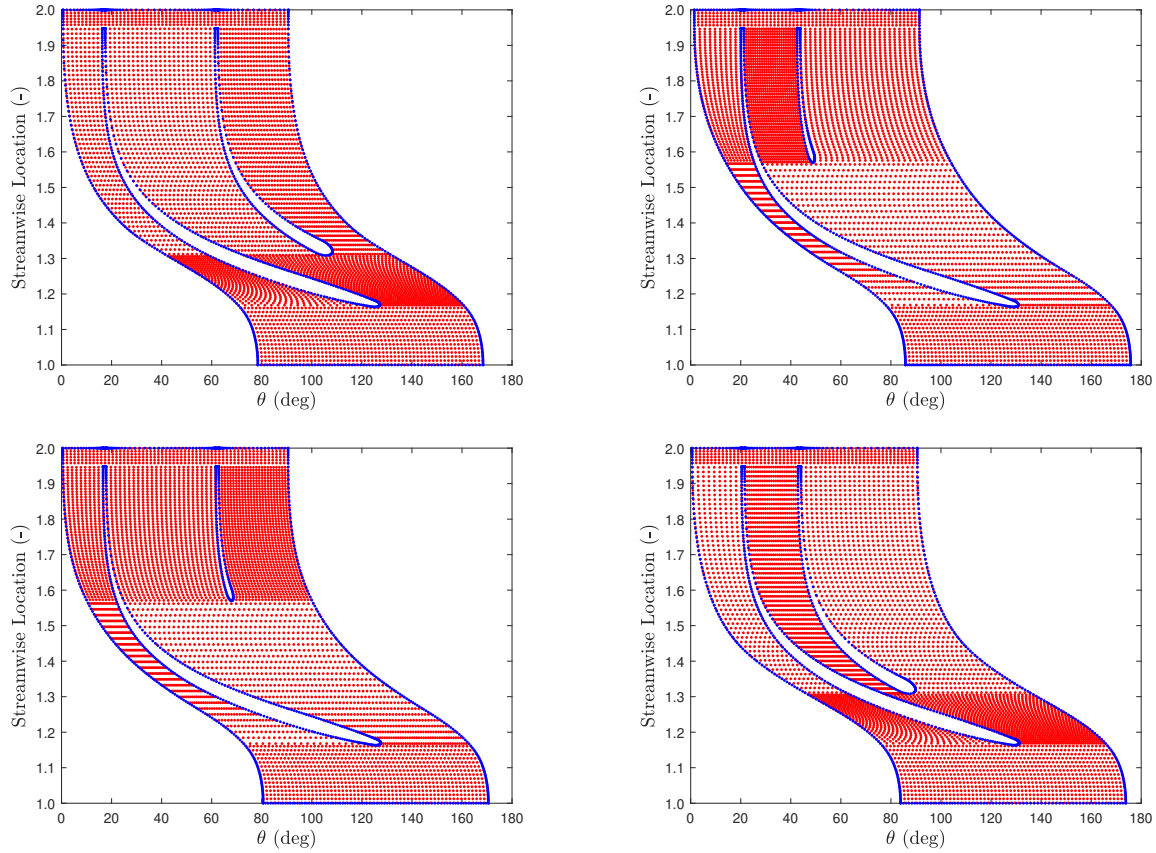


Figure 2.6: New interpolation comparison.

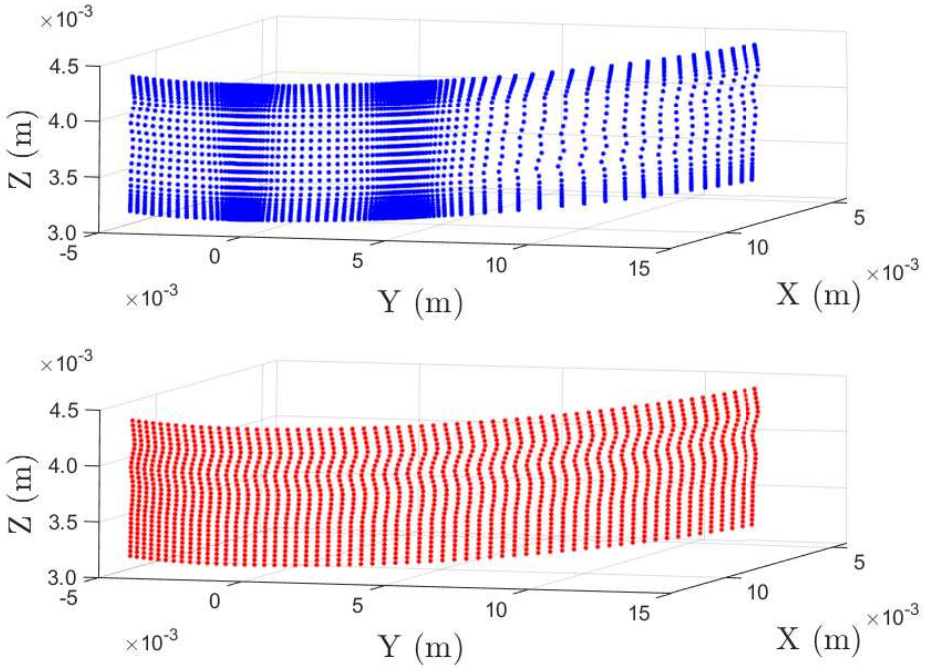
2.4.2 Inlet and outlet

A similar interpolation strategy has been applied to the outlet and inlet surfaces, as these surfaces are crucial for evaluating the performance of the turbopump, including parameters such as pressure rise (head), efficiency, and slip factor.

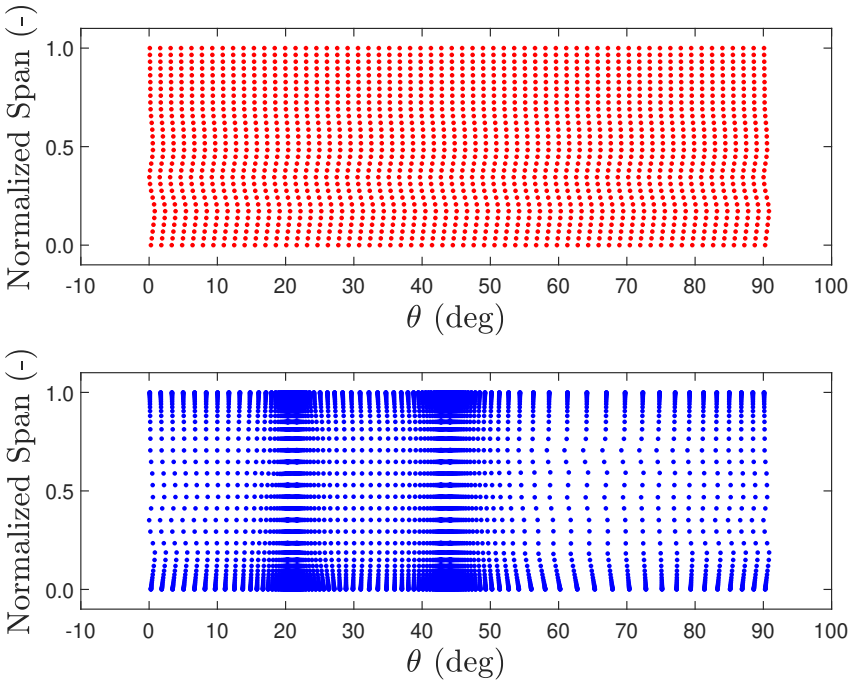
In this instance, the discretization is carried out according to the span- and θ-wise directions. This choice addresses certain limitations in results extraction from `CFXpost`, necessitating the development of a specific function to handle numerical errors present in the extracted geometrical information.

In the top part of Figure 2.7, the original outlet surface mesh is displayed. It is evident that a denser clustering of cells is present in the lower part (hub) and the upper part (shroud), are ensuring a y^+ value below two on the walls and a smooth transition to larger cells in the bulk flow. Additionally, a higher density of cells is present in correspondence with the blades. The interpolated mesh is displayed in the bottom part of Figure 2.7. It should be noted that this mesh is underdiscretized to better visualize the function's capability in properly discretizing the domain. While Figure 2.7**(a)** showcases the 3D view of the outlet, Figure 2.7**(b)** presents the outlet mesh projected in the spanwise and θ-wise directions, facilitating visualization.

While the original and interpolated mesh for the inlet projected on the Span-θ plane can be visualized in Figure 2.8, the original mesh shows a clustering of cells close to the hub and shroud, similar to the outlet. The 3D mesh is visualized in Figure A.1 and Figure A.2, which are located in the appendix.



(a) 3D visualization.



(b) 2D projection on the span-θ plane.

Figure 2.7: Original and interpolated mesh at the outlet ($m'_{le,sp} = 0.50$, $\psi_{sp} = 0.25$).

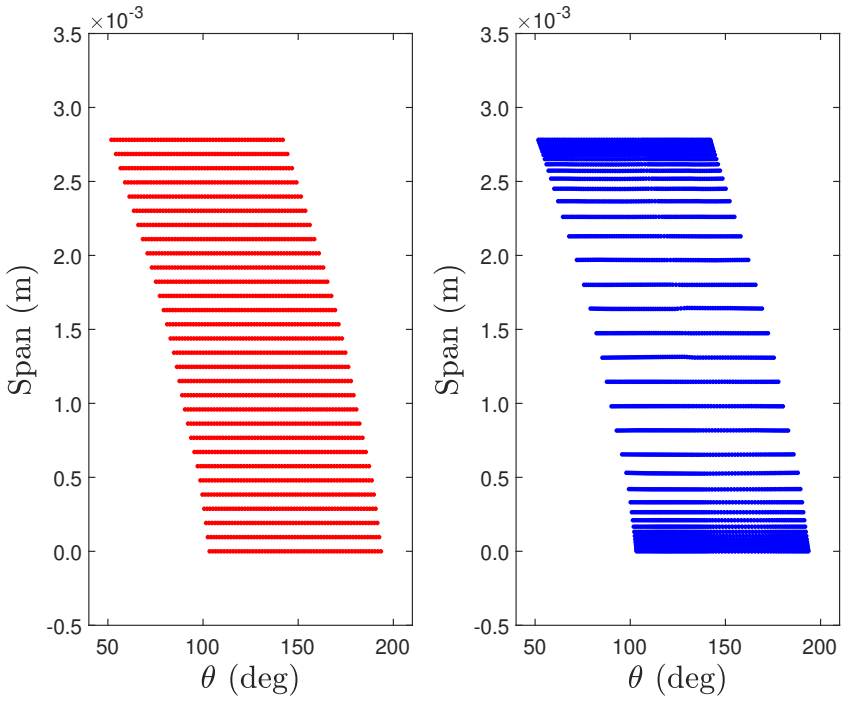


Figure 2.8: Original and interpolated mesh at the inlet, projected on the span-θ plane ($m'_{le,sp} = 0.50$, $\psi_{sp} = 0.25$).

2.4.3 Blades

Another surface that was extracted and adapted for the POD application is the blade surface, including both the main and splitter blades. These surfaces are chosen due to their significance in turbomachinery and the ultimate goal of using them to evaluate pump performance, specifically the power transferred from the shaft to the fluid.

This process involves extracting various polylines at different span levels. A polyline represents the intersection between a surface (at a defined span) and the blade surface. Since the number of cells varies between different geometrical configurations of the pump, it is necessary to ensure that each of these polylines contain the same number of points. Additionally, it is important to determine the positions of these points with respect to both the leading and trailing edges, as well as on the suction or pressure sides. Figure 2.9 displays two polylines at the same span level, one on the main blade and the other on the splitter blade.

The discretization is carried out line by line, beginning by identifying the leading edge of the blade as the starting point. The original points are then ordered such that the points on the pressure side, from the leading edge to the trailing edge, are listed first, followed by those on the suction side, from the trailing edge to the leading edge. Subsequently, this discretization is used to project the flow field of interest, and will be used to construct the POD. Figure 2.10 displays the original mesh of one polyline and its interpolation in 2D. The main blade, discretized into 40 polylines, is presented in Figure 2.11 with the pressure field projected on it.

The number of polylines used to discretize the blade can be controlled and this interpolation process is repeated along the height of the blade. A `Python` script was created to automate data extraction in `CFXpost` and ensure consistency in the extracted data.

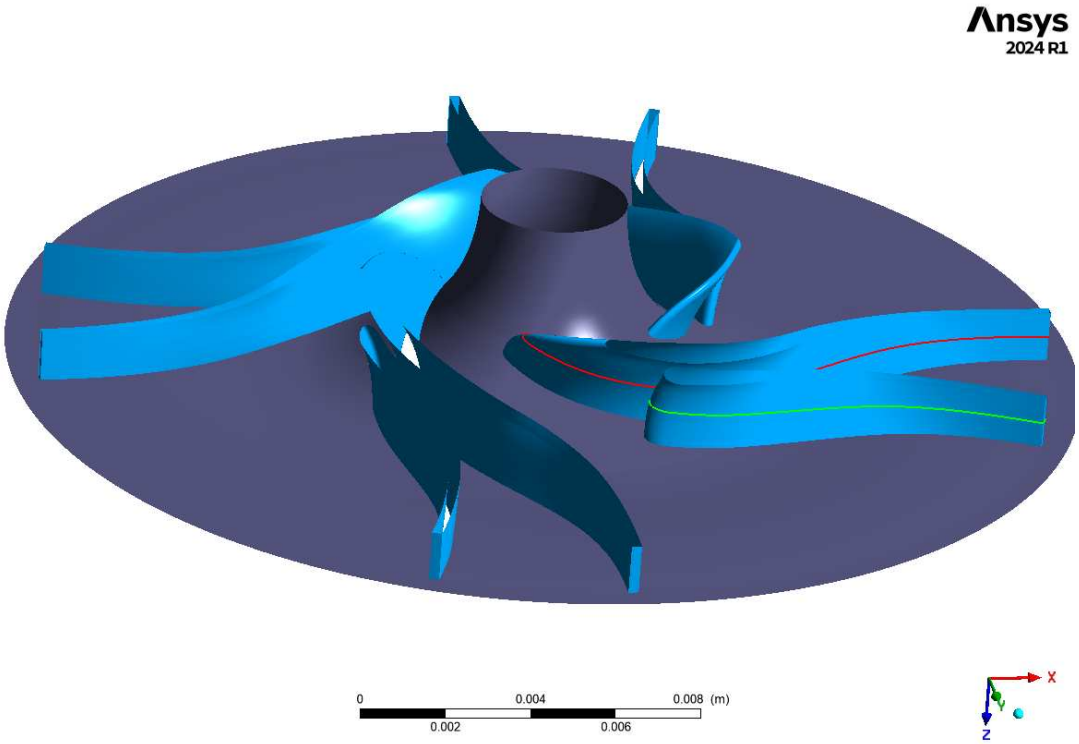


Figure 2.9: Polyline positions: The red line represents a polyline at a normalized span of 0.5 on the main blade, and the green line represents a polyline on the splitter blade ($m'_{le,sp} = 0.25$, $\psi_{sp} = 0.25$).

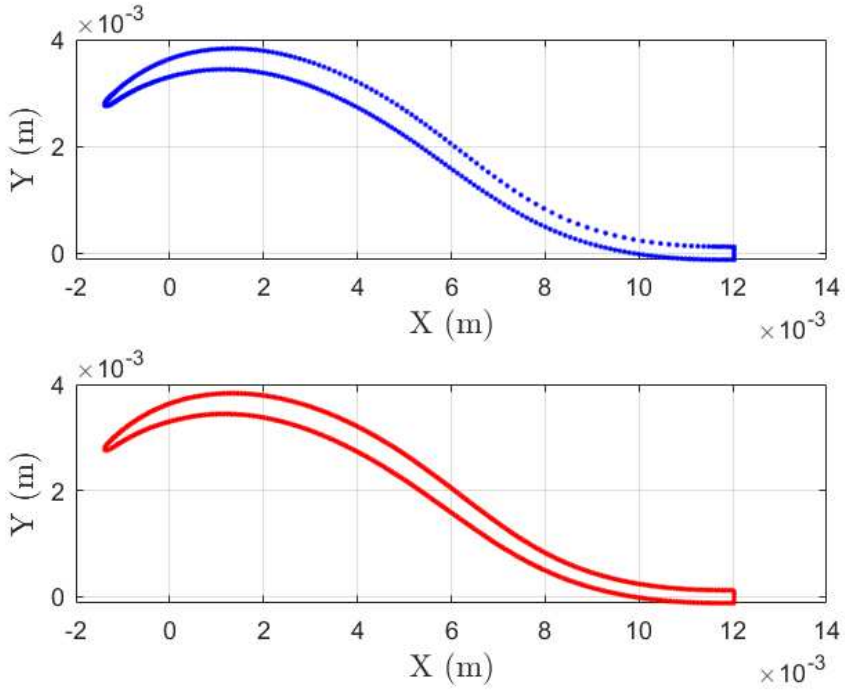


Figure 2.10: Original and interpolated mesh of the main blade polyline at a normalized span of 0.5 projected onto the XY plane ($m'_{le,sp} = 0.25$, $\psi_{sp} = 0.25$).

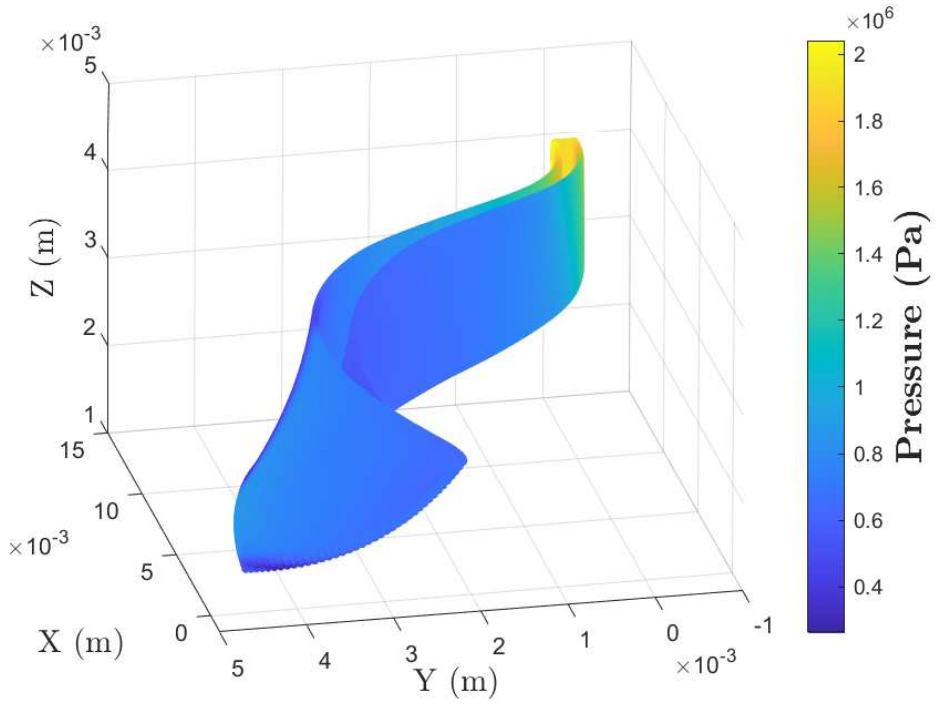


Figure 2.11: Main blade composed of 40 interpolated polylines from hub to shroud, with pressure field projection ($m'_{le,sp} = 0.25$, $\psi_{sp} = 0.25$).

2.4.4 Data format

By interpolating all of these surfaces, the data can be organized in the matrix presented in Equation 2.15 which is the required data format to apply POD. Although the absolute position of each point (x_n, y_n, z_n) may vary, their position relative to the domain boundary remains consistent. For instance, a point on the leading edge of the main blade touching the hub or a point in the center of the outlet will maintain its relative position. This consistency allows for the construction of a covariance matrix to identify the relationships and correlations between each point and itself, as well as with respect to the other points for each different geometry. Consequently, POD can be performed regardless of the mesh used for the CFD simulation, providing the flexibility to make significant modifications to the geometrical features of the turbopump design.

2.5 POD reconstruction

The utilization of Proper Orthogonal Decomposition in this study aims to reconstruct the flow field for turbopump geometries that have not been analyzed through CFD. This is achieved by constructing a POD model from high-fidelity full-order model or Computational Fluid Dynamics data. The simulations used to infer the POD/reduced order model are referred to as snapshots. Examining Equation 2.12, which is reiterated here for clarity:

$$U = \begin{pmatrix} u_{11} & \dots & u_{1n} \\ u_{21} & \dots & u_{2n} \\ \vdots & & \vdots \\ u_{m1} & \dots & u_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \phi_{11} & \dots & \phi_{1n} \\ \phi_{21} & \dots & \phi_{2n} \\ \vdots & & \vdots \\ \phi_{n1} & \dots & \phi_{nn} \end{pmatrix}. \tag{2.25}$$

The A matrix can be expressed as follows for the general scenario:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ u_{m1} & \dots & u_{mn} \end{pmatrix} = \begin{pmatrix} \tilde{a}(\mu_1) \\ \tilde{a}(\mu_2) \\ \vdots \\ \tilde{a}(\mu_m) \end{pmatrix}. \tag{2.26}$$

Each array $\tilde{a}$ comprises n coefficients representing the amplitudes for each mode necessary to reconstruct the full flow field information of each snapshot. These coefficients characterize each configuration μ_n. The interpolation methods can be used to predict the amplitude coefficients for a new configuration by leveraging the inherent relationships between inputs (parameters characterizing the simulation) and outputs (amplitude coefficients). Assuming that each configuration is correlated with modifications of multiple (q) parameters $\mu_n = [\mu_{n,1}, \mu_{n,2}, \dots, \mu_{n,q}]$.

The data for a specific snapshot, correspond to each row of the matrix U. While the eigenvectors or modes remain the same for each snapshot. The modal coefficients, represented by the matrix A, act as amplitudes required for reconstructing the flow data. Selecting a single row allows for the reconstruction of the respective snapshot, enabling reconstruction of the n^{th} snapshot by considering the n^{th} row. This property of POD can be used to extrapolate the flow field for a solution not utilized in model construction, as demonstrated in various studies such as [14], [13], [21], [22], and [23]. In existing literature, various interpolation methods have been applied, ranging from linear and cubic interpolation to neural network and Gaussian processes.

In the present study, four different interpolation methods have been tested: linear and cubic interpolation, Kriging method, and Radial Basis Function (RBF). The application of these interpolation methods will be elucidated through the simplest case, linear interpolation.

In a scenario with nine simulations, the design space of the simulation, used to build the model, where each cross is characterized by different amplitude coefficients $\tilde{a}(\mu_n)$, is graphically depicted in Figure 2.12. A linear interpolation is applied to the coefficients for a straightforward case, where each configuration is characterized by the tuning of two parameters $\mu_n = [\mu_{n,1}, \mu_{n,2}]$. To calculate the amplitude coefficients that would characterize a new configuration $\tilde{a}(\mu_{\text{new}})$, the linear interpolation approach is graphically presented in Figure 2.13. The calculation is presented as the product of the value at the desired point (depicted in black), where the product of the value at the desired point and the entire area is equivalent to the sum of the products of the value at each corner and the partial area diagonally opposite the corner (represented by the corresponding colors).

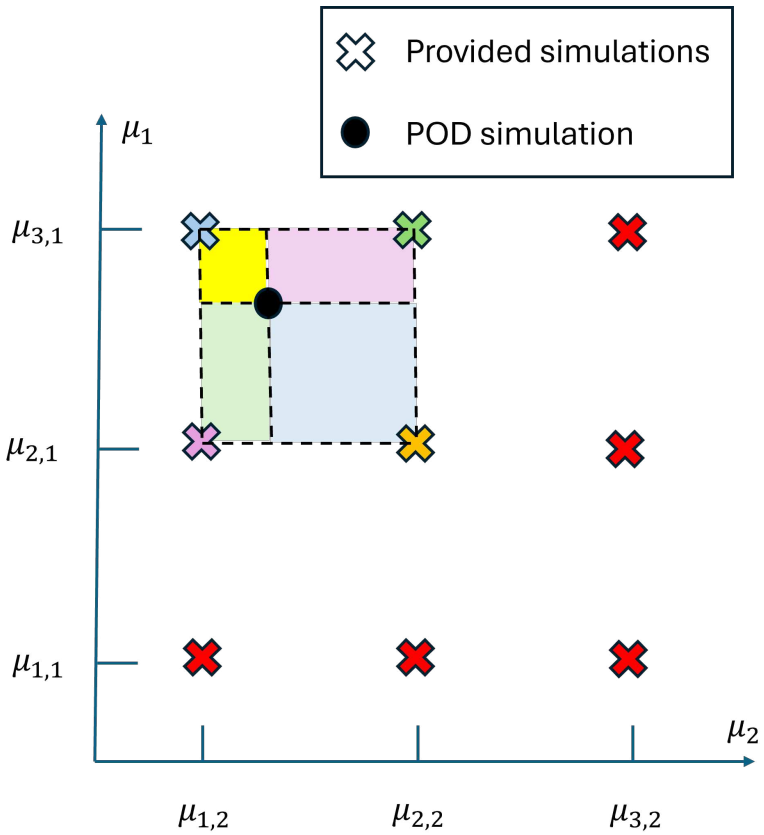


Figure 2.12: The graphical representation illustrates the design space of nine simulations employed for constructing a POD model, as well as the ongoing simulation being evaluated through the constructed model. Each simulation is distinguished by a specific combination of two parameters.

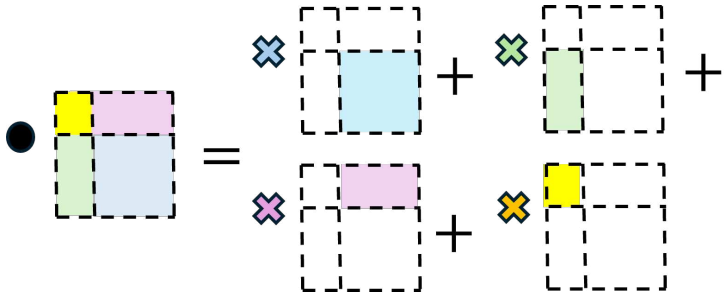


Figure 2.13: The visualization depicts bilinear interpolation geometrically.

Linear interpolation involves calculating the value at a target node or element by computing a weighted sum of contributions from neighboring nodes or elements on the source mesh. This method assumes a linear relationship between the known data points, meaning that the value between two adjacent points changes linearly with distance. As the distance from a known data point increases, its influence on the interpolated value decreases linearly. This approach provides a first-order approximation of the unknown value, making it relatively simple to implement computationally. This approach was followed by Wei et al. [13]. However, linear interpolation presents some limitations. It may struggle to accurately capture nonlinear relationships or abrupt changes in the data, as it assumes a constant rate of change between adjacent points. Consequently, while linear interpolation is useful for estimating values between known data

points, it may not always provide accurate results in scenarios with complex or nonlinear variations. The `MATLAB` function `griddata` was used, specifying as an interpolation method 'linear'.

Cubic interpolation is a method used to estimate the value at a point within the range of known data points based on the assumption of a smooth curve between these points. Unlike linear interpolation, which assumes a linear relationship, cubic interpolation approximates the data with a cubic polynomial function. Applying cubic interpolation, the value at the target point is determined by fitting a cubic polynomial curve to the nearby data points. This curve passes through the known data points and smoothly connects them, resulting in a more accurate estimation compared to linear interpolation, especially when the relationship between data points is nonlinear. Cubic interpolation offers a higher degree of precision than linear interpolation and can better capture complex variations and subtle changes in the data. However, it is computationally more intensive due to the higher-order polynomial involved in the interpolation process. Despite this, cubic interpolation remains a popular choice for applications where accuracy and smoothness of the interpolated values are crucial. Also for this case the function `griddata` has been used, specifying for the interpolation method 'cubic'.

The third interpolation method that was used is Kriging, which is a method of spatial interpolation that originated in the field of mining geology. The `ooDACE Toolbox` developed at the University of Gent [28], was utilized for this work. In particular the function `oodacefit` has been used, which implements the blind Kriging method, in `MATLAB` in accordance with [29].

Finally the `MATLAB` function `newrbe` has been tested, which is used for creating a Radial Basis Function network, a type of artificial neural network. This function trains the RBF network using a set of input-output training data pairs.

2.6 POD model construction

The construction of the Proper Orthogonal Decomposition model is a critical step in this study, particularly regarding the simulations used to train the model. Although not extensively discussed in the dissertation, various techniques can be employed to determine the placement of simulations within the design space. These techniques, known as space filling techniques, can significantly influence the quality of the final results and are related to design of experiment analysis.

The Optimal-Space Filling (OSF) design technique [30], previously utilized by Zakerholseini, was adopted to generate the initial set of simulations for this study. This technique was deemed appropriate for ensuring a comprehensive exploration of the design space. Additional simulations were incorporated according to this method. However, it is important to note that some CFD simulations did not converge and were subsequently excluded from the POD model construction.

A recap of the process necessary to construct and apply the POD model is presented in Figure 2.14. The process is divided into four steps: data extraction, preprocessing, processing, and reconstruction. This workflow provides a clearer explanation of how all the explained steps fit together.

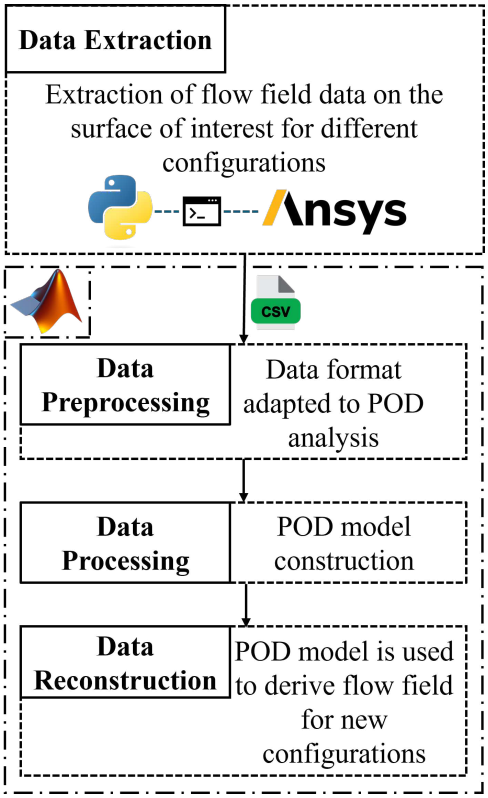


Figure 2.14: Workflow of the POD model construction and use.

2.7 Pump performance evaluation through POD

The application of Proper Orthogonal Decomposition (POD) to reconstruct the flow field for specific surfaces is used for evaluating pump performance. A model containing this information is built to perform the analysis. The primary performance indicators for a pump are the head (H) and total efficiency (η_{tot}). The pump head is defined as:

$$H = \frac{P_{t,\text{out}} - P_{t,\text{in}}}{\rho g}. \tag{2.27}$$

Here, P_t denotes the total pressure at the inlet and outlet of the turbopump, particularly the mass-averaged values on these surfaces, ρ represents the density of the working fluid (R245fa in this case), and g is the acceleration due to gravity. To compute the head using POD simulations, total pressure and mass flow fields on the inlet and outlet surfaces are required. The total efficiency is defined as:

$$\eta_{\text{tot}} = \frac{\dot{m}_{\text{in}} H g}{P_{\text{shaft}}}. \tag{2.28}$$

The shaft power (P_{shaft}) must be calculated to determine the total efficiency, which cannot be derived solely from the inlet and outlet surfaces. Therefore, blade surface data becomes essential. The shaft power is defined as:

$$P_{\text{shaft}} = \sum_{pp=i,j,k} (\text{Velocity in Stationary Frame} \cdot \vec{pp}) \times (\text{Force} \cdot \vec{pp}), \tag{2.29}$$

where i, j, and k denote the three main directions (X, Y, and Z), and the force (Force) is given by:

$$\text{Force} = P_t \times \text{Area}. \tag{2.30}$$

These calculations are performed for each point of the spatial discretization. To enable such calculations, a POD on the blade surfaces is conducted for the pressure, velocity in the stationary frame of reference, and normal vectors. In the analysis of pressure and velocity fields, the projection of the flow field, as detailed in subsection 2.4.3, is initially utilized, followed by the application of POD.

To predict the area of both the main and splitter blades, the size of each cell is considered, leveraging the fact that all polylines are equidistant along the height due to their geometric definitions. This area prediction is accomplished using the blind Kriging method to calculate the overall area of both the main and splitter blades and the geometric information to calculate the area of each individual cell. For calculating normal vectors for each cell, a consistent mesh discretization is ensured by controlling the number of cells per simulation. The normal vectors are computed through triangulation formed by the cell center, the adjacent point on the same polyline, and the corresponding point on the polyline above. This involves applying the cross-product between the vector from the center to the previous point and the vector from the center to the point above. These geometric calculations, combined with the measurements of pressure and velocity fields in the stationary frame, are essential for accurately calculating the power output of the shaft. It is noteworthy that the total power transferred to the shaft should include the forces induced by shear stresses on the blades, hub, and shroud. However, after analyzing their impact on overall power calculations, it was found that they contribute to a maximum of 5 % of the total power transferred to the shaft, so it was decided to neglect them.

From the flow field obtained via POD simulations, additional performance markers can be derived. These include the minimum pressure the blade withstands to assess cavitation risk, which is evaluated through the Net Positive Suction Head (NPSH) in subsection 3.4.1. Another important marker is the pressure difference on the pressure side of the main blade between the hub and shroud. This pressure difference often causes secondary flows in turbomachinery bulk flow, leading to losses in pumps.

A visual representation of how the model is built starting from the POD model based on specific surfaces is provided in Figure 2.15. The overall model is constructed by performing the POD for four different regions of the turbopump: inlet, outlet, main, and splitter blades. For the inlet and outlet, two separate PODs are performed on the total pressure and mass flow fields. For the main and splitter blades, the POD is applied to the pressure and velocity fields. Additionally, the geometrical information of the main and splitter blades, such as area and normal vectors, are necessary.

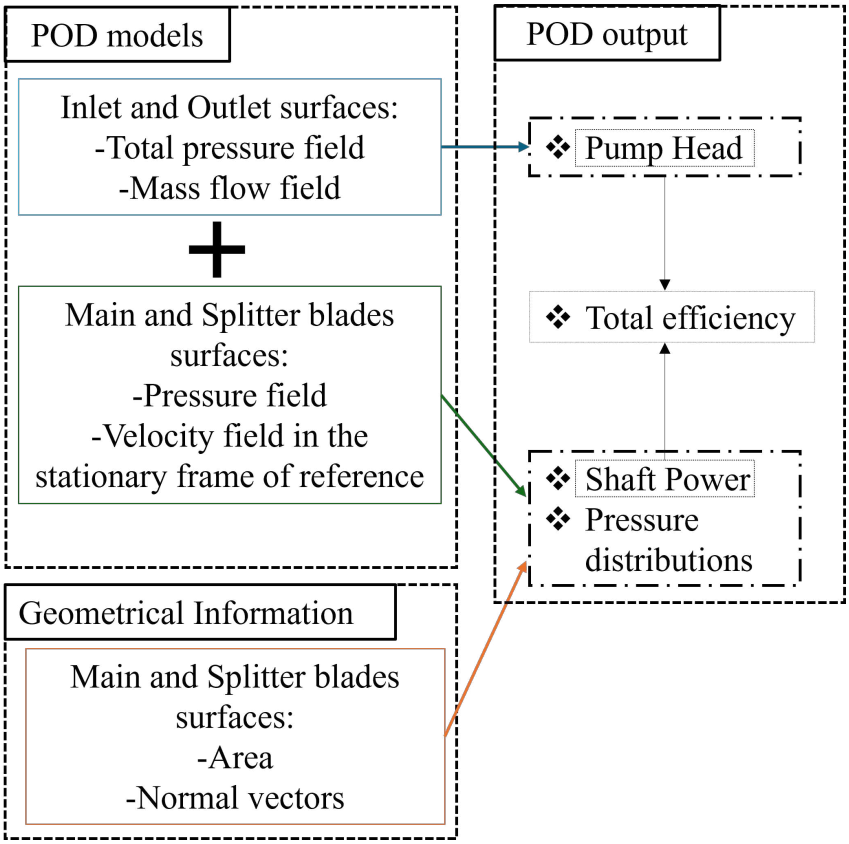


Figure 2.15: POD model configuration and use in predicting the turbopumps performance.

Chapter 3

Results

3.1 Surface flow field POD

The initial assessment of the POD capabilities involved a constrained dataset, consisting of a limited number of simulations and focusing solely on the variation of two parameters across these simulations. This constrained approach was adopted to facilitate a clearer understanding of the limitations inherent in the POD model and to gain deeper insights into the obtained results. The chosen design space encompasses adjustments to the splitter pitch and leading-edge positions, as these parameters are pivotal in defining the turbopump's performance.

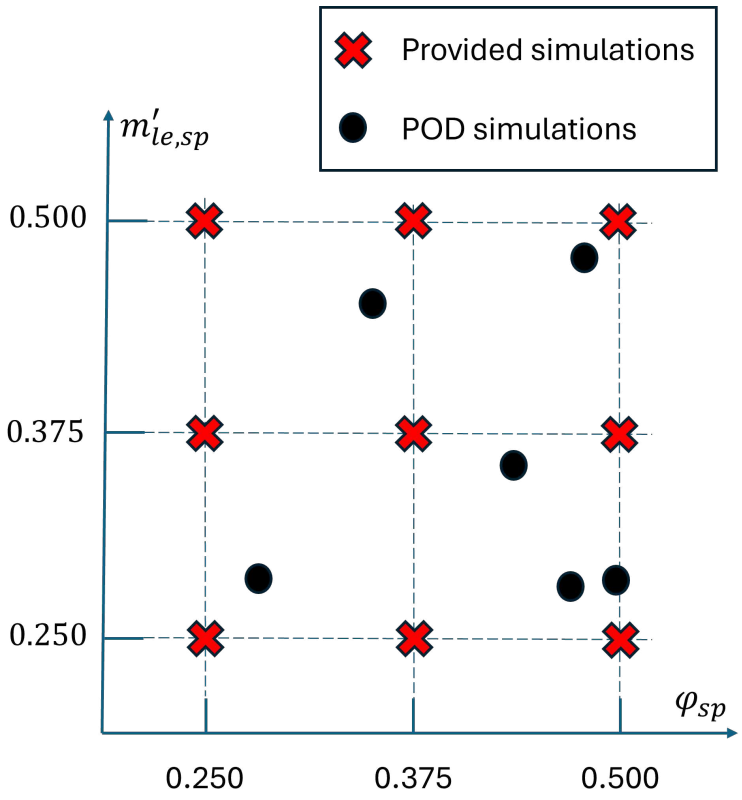


Figure 3.1: Design space for building the first POD model of the turbopump.

The visual depiction of the design space is presented in Figure 3.1. Here, nine CFD simulations were employed to construct the POD model. Six random geometrical configurations, inside the design space, have been selected to serve as benchmarks for assessing the model's performance. The simulations for building the POD model were based on a turbopump designed to operate at a rotational speed of 25 krpm, featuring a shrouded four-blade configuration with an outlet angle of 90° and operated at a rotational speed of 40 krpm. The POD model utilizes the Kriging method to predict modal coefficients, as this method has proven to be the most accurate. A short discussion on the performance of the various interpolation methods can be found in section 4.3.

Parameter	S1	S2	S3	S4	S5	S6
ψ_{sp}	0.275	0.300	0.500	0.460	0.400	0.470
$m'_{le,sp}$	0.275	0.450	0.275	0.260	0.350	0.470

Table 3.1: Parameters of the simulations used to test the first POD model (S is the abbreviation of Simulation).

3.1.1 Blade to blade surface

The POD model was firstly tested in the prediction of the velocity field on a blade to blade surface at a normalized span level (S_N) of 0.5, where the length and position of the splitter blades is being tuned. After the application of the POD, nine modes have been identified. The energy (TKE) carried by each mode is displayed in Figure 3.2.

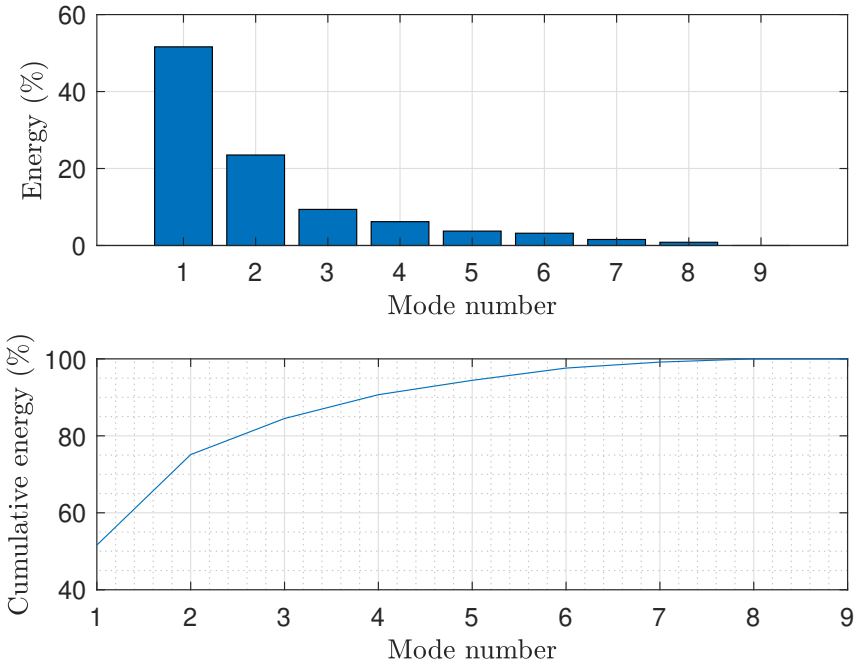


Figure 3.2: Energy distribution of POD eigenmode velocity fields in the blade to blade surface. In the upper subplot of the figure, the energy distribution across each mode is depicted. The lower subplot illustrates the cumulative energy.

The first mode contains half of the total energy in the flow data, whereas the ninth mode contributes approximately 0 % to the total energy, making it the least significant. The first six modes identified through the POD analysis account for 97 % of the total energy, making them suitable for reconstructing flow information for new parameter combinations. An error analysis, shown in Figure 3.3, involves cumulatively adding modes and evaluating the resulting accuracy. It becomes clear that after incorporating the first six modes, the improvements in accuracy are negligible. The mean relative error is averaged across the six simulations used to validate the quality of the POD model.

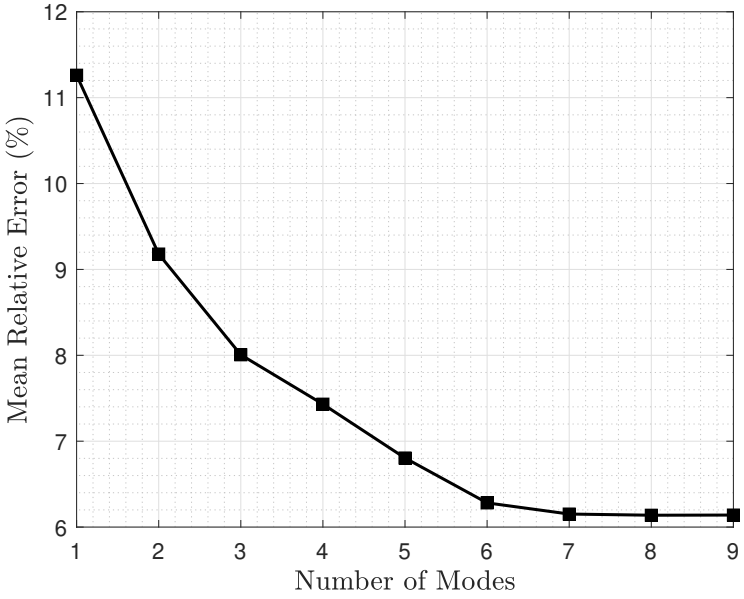


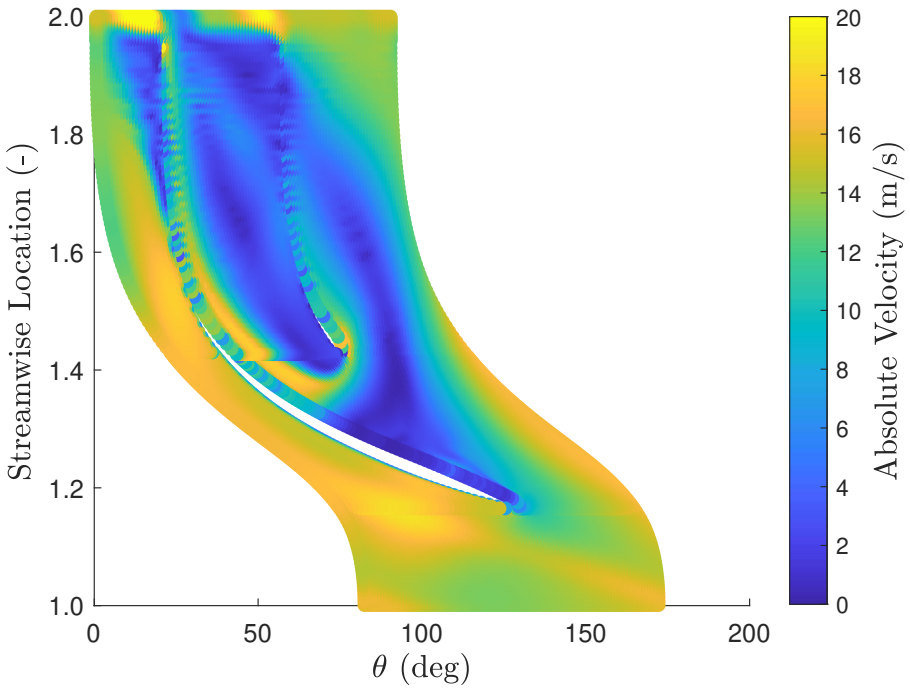
Figure 3.3: Mean relative error for the velocity field prediction of the six simulations used to test the first POD model presented in Table 3.1, with respect to mode addition.

All nine modes were utilized, however, as their inclusion does not noticeably impact computational performance. It's worth noting that predicting the flow field for the required simulation takes only half a second on a standard workstation with an Intel(R) Core(TM) i7-10870H CPU running at 2.20 GHz and equipped with 16 GB of RAM, which is common for a standard laptop in 2024.

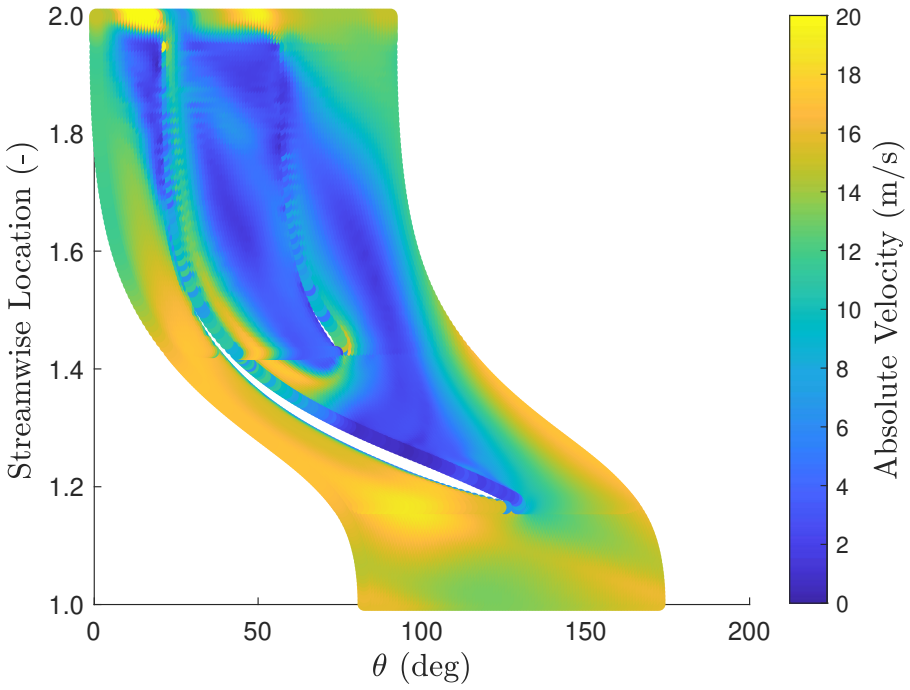
A visual comparison of the predicted velocity field for the fifth simulation (S5, Table 3.1) is depicted in Figure 3.4, while a quantitative comparison is provided in Figure 3.5, illustrating the relative error between the CFD and POD results across the entire domain. In the visualization of the original CFD flow field and the predicted POD flow field, steep changes in velocity can be observed along the boundaries of the blades. These changes are due to numerical errors in the interpolation of the velocities when projected onto the new mesh. The original mesh is highly clustered around the blade edges, and is characterized by steep changes in the velocity. These discrepancies arise due to the interpolation function used in `MATLAB` (`griddata`), which struggles to accurately interpolate the values in these regions. This aspect does not ruin the quality of the POD in predicting the flow field, but suggest the necessity of checking the quality of the projection of the flow field on the new mesh.

The relative error for the velocity field (RE_v) is calculated as:

$$RE_v = \frac{|\mathrm{POD}_{v,i} - \mathrm{CFD}_{v,i}|}{\overline{\mathrm{CFD}}_v}, \tag{3.1}$$



(a) CFD results for the velocity flow field at $S_N = 0.5$.



(b) POD results for the velocity flow field at $S_N = 0.5$.

Figure 3.4: Visual comparison of the CFD and POD results ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

where $\overline{\text{CFD}}_v$ denotes the mean velocity obtained from the CFD simulations in the plane of interest:

$$\overline{\text{CFD}}_v = \frac{1}{n}\sum_{i}^{n}(\text{CFD}_{v,i}). \tag{3.2}$$

This normalization is done to prevent the relative error from being excessively influenced by division by values close to zero. If the denominator would be the instantaneous velocity instead of the mean velocity, it could potentially bias the interpretation of the results quality.

While examining the results, it's important to note that in certain regions, the relative error may approach values of 100 %. This is primarily due to the presence of close-to-zero values in the CFD simulation data at those locations. The POD, by its nature, tends to smooth out such outliers. The mean relative error for the velocity field is approximately 6 %, which, given the low number of simulations provided to the POD model, can be considered a good result. The same analysis can be done for the pressure field, and the results are presented in section A.2. However, it should be emphasized that given the higher uniformity of the flow field, the POD showed better performance with a mean relative error of approximately 3 %.

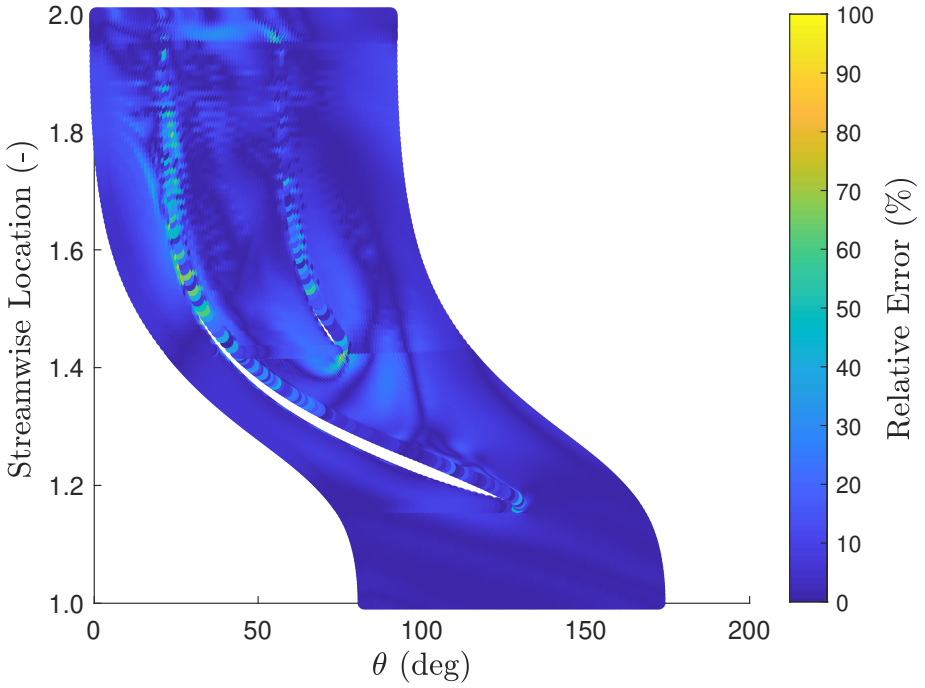


Figure 3.5: Relative error between CFD results and POD results for the velocity field on the blade to blade surface at $S_N = 0.5$ ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

Error analysis

In Table 3.2, a detailed error analysis for the velocity field predictions of the blade to blade surface using POD across six different simulations is presented. The table summarizes the median relative error, the 25^{th} and 75^{th} percentiles, and the proportion of outliers, defined as errors exceeding the upper whisker[1], calculated as 1.5 times the interquartile range above the 75^{th} percentile and below the 25^{th} percentile. This range between the first and third quartiles illustrates the middle 50 % of values, providing a measure of the spread and variability in the data.

[1] maximum and minimum values that are not considered outliers

Parameter	S1	S2	S3	S4	S5	S6	Overall
Median (%)	3.4	5.9	4.5	6.7	4.9	3.8	4.9
25^{th} percentile (%)	1.5	1.6	1.3	2.0	1.8	1.3	1.6
75^{th} percentile (%)	9.3	14.1	10.3	13.9	9.1	9.1	11.0
Outliers (%)	8.5	3.4	4.8	4.1	5.9	4.5	5.2

Table 3.2: Error analysis on the velocity field prediction of the first blade to blade surface POD (S is the abbreviation of Simulation).

To visually complete this analysis, Figure 3.6 illustrates the performance of the POD predictions for the six simulations concerning the prediction of the blade to blade surface velocity field. This box plot visually highlights the distribution of errors and identifies any extreme values or trends across simulations.

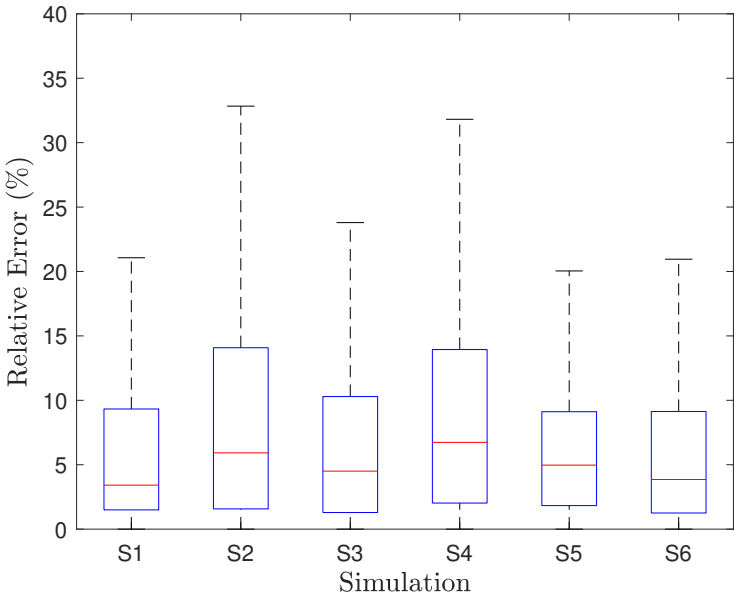


Figure 3.6: Box plot of the relative error for the six simulations used to evaluate the POD model's prediction quality for the velocity field on the blade-to-blade surface.

3.1.2 Inlet and outlet surfaces

The design space depicted in Figure 3.1 was used to construct two separate POD models for reconstructing the flow field information at both the inlet and outlet sections of the turbopump. This approach allows for targeted analysis of each section. The focus of this analysis is on the total pressure within a stationary frame of reference, which is vital for accurately assessing pump performance, especially in evaluating the pumps head.

The energy distribution plots, which provide detailed insights into the energy characteristics at both the inlet and outlet, are included in the Appendix. Specifically, the plot for the inlet POD model is shown in Figure A.5**(a)**, and the plot for the outlet POD model is in Figure A.5**(b)**.

As illustrated in Figure A.5**(a)**, approximately 90 % of the energy is concentrated by the first mode at the inlet, indicating that the remaining modes are significantly less relevant. This

predominance can be attributed to the boundary conditions set during the CFD simulation, specifically the mass flow definition at the inlet, which results in a relatively uniform flow field across all geometries, dominated by a single primary mode.

Conversely, in Figure A.5**(b)**, the energy distribution at the outlet surface is more dispersed, with the first three modes together accounting for roughly 90% of the total energy. This suggests a more complex flow behavior at the outlet compared to the inlet, where multiple modes play a substantial role in the dynamics of the flow. This complexity is a direct consequence of the fluid flowing from the inlet to the outlet and being influenced by geometrical changes along the path.

A comparison of CFD and POD results for a specific geometrical configuration is shown for the inlet surface in Figure A.6, located in the Appendix. Visually, the pressure fields produced by the CFD simulations and those reconstructed using POD are indistinguishable. This similarity is quantitatively supported by the analysis in Figure 3.7, where the relative error is detailed. The maximum error recorded is 2.7 %, with the mean relative error remaining exceptionally low, below 1 %. These results affirm the high accuracy of the POD approach, while also highlighting its efficiency with computational times of less than a second required for this analysis.

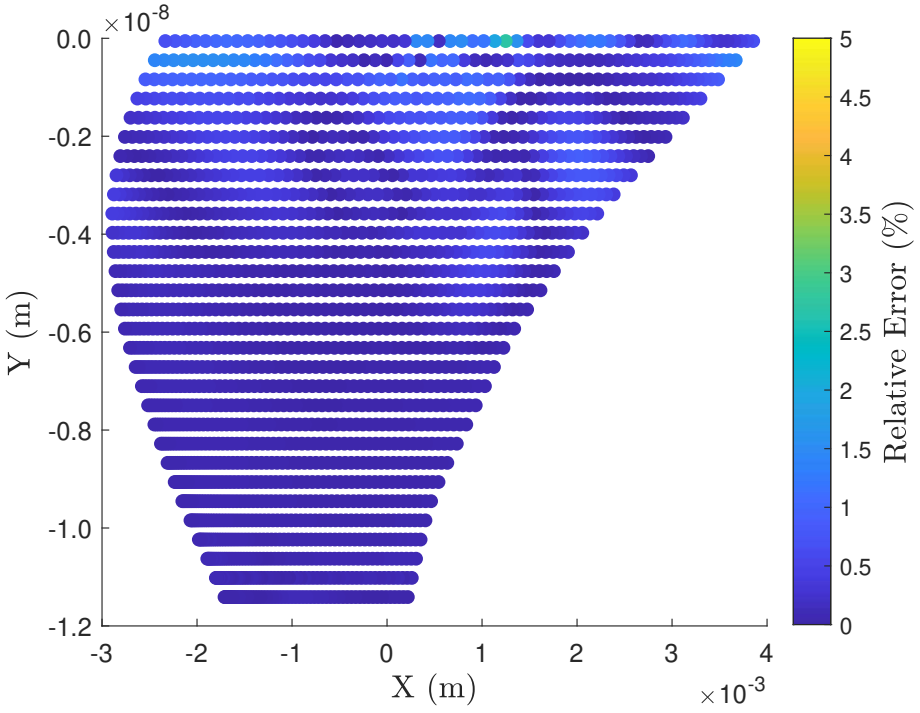


Figure 3.7: Relative error between CFD results and POD results of the inlet surface for the total pressure field ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

Similar conclusions regarding the accuracy of simulations can be drawn for the outlet surface. Figure A.7 offers a visual comparison of the pressure fields from both CFD and POD results. The differences between these simulations are quantified in Figure 3.8, which shows that the maximum error values reach approximately 8 %, while the mean relative error is around 2 %.

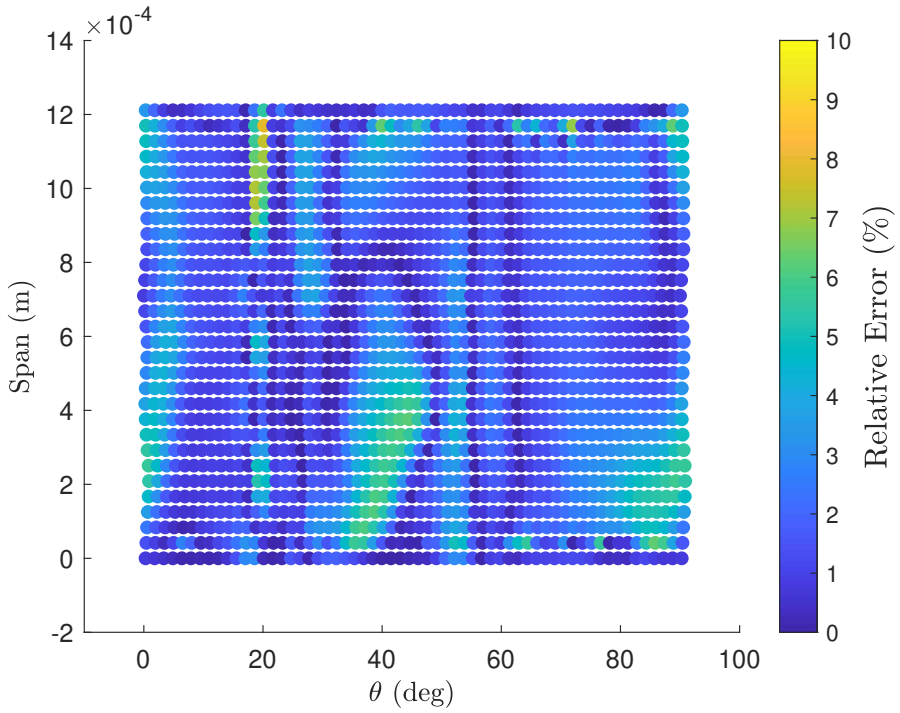


Figure 3.8: Relative error between CFD results and POD results of the outlet surface for the total pressure field ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

Error analysis

All six simulations used to test the POD model quality demonstrated high-quality results. Consequently, it was decided, not to present these results in a tabular format as done previously in Table 3.2. Instead, a box plot is used to visually summarize the findings for the inlet surface, as shown in Figure 3.9.

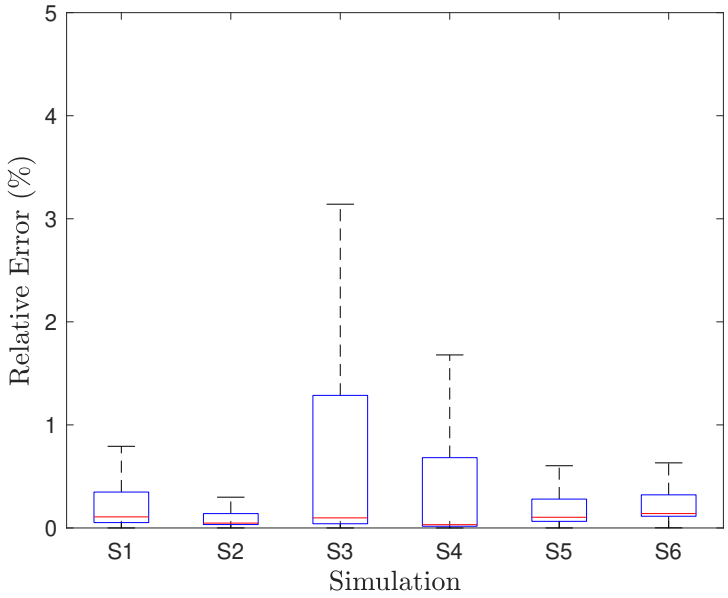


Figure 3.9: Box plot of the relative error for the six simulations used to evaluate the POD model's prediction quality for the total pressure field on the inlet surface.

For the outlet simulations, the quality of the POD in predicting total pressure within the stationary frame of reference was evaluated. Statistically significant indicators of relative error are presented in Table 3.3 and visually summarized through a box plot in Figure 3.10. The overall capability of the POD model in predicting the flow field for the inlet and outlet surfaces is shown to be of very high quality.

Parameter	S1	S2	S3	S4	S5	S6	Overall
Median (%)	1.5	2.4	5.5	3.7	1.1	2.9	2.9
25^{th} percentile (%)	0.5	0.8	2.4	1.5	0.5	1.1	1.1
75^{th} percentile (%)	4.5	6.7	10.8	5.8	2.7	5.1	5.9
Outliers (%)	5.2	8.2	1.4	1.8	2.3	7.6	4.1

Table 3.3: Error analysis of the first POD model for the total pressure field on the outlet surface (S is the abbreviation of Simulation).

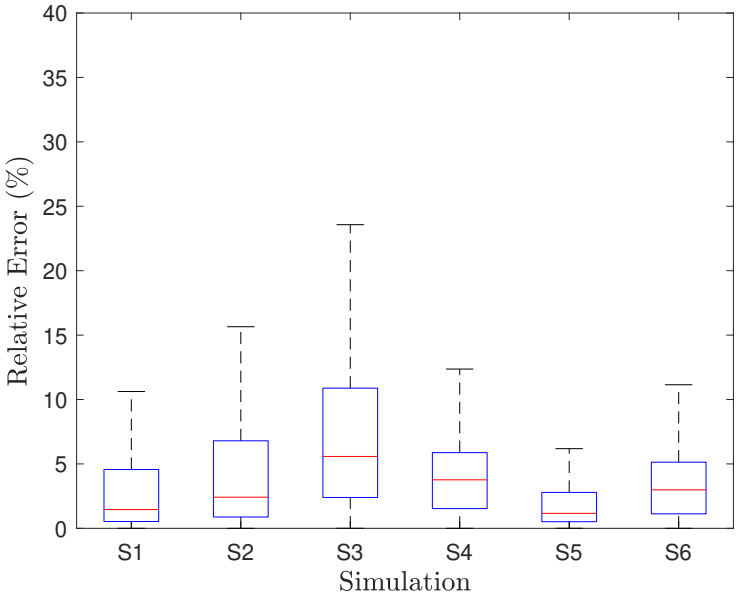


Figure 3.10: Box plot of the relative error for the six simulations used to evaluate the POD model's prediction quality for the total pressure field on the outlet surface.

3.1.3 Main and splitter blade surfaces

Finally, two POD models aimed at predicting the pressure field distribution on the main and splitter blade surfaces were constructed. This involved building a snapshot matrix where various polylines from the hub to the shroud were stacked (40 for the current case), ensuring a thorough discretization of the entire blade surface. Only the energy distribution across the modes and the relative error are presented. The energy distribution are depicted in Figure A.8. While, the relative errors are presented in Figure 3.11 and Figure 3.12, for the main and splitter blades respectively.

In the main blade POD analysis, nearly 60 % of the field energy is concentrated in the first mode and the first three modes would carry 90 % of the total field energy. From the quantitative results of the relative error for one specific configuration, Figure 3.11, the POD reconstruction exhibits higher relative errors on the pressure side of the blade. In this position the fluid experiences higher acceleration due to the blade turning, resulting in a pressure drop. Nevertheless, the maximum relative error is approximately 15 %, and the mean error across the entire blade surface is around 1.7 %, indicating that the POD predictions provide satisfactory results also for a 3D geometry as for the current case.

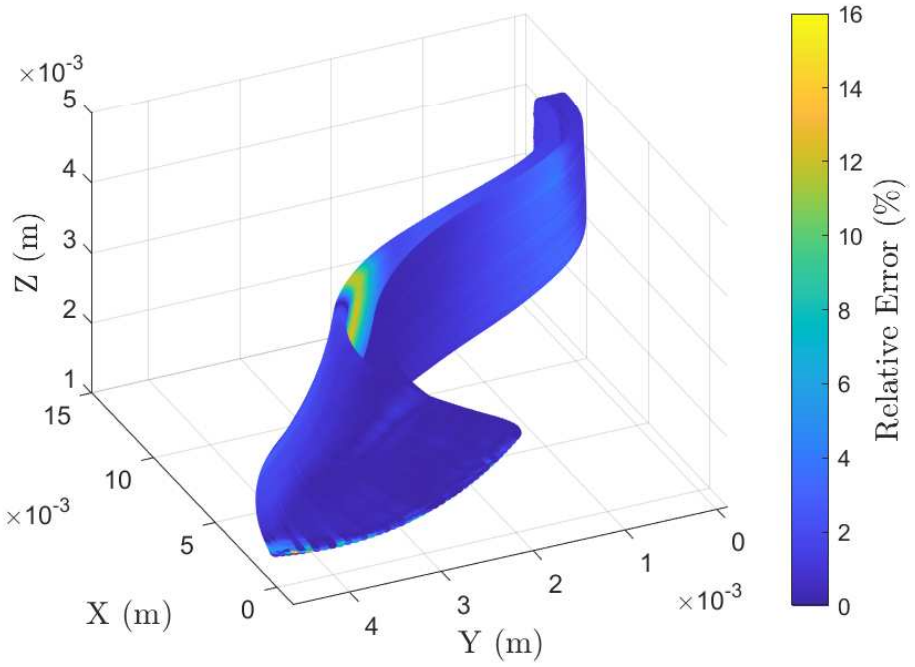


Figure 3.11: Relative error on the pressure field prediction between CFD and POD results of the total pressure field on the main blade surface($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

Next, the same analysis is conducted and presented for the splitter blade. The mode ranking based on energy considerations is presented in Figure A.9, and the relative error of the prediction for the pressure field on the splitter blade is depicted in Figure 3.12. It was found that almost 100 % of the energy is carried by the first mode. Regarding the performance of the POD, the only area that was found to be concerning is the leading edge of the splitter blade, where the maximum relative error is approximately 13 %, and overall the mean error is 2.6 %.

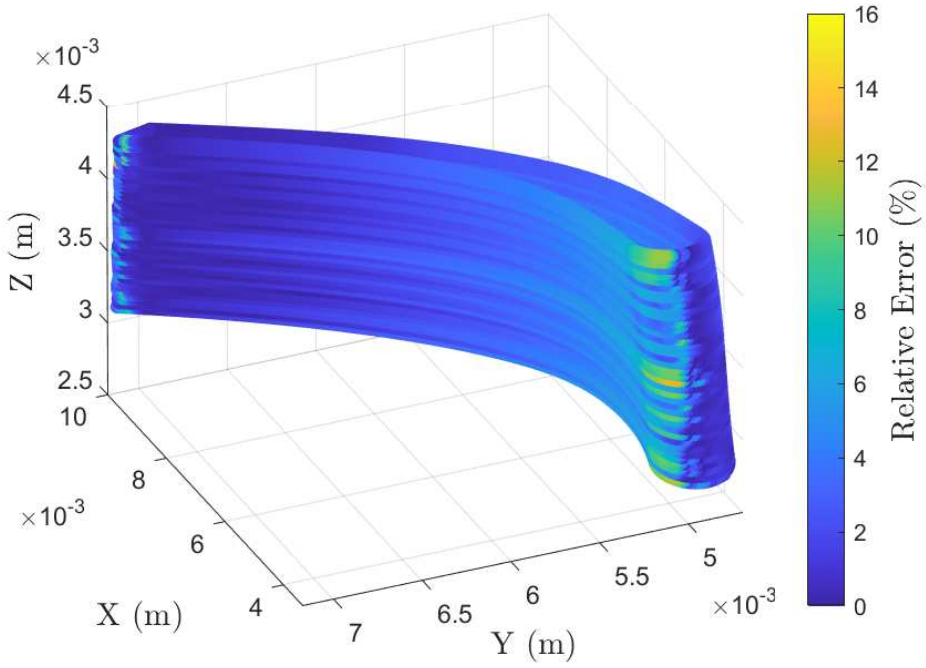


Figure 3.12: Relative error on the pressure field prediction between CFD and POD results of the total pressure field on the splitter blade surface($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

Error analysis

The six simulations initially used to test the quality of the POD are employed again to analyze the POD prediction for the flow on the main blade. Pertinent error statistics for this part of the study are detailed in Table 3.4, and a visual representation of the POD reconstruction performance for the main blade is shown in Figure 3.13. The results demonstrate high quality, considering the minimal computational demand required.

Parameter	S1	S2	S3	S4	S5	S6	Overall
Median (%)	1.5	1.7	0.9	1.4	1.7	0.7	1.3
25^{th} percentile (%)	0.4	0.4	0.3	0.6	0.3	0.2	0.4
75^{th} percentile (%)	2.6	2.3	1.4	2.4	2.4	1.3	2.1
Outliers (%)	7.7	2.7	≈ 0	≈ 0	3.8	3.3	2.9

Table 3.4: Error analysis of the first POD model for the total pressure field on main blade surface (S is the abbreviation of Simulation).

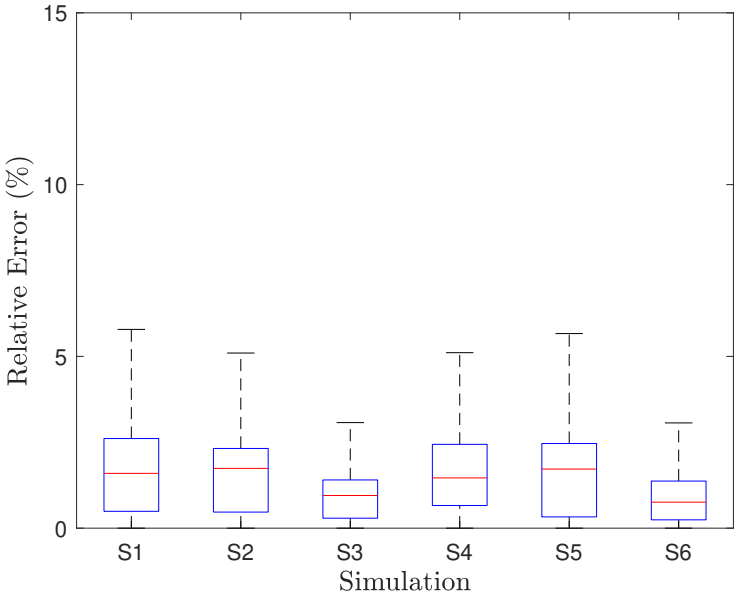


Figure 3.13: Box plot of the relative error for the six simulations used to check the POD model quality for the main blade surface.

For the splitter blade, a similar analysis is conducted, with findings presented in Table 3.5 and Figure 3.14. It is observed that the error is slightly greater for the splitter blade than for the main blade, even though only one mode predominantly carries the total data energy for both components. This difference can be attributed to how the interpolation of the coefficients affects the results. Inaccurate predictions of the amplitude of the first mode can degrade the results. However, these outcomes are still deemed acceptable given the overall context of the study.

Parameter	S1	S2	S3	S4	S5	S6	Overall
Median (%)	2.8	3.0	1.6	2.3	2.2	1.1	2.2
25^{th} percentile (%)	1.6	2.2	1.1	1.2	1.2	0.6	1.3
75^{th} percentile (%)	3.9	3.4	2.2	3.4	3.2	1.5	2.9
Outliers (%)	3.2	7.1	3.0	1.4	0.3	0.8	2.6

Table 3.5: Error analysis of the first POD model for the total pressure field on splitter blade surface (S is the abbreviation of Simulation).

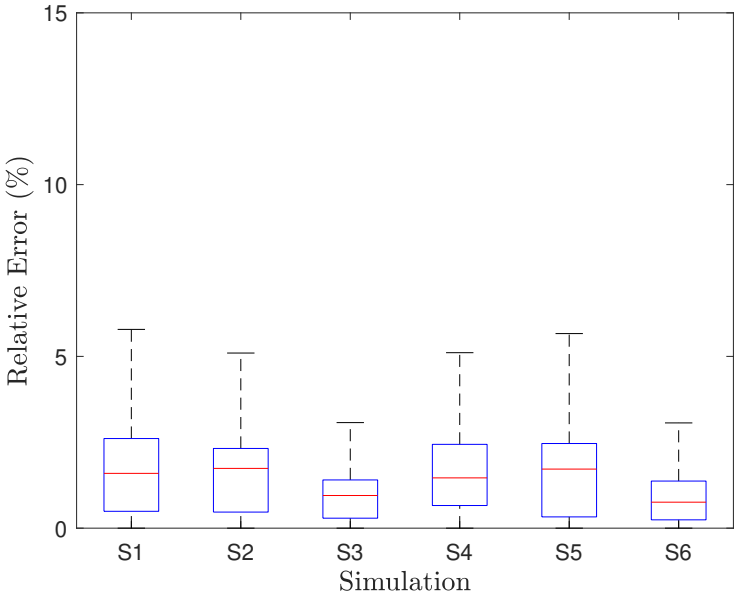


Figure 3.14: Box plot of the relative error for the six simulations used to check the POD model quality for the splitter blade surface.

3.2 Modal visualization

One of the features of Proper Orthogonal Decomposition (POD) is its capability to identify the main modes of the flow field. Therefore, the modes within the studied design space were analyzed. The following section presents the modes in the blade to blade surface, including both the modal coefficients and the corresponding spatial modes. In particular, the left-side figure shows the modal coefficients as a contour plot within the design space, while the right-side figure presents the spatial modes. Only the first three modes are reported in Figure 3.15, with the remaining six modes presented in the appendix in Figure A.10 and Figure A.11. The energy associated with each mode is presented in Figure 3.2. In the right side of the figures, where the spatial modes are represented, some horizontal lines are visible. These lines are due to the overlapping of the meshing cells in `MATLAB` and are thus purely a visualization artifact.

The various modes are presented so that if one of the modes identified by the POD needs to be reconstructed by a specific simulation, it is possible to do so by choosing the modal coefficient corresponding to the analyzed configuration and multiplying it by the spatial mode. In this case, the contour plot of the modal coefficients has been identified using the built POD model, with the Kriging method to perform interpolations. From the spatial mode, it can be observed, for example, that in the first mode, the yellow and blue patches are opposite in sign and comparable in absolute value. If the spatial mode is multiplied by a modal coefficient, this either enhances or decreases both of them, but their signs will remain opposite. To reconstruct the full flow field, it is necessary to identify the correct mode, by multiplying the right amplitude to the spatial mode, for each configuration of the turbopump and sum them together. For example, when analyzing the third mode in Figure 3.15, it is possible to state that for a configuration with a short splitter blade close to the main blade, $m'_{le,sp} = 0.25$ and $\psi_{sp} = 0.25$, the flow near the main blade experiences deceleration, especially considering that the spatial mode is being multiplied by a negative modal coefficient. Conversely, for a configuration characterized by $m'_{le,sp} = 0.30$ and $\psi_{sp} = 0.45$, the flow on the main blade is further accelerated.

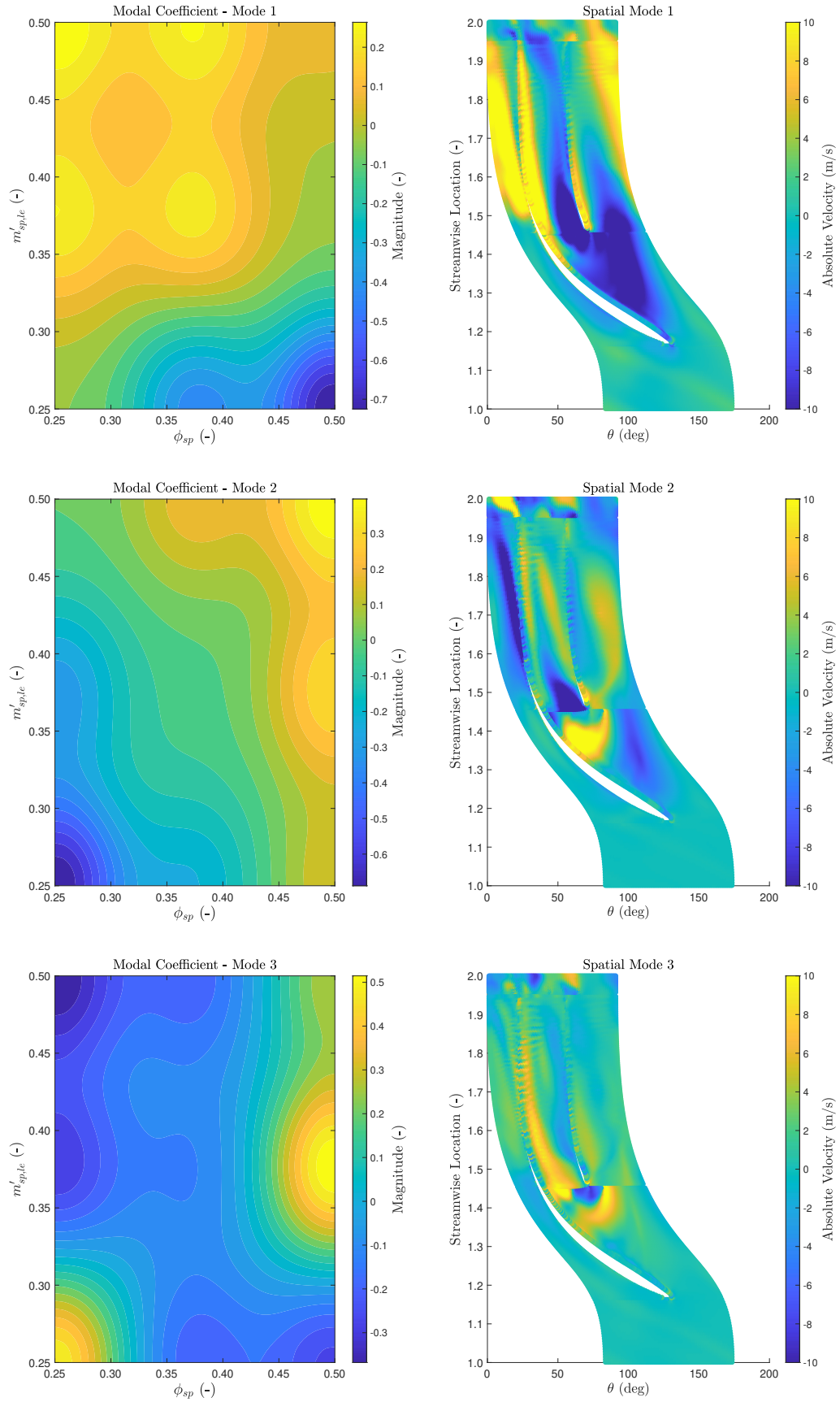


Figure 3.15: First three modes obtained from the first POD model, of the blade to blade surface for the velocity field.

Interpreting the various modes proved challenging. The usual approach from classical fluid mechanics in visually analyzing the flow is more difficult to apply, given that the presented modes are identified through a statistical analysis. This difficulty may be attributed to the nature of absolute velocity, which, despite being an absolute value, is still related to a vector field. The visualization of the spatial modes might provide more insight if applied to a scalar field or a vector field confined to the plane used for visualization.

To simplify the analysis, POD was applied while tuning only one parameter of the turbopump to observe how the weight of each mode changes according to the geometrical configuration. The outlet blade angle varied between 35° and 90°. The energy rankings are presented in Figure 3.16.

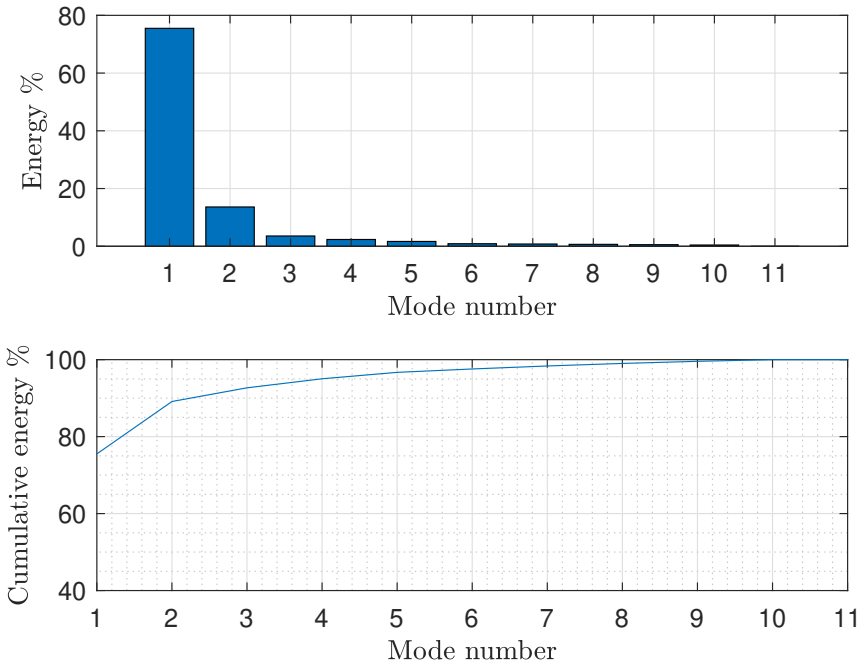


Figure 3.16: Energy distribution of POD eigenmode velocity fields in the blade to blade plane where only the outlet angle is being changed.

The visualization of the first three modes are presented in Figure 3.17. For the first mode, it is easier to identify a trend in the modal coefficients, especially for an outlet angle greater than 60°. In this range, the modal coefficient has an increasing trend and presents a positive value. Conversely, for an angle smaller than 60°, the trend is increasing in absolute value as the modal coefficient is negative. This observation also explains the trend of the second mode, where the modal coefficients present a minimum for a blade angle of 60°. This occurs because the identified modes must balance each other to reconstruct the full flow field accurately. In the third mode, the spatial mode is characteristic of one of the analyzed simulations, specifically the one with an outlet angle of 55°. Unfortunately, even in this case, while POD more clearly identifies the main modes in the flow field, the information provided is not easily applicable for enhancing the design process.

Coherent structures can be identified in the zones of correlation (yellow and blue patches in the spatial modes), especially when the corresponding energy is dominant compared to other modes. However, these structures do not provide clear indications of the flow dynamics. Nevertheless, POD can be very useful because it enables the reconstruction of a flow with only a few of the most energetic modes, and may allow to identify motions that are sometimes not easily spotted in the raw data.

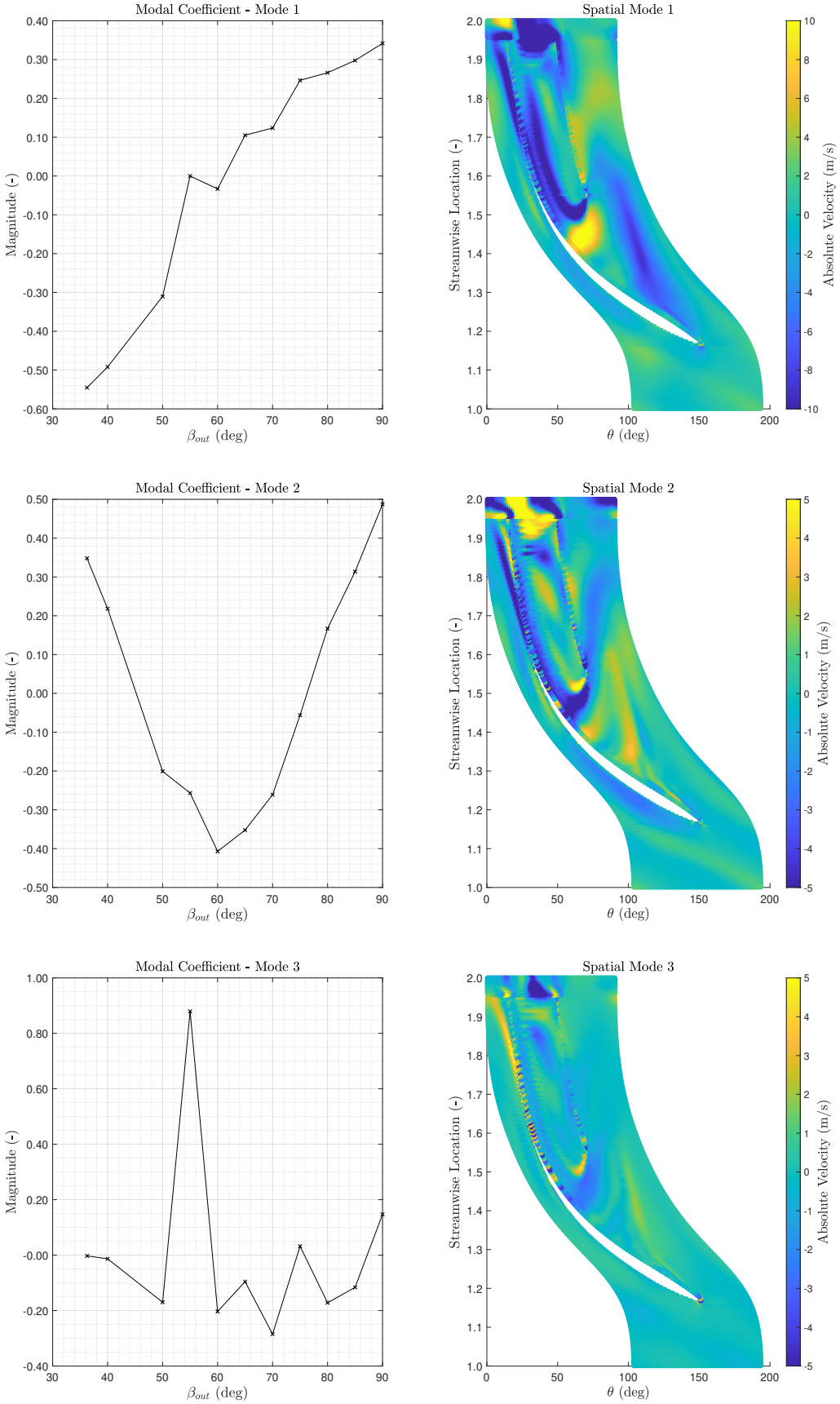


Figure 3.17: First three modes obtained from the POD model, of the blade to blade surface for the velocity field where the blade outlet angle is being tuned.

3.3 First POD model for pump performance prediction

After confirming the effectiveness of POD in reconstructing the flow fields across critical surfaces and verifying the model's quality, a new POD model is developed as outlined in section 2.7. This model aims to predict the pump performance for new geometrical configurations. The design space utilized, shown in Figure 3.1, focuses on adjustments to the position and length of the splitter blade.

The construction of the initial POD model took approximately 50 seconds, a duration expected to increase with the addition of more simulations. This time primarily accounts for the interpolation of various surface meshes and the projection of data necessary to build different snapshot matrices.

Once set up, the model efficiently computes critical performance metrics, such as shaft power, pump head, total efficiency, the lowest pressure on the blades, and the pressure difference between the hub and shroud on the pressure side of the main blade. These calculations take about 2 seconds for each specific configuration. The prediction quality is initially verified against the solutions used to construct the model, with the mean error and standard deviation detailed in the following table.

Parameter	Pump Head (%)	Shaft power (%)	Total efficiency (pt)
Mean error	1.34	1.14	0.22
Standard Deviation	0.84	0.44	0.34

Table 3.6: Mean error and standard deviation on the prediction of the main performance indicators for the simulations used to train the model.

The results for one specific configuration, identified as S5 from the configurations described in Table 3.1, are presented in the table below, to better visualize the accuracy of the results also in absolute terms:

	CFD	POD	
	Results	Results	Delta
Pump head (m)	205	206	✓
Shaft power (W)	592	611	3.2 %
Total efficiency (%)	95.37	92.5	2.7 pt

Table 3.7: Comparison of results obtained from the CFD and those obtained from the POD for configuration S5 in Table 3.1 ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

The pump head predictions are highly accurate, but the model tends to overpredict shaft power, which adversely affects the accuracy of the total efficiency. The primary source of this error lies in the inaccuracy of the POD's prediction of the pressure distribution on the pressure side of the main blade, where significant errors are observed, as depicted in Figure 3.11.

The error analysis for the six simulations used for the model testing is detailed in Table 3.8. The results indicate that both pump head and shaft power are predicted with good accuracy. However, issues emerge when evaluating total efficiency, where the compounded errors from both head and shaft power predictions are magnified. Despite these challenges, the results are

deemed satisfactory, particularly considering that they were obtained in approximately 2 seconds for each configuration. While it would require about 2000 times this amount of time to extract this information from a CFD simulation.

Parameter	S1	S2	S3	S4	S5	S6	Mean	Std. Dev.
Pump head RE (%)	4.2	3.6	1.8	1.6	0.1	1.6	2.2	1.4
Shaft power RE (%)	3.7	1.5	2.5	0.6	3.2	0.6	2.0	1.2
Total efficiency Δ (pt)	0.5	2.0	1.8	2.1	2.7	0.9	1.7	0.8

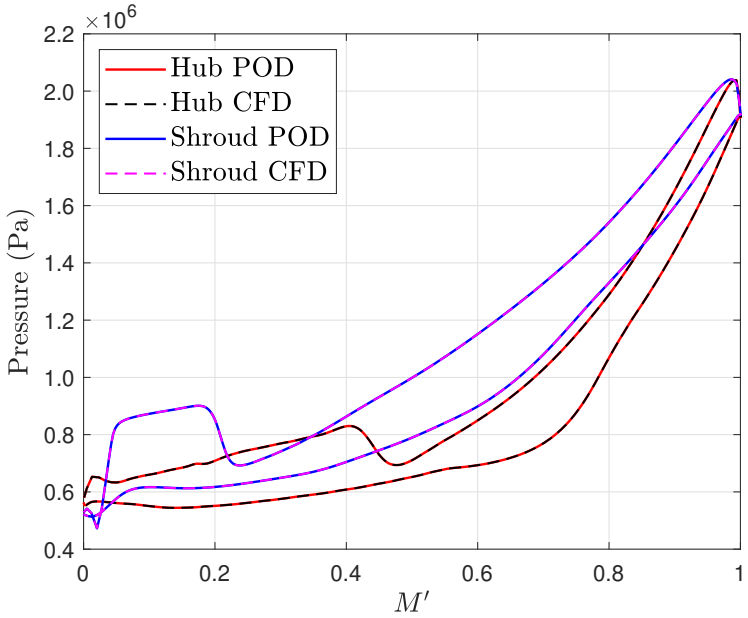
Table 3.8: Error analysis Performance POD (S is the abbreviation of Simulation).

The constructed POD model not only predicts basic performance metrics but also enables the extraction of additional indicators and deeper insights into pump operation. For example, Figure 3.18 illustrates the pressure distribution near the hub and shroud of the main blade, comparing data from a model training simulation and simulation S2 from Table 3.1. The training set simulation shows a high degree of accuracy, whereas the S2 simulation reveals minor errors, notably in the pressure drop caused by flow detachment.

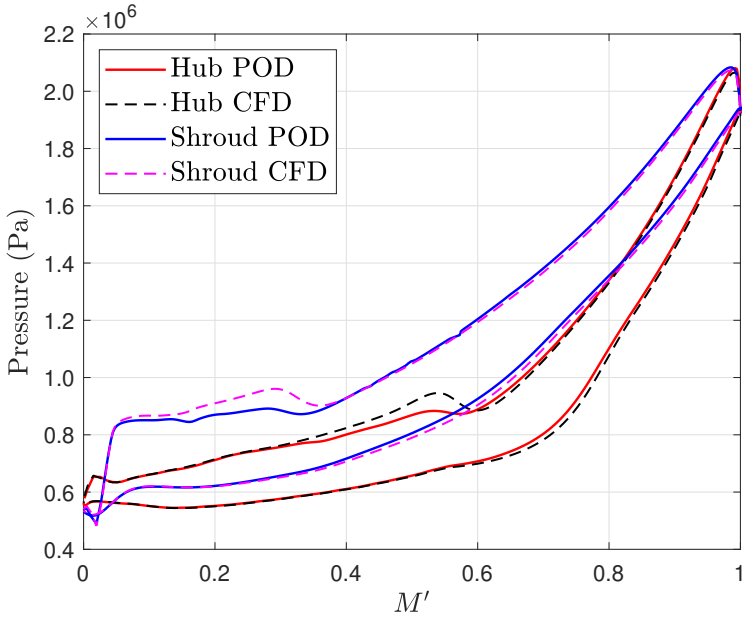
Furthermore, this model facilitates the assessment of potential cavitation by calculating the Net Positive Suction Head (NPSH) and analyzing the pressure differential between the hub and shroud. This differential can serve as a critical parameter to minimize in order to reduce secondary flows within the pump's bulk flow, enhancing efficiency.

The design space can be investigated performing 400 simulations, for different combinations of pitch positions of the splitter blade and meridional positions of the leading edge of the splitter blade. To perform the 400 simulations it took 12 minutes. Information regarding the pump head, total efficiency, NPSH and pressure distribution at different spans of the blade are retrieved. With the two main performance indicators being presented in Figure 3.19. While the NPSH was chosen not to be reported, since no big changes were presents in the values. All of the obtained results presented a NPSH between 29.5 and 30.2 meters.

Valuable design insights have been extracted from these simulations regarding how to optimize pump head and efficiency using the POD model. The turbopump under study, characterized by four blades, a shroud, and a rotational speed of 40,000 rpm, and an outlet beta angle of 90°. Specifically, the POD model reveals that positioning the splitter blade further from the main blade maximizes both the pump head and efficiency. Additionally, the model indicates that a shorter splitter blade enhances pump head, while a longer blade boosts efficiency. These findings align with results previously reported by Zakeraholseini [7], further validating the trends observed in this study and showcasing the POD model's capability to accurately simulate and analyze different geometrical configurations in pump design.

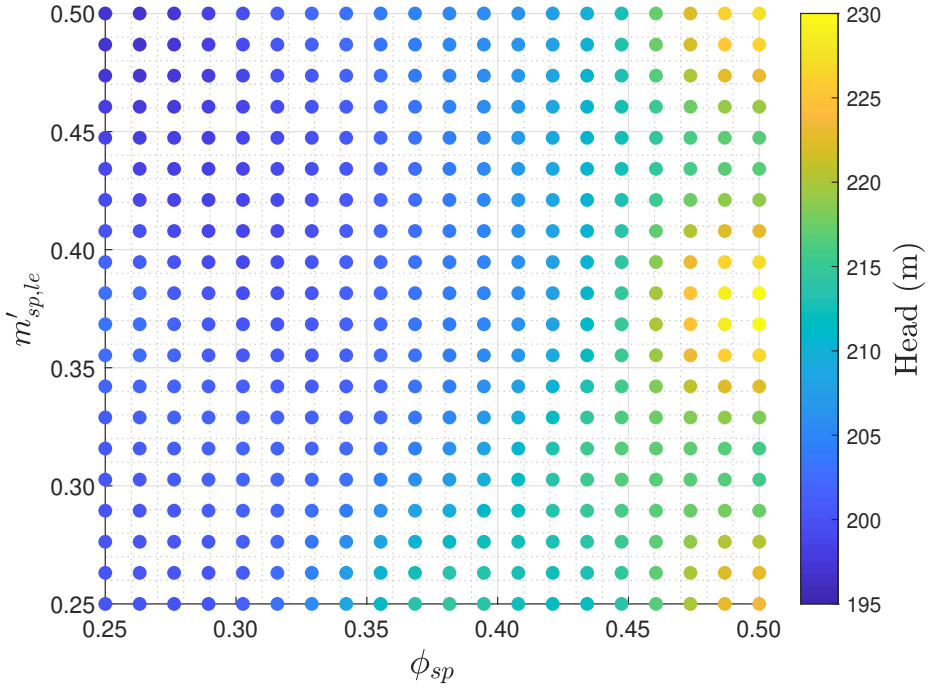


(a) Pressure distribution on training configuration ($m'_{le,sp} = 0.25$, $\psi_{sp} = 0.50$).

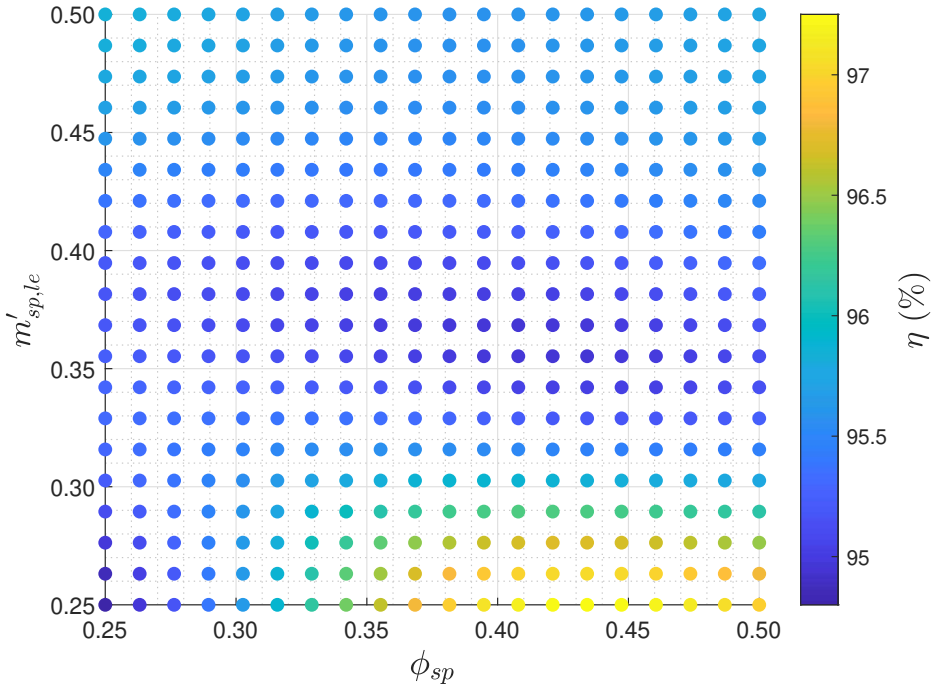


(b) Pressure distribution on test configuration ($m'_{le,sp} = 0.30$, $\psi_{sp} = 0.45$).

Figure 3.18: Comparison between POD and CFD results for the pressure distribution on the main blade close to hub and shroud.



(a) Head prediction.



(b) Total efficiency prediction.

Figure 3.19: Design space evaluation through POD while tuning splitter blades pitch and leading edge positions.

3.4 POD model tuning three parameters

3.4.1 Four blades turbopump

After successfully validating the model with adjustments to two pump parameters, the POD model was extended by incorporating simulations that address changes in the blade outlet angle. This angle was restricted to a range of 50° to 90°, supported by the necessary number of existing simulations. The simulations to build the POD model were based on a turbopump designed to operate at a rotational speed of 25 krpm, featuring a shrouded four-blade configuration and operated at a rotational speed of 40 krpm. For the POD model training 27 simulations have been used, nine simulations characterized by the splitter blade pitch and leading edge combinations as in Figure 3.1 for three different values of outlet blade angle, corresponding to 50°, 70° and 90°. The new design space is presented in Figure 3.20.

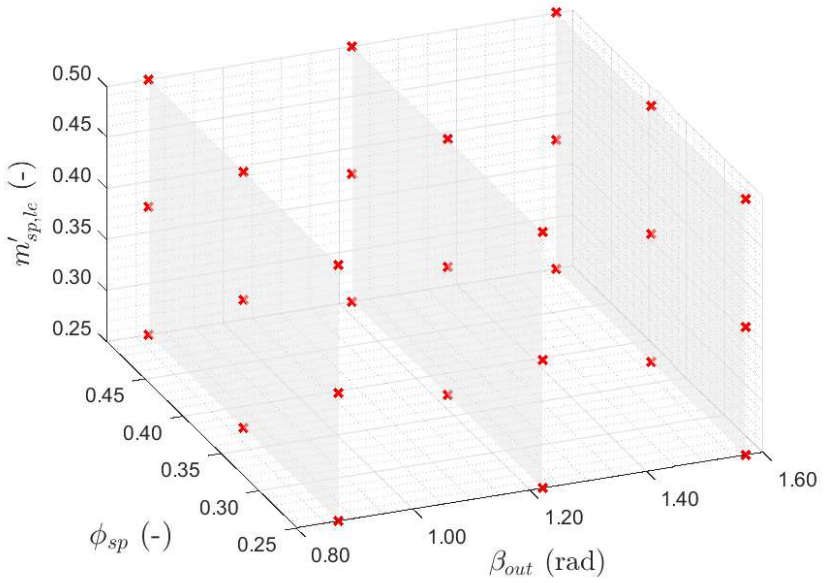


Figure 3.20: Design space for building the POD model of a shrouded turbopump with four blades, where three parameters are being tuned.

The POD model is firstly built and this required approximately 5 minutes. The model is subsequently evaluated using 20 random simulations performed via CFD, validating the POD results for head, total efficiency, and shaft power. The parameters of the configurations used for the test are reported in Table A.6. Overall, the results demonstrated good quality, even if discrepancies between the true and predicted total efficiency could introduce inaccuracies when trying to identify the optimal combination of parameters for the turbopump configuration. The mean error and standard deviation over the 20 simulations used to evaluate the quality of the model are presented in Table 3.9.

Parameter	Pump Head (%)	Shaft power (%)	Total efficiency (pt)
Mean error	1.53	1.47	-0.23
Standard Deviation	1.90	1.36	0.73

Table 3.9: Mean error and standard deviation on the prediction of the main performance indicators for 20 random simulations used to test the model.

The POD model, characterized by its rapid simulation capabilities, was employed to extensively analyze the design space depicted in Figure 3.1, resulting in 8000 equally spaced simulations. These simulations offered new insights into pump head, total efficiency, and NPSH, as detailed in Figure 3.22 and 3.21. While interpreting specific data from these plots may be complex, due to the big quantity of information contained in them. The broader value of these findings lies in their high quality and consistency with the CFD results, as well as the efficiency of the simulation process itself. This analysis is particularly useful in understanding the influence of design modifications on key performance indicators.

Further exploration of the design space is conducted through targeted analyses of certain geometric parameters. For example, Figure 3.23 investigates the effects of setting the leading edge position to 0.5, demonstrating significant increases in both head and efficiency when the splitter blade is further from the main blade, with optimal results at specific outlet angles. The highest head is noted at an outlet angle near 90°, while peak efficiency is observed at 60°.

Similarly, Figure 3.24 explores the impact of the splitter blade length and outlet angle when the pitch position is fixed at a value of 0.35. Here, as for the previous case, an outlet angle of 90° delivers the highest pump head, whereas an outlet angle between 70° and 80° provides the greatest total efficiency. Best designs are identified when the leading edge is set at 0.25. Meaning that a longer blade would help to increase both the head and total efficiency of a turbopump.

Lastly, Figure 3.25 reveals that an outlet angle of 60° optimizes both efficiency and head when the splitter blade is positioned closer to the main blade, with the leading edge position set at 0.37 to maximize the head and to a value of 0.25(longer blade) to maximize the efficiency. These findings underscore the interplay between geometric adjustments and performance outcomes, enriching the understanding of turbopump design optimization.

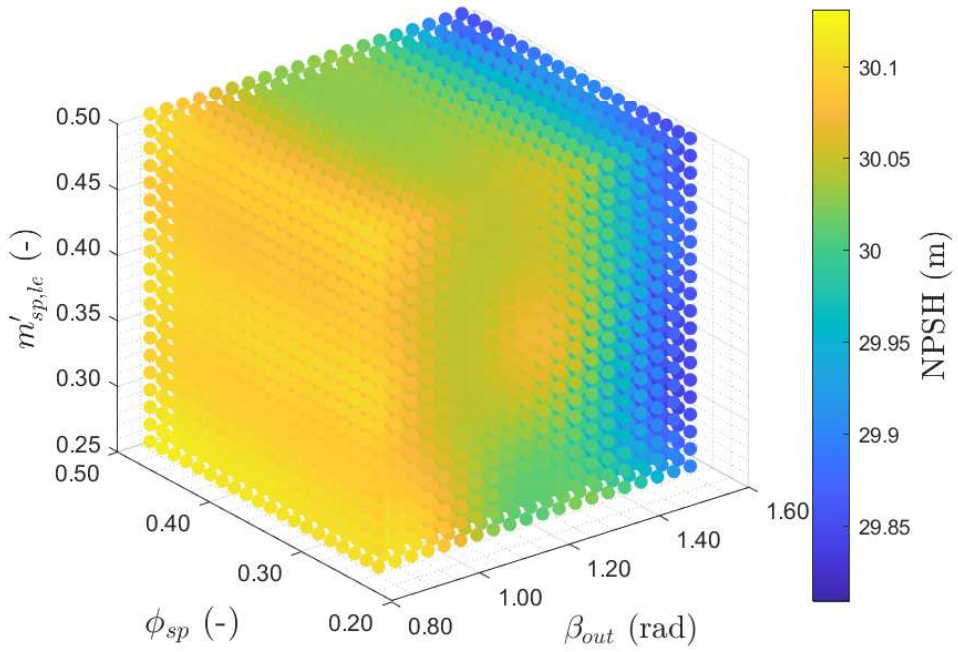
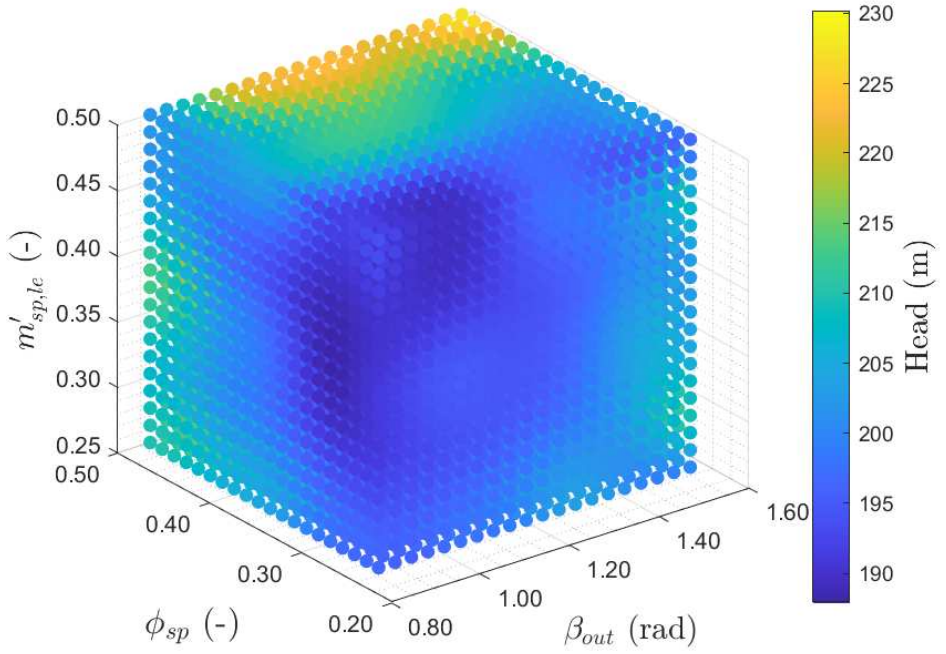
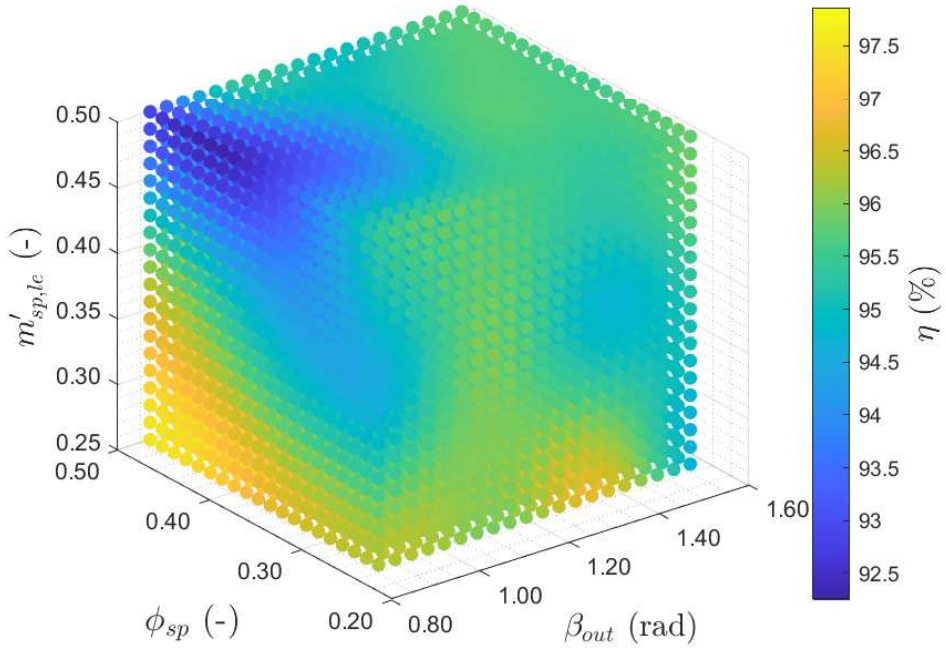


Figure 3.21: NPSH prediction through POD while tuning splitter blades pitch, leading edge positions and blade outlet angle.

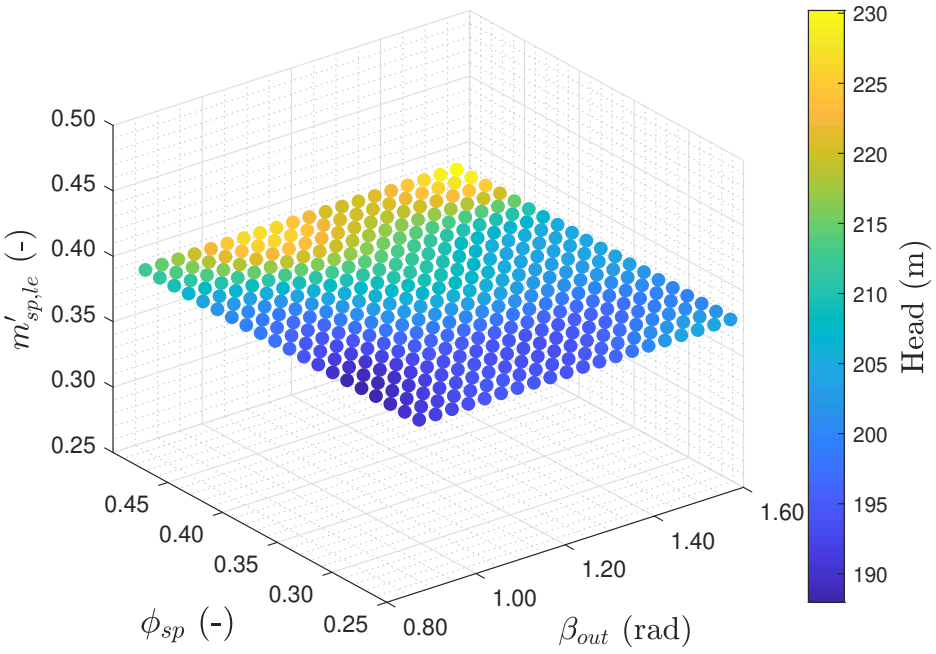


(a) Head prediction.

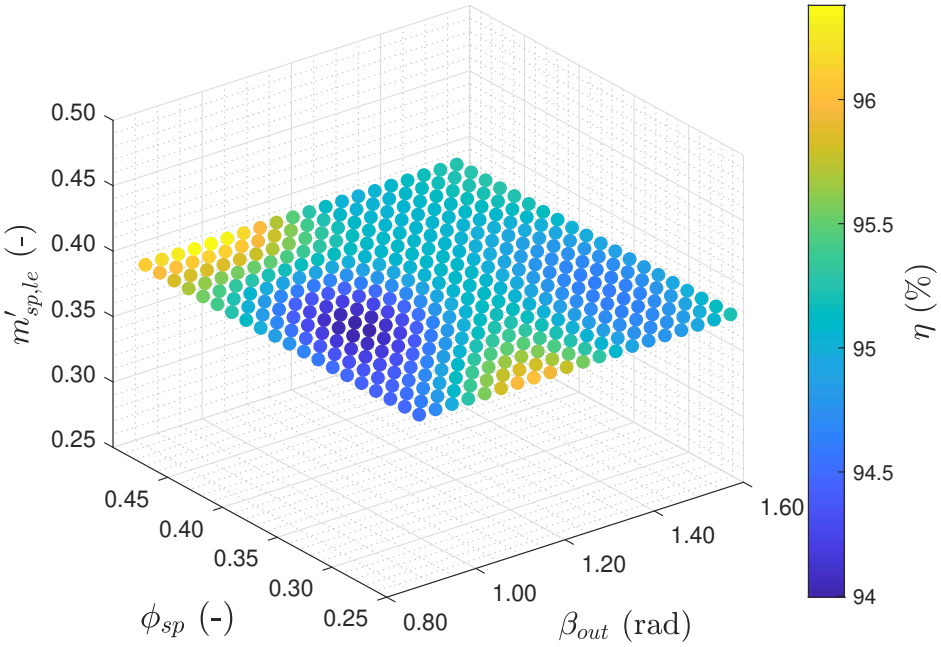


(b) Total efficiency prediction.

Figure 3.22: head and total efficiency prediction through POD while tuning splitter blades pitch, leading edge positions and blade outlet angle.

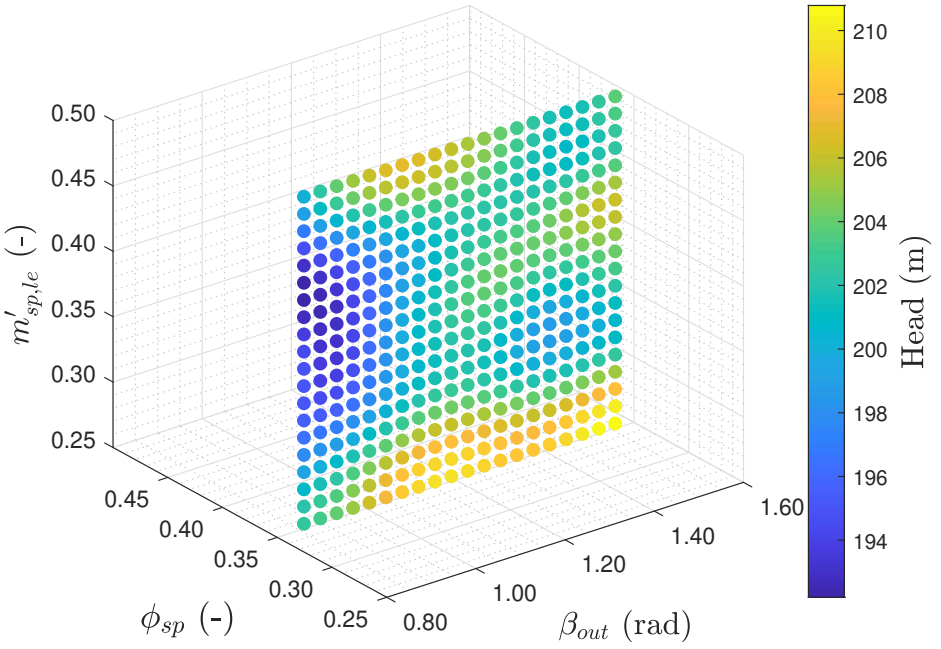


(a) Head prediction.

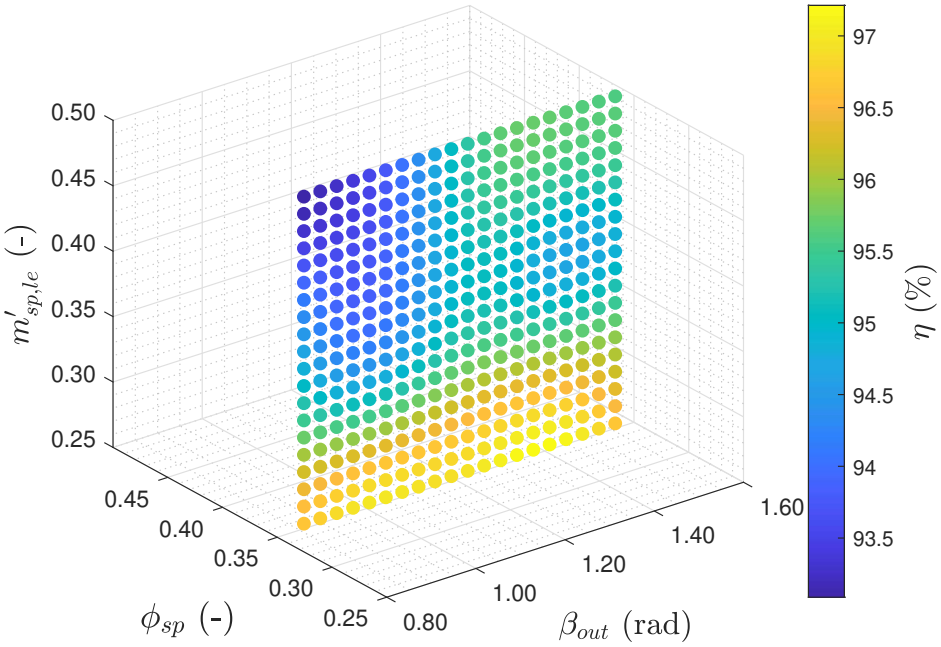


(b) Total efficiency prediction.

Figure 3.23: POD simulation for different splitter blade pitch position and outlet blade angle. Splitter blade leading edge position fixed to 0.38.

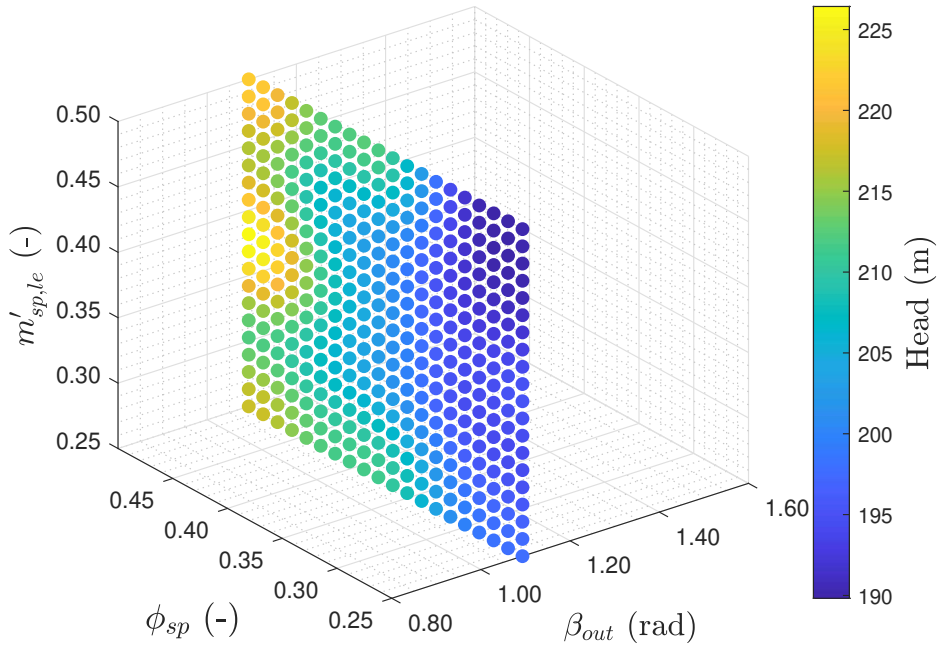


(a) Head prediction.

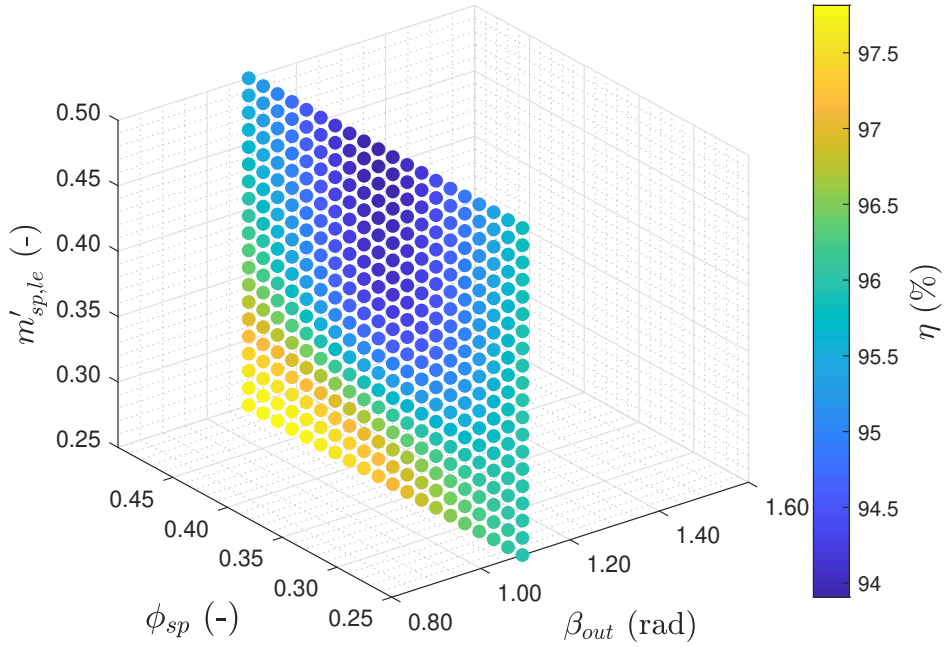


(b) Total efficiency prediction.

Figure 3.24: POD simulation for different splitter blade leading edge position and outlet blade angle. Splitter blade pitch position fixed to 0.35.



(a) Head prediction.



(b) Total efficiency prediction.

Figure 3.25: POD simulation for different splitter blade leading edge position and pitch position. Blade outlet angle fixed to 60°.

Out of the 8000 simulated configurations a Pareto front analysis is performed and the non-dominated solutions are identified. The Pareto front is displayed in Figure 3.26, where 25 non-dominated individuals have been found. Details on this configurations, including the corresponding values for head and efficiency along with their positioning in the design space, are presented in the appendix in Figure A.12. Subsequently, the individuals identified through the POD model are tested, to confirm the quality of the Pareto front search. CFD simulations are then performed on two specific turbopump configurations: the one identified as to maximize the total efficiency and the other to maximize pump head. To validate the parameter combinations suggested by the POD model, these two configurations were compared against the 27 configurations used to construct the POD model, which can be considered as the baseline for the optimization process. The parameters of these two simulations are documented in Table 3.10, and the performance of each configuration evaluated through CFD is present in Figure 3.27, where the performances predicted through CFD are compared with the one of the initial configuration to affirm the capability of the POD model in identifying an optimized solution. As it is possible to see, the model successfully identified non-dominated solutions and showed high predictive accuracy, also in the values of the performance indicators obtained. While the head prediction shows consistency between the POD and CFD results, a discrepancy emerges in the efficiency results for the configuration intended to maximize head: the POD model predicted an efficiency of approximately 95 %, whereas the CFD simulation indicated an efficiency of approximately 97 %. This highlights a slight variation in the efficiency predictions, demonstrating the model's overall effectiveness yet indicating areas for refinement in its predictive capabilities.

Configuration	β_{out}	ψ_{sp}	$m'_{le,sp}$
Max Head	90°	0.5	0.38
Max Efficiency	65°	0.48	0.25

Table 3.10: Configuration maximizing head and total efficiency for turbopumps with four blades.

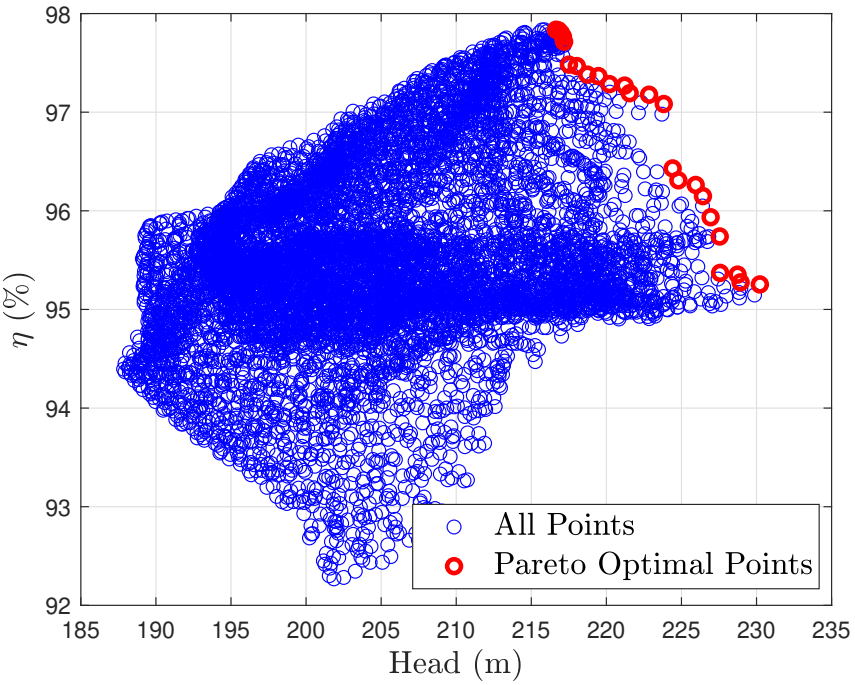


Figure 3.26: Population performance results and Pareto front identification. Case with four blades.

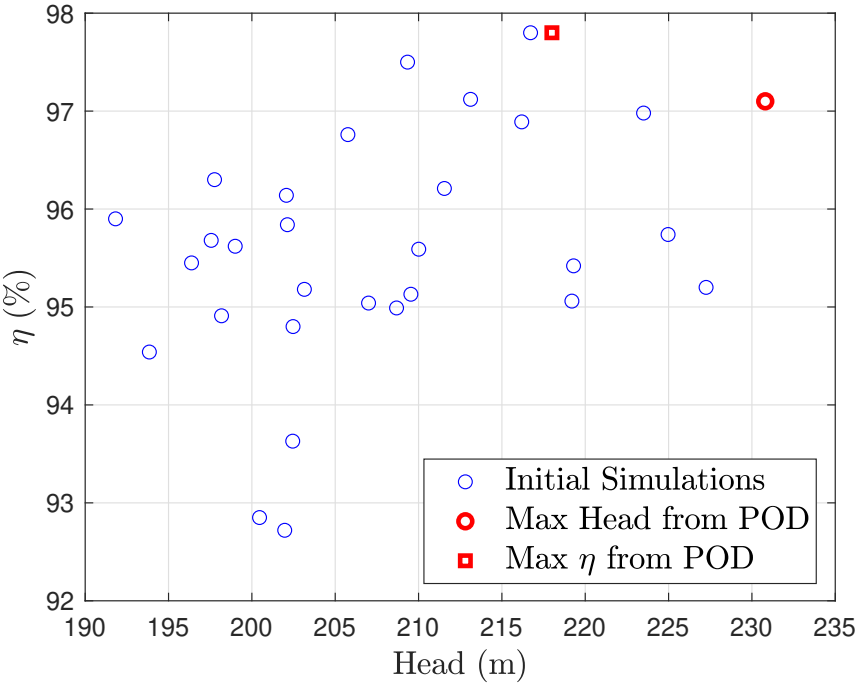


Figure 3.27: Comparison between best solution provided from POD and configurations used to build the POD model for a four blades turbopump.

3.4.2 Five blades turbopump

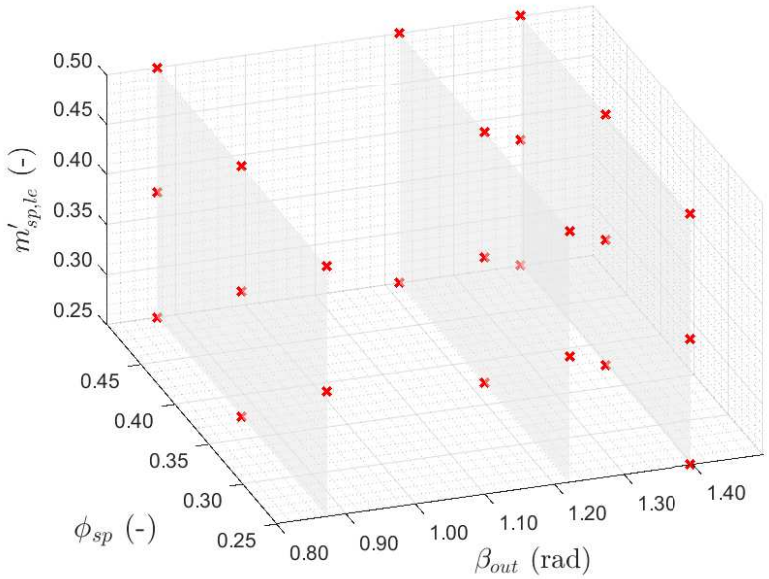


Figure 3.28: Design space for building the POD model of a shrouded turbopump with five blades, where three parameters are being tuned.

The POD model was successfully adapted to accommodate a change in the number of blades, enabling the development of a model based on CFD simulations of turbopumps with a shrouded design and five blades. However, due to the non-convergence of most simulations with an outlet blade angle of 90°, the geometry of the turbopump configurations analyzed through CFD simulations to train the model was limited to outlet angles ranging from 50° to 80°. The design space for this revised POD model is depicted in Figure 3.28. It is noted that three simulations are missing to fully populate the design space as was done in the previous model, resulting in the use of 24 CFD simulations for constructing the new POD model.

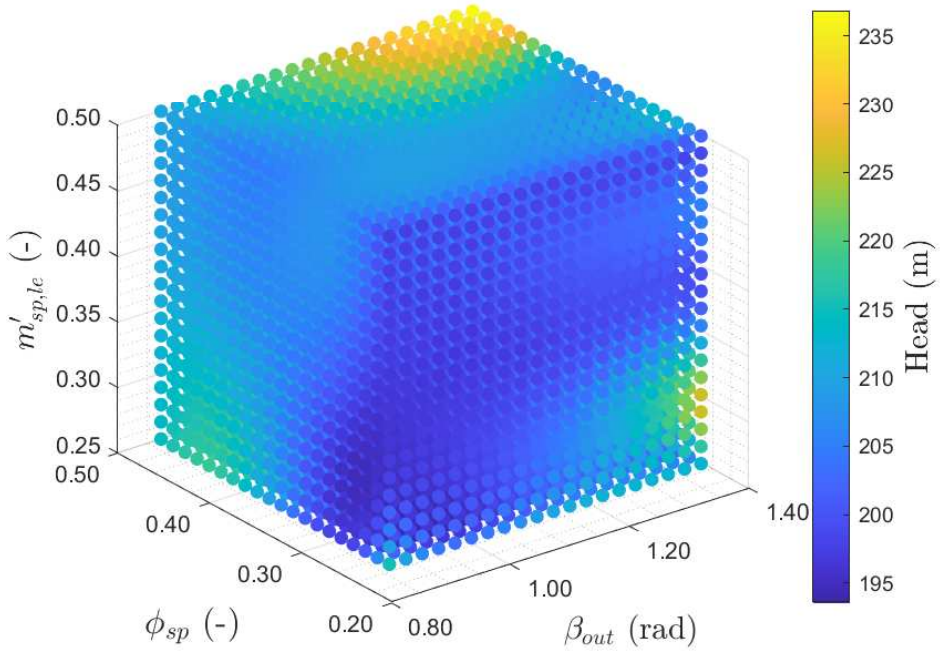
The accuracy of the updated POD model was evaluated using 20 random simulations conducted through CFD. The parameters of the configurations used for the test are reported in Table A.7. The mean error and standard deviation of these tests are reported in Table 3.11. This evaluation revealed a higher error in the prediction of the pump head, whereas the error in predicting the shaft power was lower, compared to the previous model. Despite these variations, the quality of the model is still considered satisfactory, especially given its efficiency in rapidly providing results for each configuration.

Parameter	Pump Head (%)	Shaft power (%)	Total efficiency (pt)
Mean error	1.54	0.16	1.21
Standard Deviation	2.58	0.69	0.68

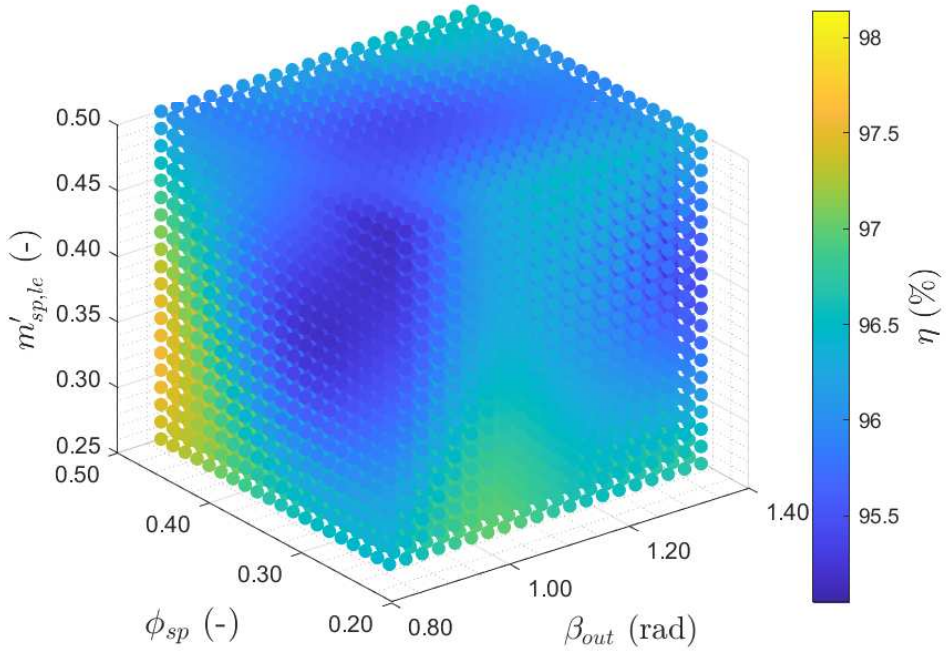
Table 3.11: Mean error and standard deviation on the prediction of the main performance indicators for 20 random simulations used to test the model of the turbopumps characterized by five blades.

This time the POD domain is restricted in terms of beta angle at the outlet between 50° and 80°, also in this occasion 8000 simulations have been performed, to extract information as done for the previous case. The results for the head and total efficiency are presented in Figure 3.29. The analysis, where the design space is examined by fixing one of the three parameters being tuned, is documented in the appendix under section A.7.

As done previously the results provided from the POD model for the 8000 different configurations are used to identify the Pareto front. The non-dominated solutions are highlighted in Figure 3.30. The positioning of these solutions within the design space can be found in the appendix in Figure A.16. This iteration of the Pareto front, identified 38 individuals.



(a) Head prediction.



(b) Total efficiency prediction.

Figure 3.29: POD prediction through POD while tuning splitter blades pitch, leading edge positions and blade outlet angle. Case with five blades.

The parameter combinations predicted by the POD model to deliver the highest head and the highest efficiency are detailed in Table 3.12. These two configurations were tested through CFD simulations and compared against the simulations used to construct the POD model, which are once again used as the baseline for optimization. The results of this comparison are presented in Figure 3.31, demonstrating once more that the POD model effectively identifies optimized turbopump geometrical configurations. This also illustrates the model's adaptability in handling changes in the number of blades in the pump. In contrast to the findings with the previous model, this iteration shows a higher degree of accordance between the predicted total efficiency from the POD model and the actual results obtained from the CFD simulations. However, discrepancies remain in the head prediction results. Specifically, for the configuration predicted to yield the highest pump head, the POD model overestimated the head rise, predicting approximately 235 meters, while the actual value measured from the CFD was 230 meters. This highlights a slight overestimation by the POD model in this specific aspect of pump performance, indicating the need for further refinement to enhance its predictive accuracy. For the current case, this can be due to the not optimal space filling of the simulations used to build the model.

Configuration	β_{out}	ψ_{sp}	$m'_{le,sp}$
Max Head	80°	0.50	0.47
Max Efficiency	60°	0.50	0.33

Table 3.12: Configuration maximizing head and total efficiency for turbopumps with five blades.

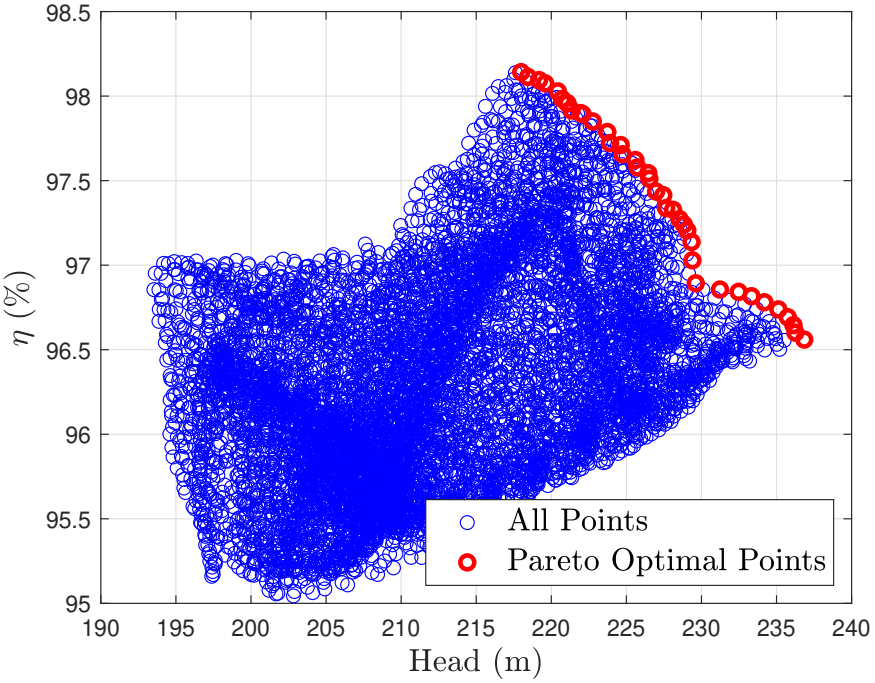


Figure 3.30: Population performance results and Pareto front identification. Case with five blades.

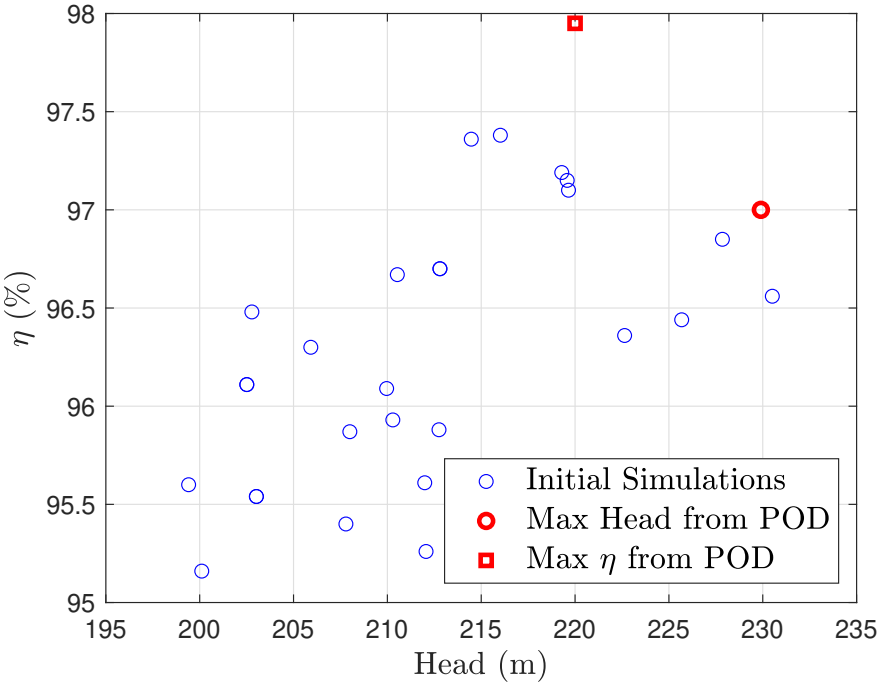


Figure 3.31: Comparison between best solution provided from POD model and configurations used to build the POD model for a four blades turbopump.

3.4.3 Six blades turbopump

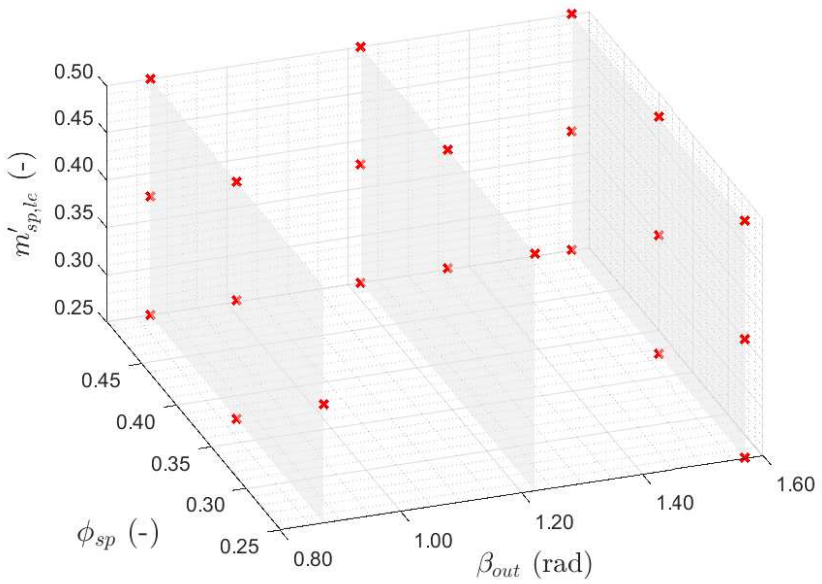


Figure 3.32: Design space for building the POD model of a shrouded turbopump with six blades, where three parameters are being tuned.

The POD model is finally tested for turbopumps characterized by a shrouded design with six blades. The design space for this revised POD model is depicted in Figure 3.32. It is noted that five simulations are missing to fully populate the design space as was done for the first model, resulting in the use of 22 CFD simulations for constructing the new POD model. Still the capability of the model in identifying optimized solutions is tested.

The accuracy of the updated POD model was evaluated using 20 random configuration in the design space, the simulations have been conducted through CFD. The parameters of the configurations used for the test are reported in Table A.8. The mean error and standard deviation of these error analysis are reported in Table 3.13. For the POD model used to predict the performance of the turbopump design characterized by six blades, the pump head prediction accuracy is shown to be worse than that of the first model for a four-blade design. While, the prediction for the power transferred to the shaft is still of acceptable accuracy. The inaccuracies are attributed to the sub optimal filling of the design space with CFD simulations.

Parameter	Pump Head (%)	Shaft power (%)	Total efficiency (pt)
Mean error	4.42	1.24	1.45
Standard Deviation	1.77	0.29	1.10

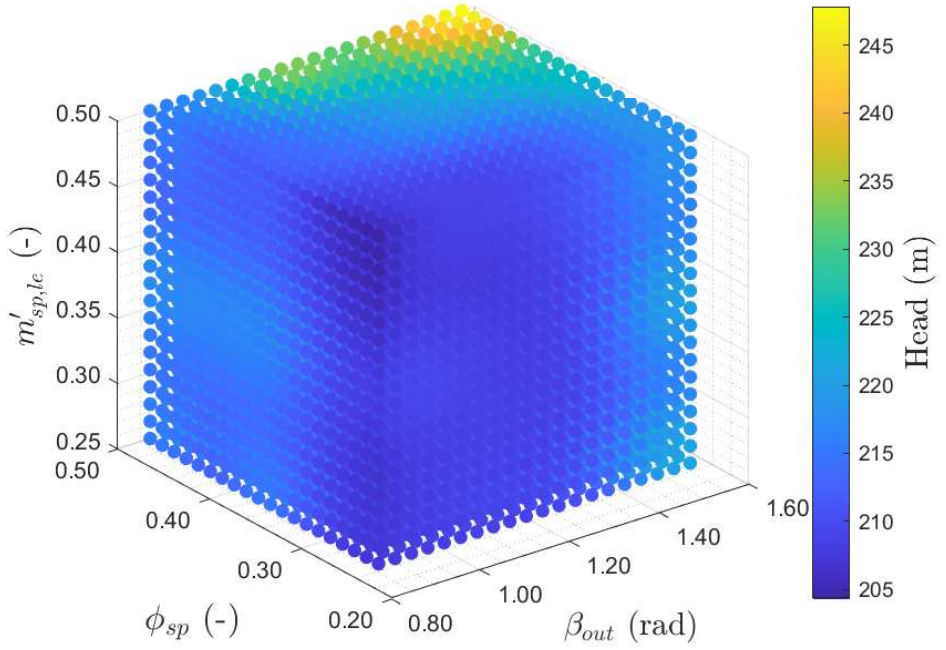
Table 3.13: Mean error and standard deviation on the prediction of the main performance indicators for 20 random simulations used to test the model of the turbopumps characterized by six blades.

Once more 8000 simulations have been performed, to extract information as done for the previous case. The results for the head and total efficiency are presented in Figure 3.29. The analysis, where the design space is examined by fixing one of the three parameters being tuned, is documented in the appendix under section A.8.

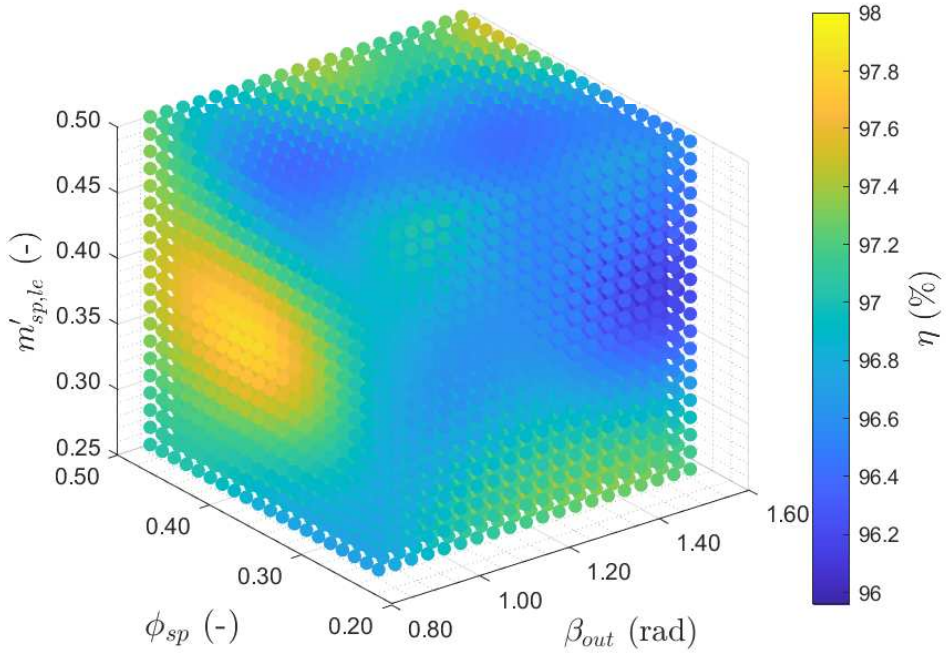
In the main body of the document, the Pareto front is displayed, and the non-dominated solutions are highlighted in Figure 3.34. The positioning of these solutions within the design space can be found in the appendix in Figure A.20. This iteration of the Pareto front includes 26 individuals. The parameter combinations predicted by the POD model to deliver the highest head and the highest efficiency are detailed in Table 3.14. These two configurations were tested through CFD simulations and compared against the simulations used to construct the POD model, which are once again considered as the baseline for optimization in Figure 3.35. For the current case the POD model failed in identifying, the geometrical configuration that would provide the highest efficiency, while the configuration that yields the highest head is once more identified, even if for the current case it does not provide any additional information if compared with the starting simulations. Since the simulation providing the highest head is one of those used to build the model. The significant error presented by the last model can once again be attributed to the lack of sufficient simulations to properly populate the design space.

Configuration	β_{out}	ψ_{sp}	$m'_{le,sp}$
Max Head	90°	0.50	0.50
Max Efficiency	65°	0.50	0.38

Table 3.14: Configuration maximizing head and total efficiency for turbopumps with six blades.



(a) Head prediction.



(b) Total efficiency prediction.

Figure 3.33: POD prediction through POD while tuning splitter blades pitch, leading edge positions and blade outlet angle. Case with six blades.

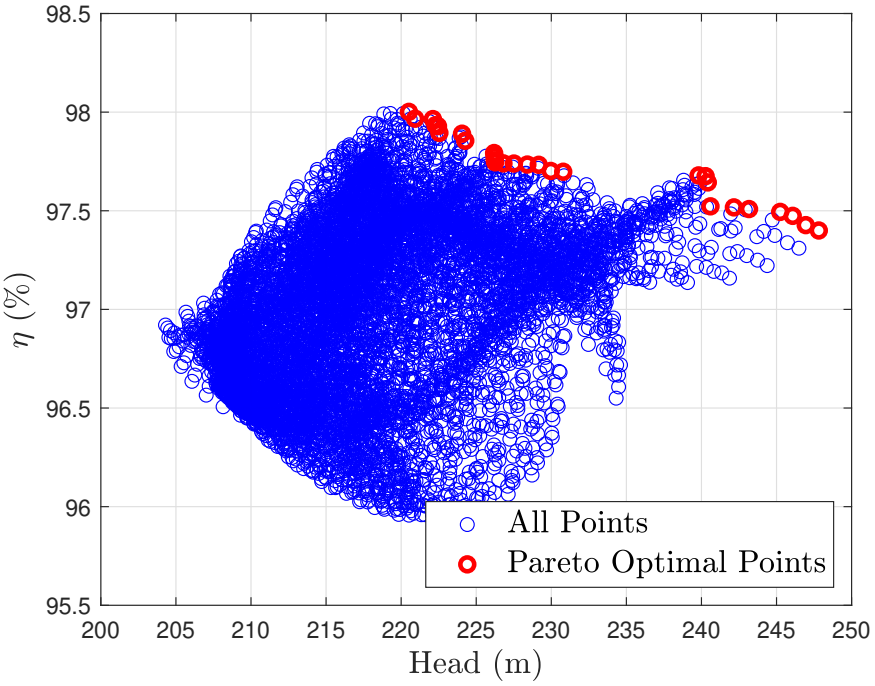


Figure 3.34: Population performance results and Pareto front identification. Case with six blades.

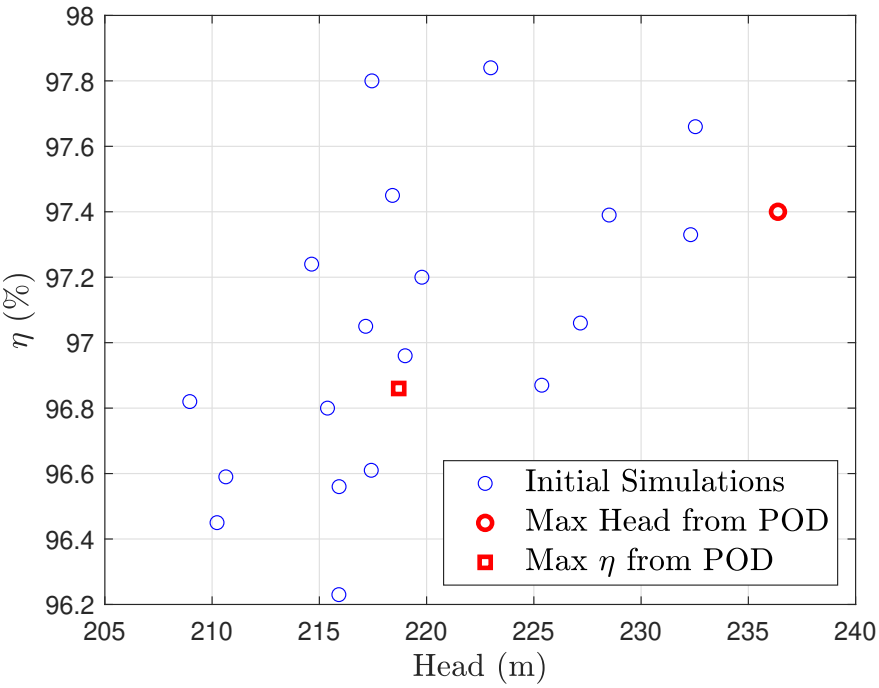


Figure 3.35: Comparison between best solution provided from POD and configurations used to build the POD model for a six blades turbopump.

Chapter 4

Discussion

4.1 Mesh interpolation and data extraction

This work proposes a novel method for building a reduced order model, connecting it a posteriori to the innovative approach by Zakerholseini that enabled the generation of a large number of distinct geometric configurations.

From the start, challenges arose in applying the Proper Orthogonal Decomposition to data extracted from `CFX`. Expected issues included inconsistencies in the mesh topology due to significant geometric changes. This challenge was turned into an opportunity by developing a zonal-focused POD analysis, which targets the surfaces most relevant for extracting crucial information for turbomachinery analyses. This aspect represents the core innovation of this work.

During the development process, several challenges were addressed, beginning with the need for a consistent mesh topology to conduct an accurate analysis and construct the covariance matrix. This led to the creation of three `MATLAB` functions to achieve the results presented in section 2.4. These functions rely on data extracted from `CFX` using a script. One significant challenge was dealing with numerical inconsistencies in the extracted information, primarily due to discrepancies arising from data transfer between the meshing software, `Turbogrid`, and the CFD software, `CFX`. While the meshing software organizes information according to the mesh topology (cell connectivity and boundary conditions), `CFX` uses geometrical positions. This discrepancy led to small numerical errors that made data processing problematic.

In Figure 4.1, the entire geometry of the turbopump impeller is presented, highlighting the outlet surface in red. Although the full geometry is shown, only one section is used for the CFD simulation, leveraging symmetry boundary conditions. This creates a lateral surface automatically generated by the software. In Figure 4.2, one of these boundaries is visible on the outlet surface. The different layers between the hub and shroud are visible vertically, while the layers in the theta direction are horizontal.

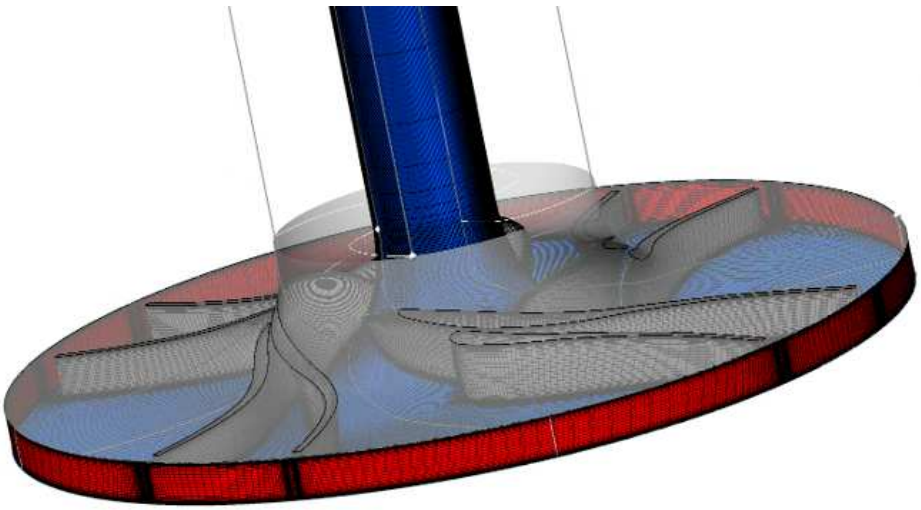


Figure 4.1: Turbopump geometry with the outlet surface highlighted.

From this figure, two challenges in creating a mesh with a consistent topology are evident. The first is the skewed boundary end, requiring precise border detection for accurate interpolation. The second involves numerical errors in `CFX` during data extraction. Although meshing software typically tags each layer sequentially, `CFX`'s reliance on geometric values sometimes misplaces layers, with lower layers appearing above higher ones. This creates difficulties in identifying boundaries for a common mesh. The developed function addresses this issue for both the outlet and inlet surfaces but could not solve similar problems in the stream-wise direction, for this case a 3D mesh consistent between geometries was to be created.

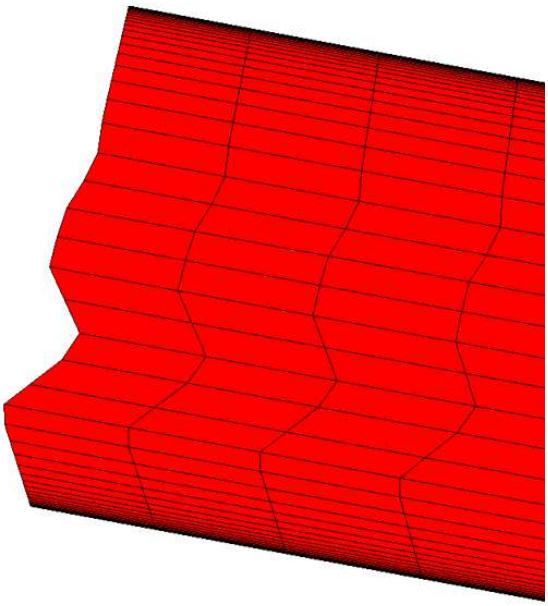


Figure 4.2: Outlet surface zoom on boundary.

Similar challenges arose when creating a consistent mesh surface for the blade to blade surface. However, in this case, the problem was addressed differently, as explained in subsection 2.4.1, leading to successful results as shown in Figure 2.6. Initially, the goal was to use this surface to accurately visualize flow features, and to apply multiple planes at various heights to generate a complete 3D discretization of the domain. Ultimately, this approach was abandoned because projecting information onto multiple planes at different heights would have introduced numerical errors, primarily due to the difficulty of maintaining the conservation of certain variables, thereby affecting the quality of the results.

It's important to note that, without this issue, the strategy could have presented an intriguing process that would have opened an opportunity in the implementation of physics-based constraints. This would have transformed the POD approach into a method that incorporates both data-driven and physics-based principles.

Some other minor numerical problems in the data extracted from `ANSYS` is the reasons for which the strategy that was presented for the creation of the mesh for the blade through polylines was chosen.

Before implementing the strategy presented in this work, other approaches were attempted. The first involved internally defining points from which to extract information about certain variables of interest. However, this method encountered problems and failed to provide consistent discretization across the entire domain due to significant changes in the geometry.

Another approach tried was creating a point cloud in `MATLAB` to extract data at specific positions defined according to the boundaries, leveraging the software's capabilities and using these positions as probes. Unfortunately, `CFX` required internal information that could not be provided, and the `ANSYS` technical support could not offer a solution to this issue.

Through both the literature review and practical experience in applying the POD method efficiently, it became evident that the best approach involves ensuring a consistent mesh topology from the outset when setting up CFD simulations. However, this is not feasible in all cases, especially when significant geometric changes are applied to the studied structure. According to the literature review, POD was often combined with in-house software, while the only commercial software used was `NUMECA`.

In this work, however, it was demonstrated that creating surfaces to project results and maintain consistent topology across simulations could be a viable strategy and provides an interesting alternative. Despite some challenges that have been addressed, this approach shows promise for achieving robust and consistent data analysis.

4.2 POD model construction

One of the steps in the POD model construction, that significantly affects the results is the chosen space-filling technique. This effect is particularly evident in the three models built for a turbopump with different numbers of blades. The primary reason for the greater discrepancy in the model with six blades compared to the four-blade configuration is the fewer number of simulations used. This is also verified by decreasing the number of simulations for the four-blade case and observing that the error on the 20 test configuration increases. Although there is no clear linear relationship between the number of simulations and the error, it underscores the importance of not only the quantity of simulations but also their placement in the design space.

4.3 POD reconstruction

The quality of the POD model prediction was also highly dependent on the interpolation method used for the modal coefficients. For the POD of a blade to blade surface, the linear interpolation method yielded a mean relative error of approximately 15 % with respect to the design space presented in Figure 2.12. While this error could be considered acceptable given the complexity of the flow, linear interpolation produced many outliers due to the flow's non-linearity.

Cubic and RBF (Radial Basis Function) interpolation methods performed similar, each with a mean relative error of approximately 10 %. However, the RBF interpolation showed a higher sensitivity to the number of simulations used for constructing the POD, resulting in worse outcomes with an increased number of simulations.

The Kriging method demonstrated superior results, providing high-quality predictions and good adaptability to different surfaces and variable fields.

Interpolation Method	Kriging	Linear	Cubic	RBF
Mean RE_v (%)	7.2	15.2	10.1	10.6
Outliers (%) $RE_v > 100\%$	0.3	10.4	5.7	7.2

Table 4.1: Quality of the velocity field prediction on the blade to blade surface.

This analysis suggested that the Kriging method was the most suited for the POD model of the turbopump, the values presented in Table 4.1 are mean values between the six simulations used to check the quality of the blade to blade surface prediction. Their parameters are presented in Table 3.1. In all of the results presented the Kriging method was used, and the overall quality of the results can be considered of good quality.

4.4 POD capabilities

The POD capability to build a reduced order model for a small-scale turbopump has been tested. Firstly, the accuracy in predicting the flow field for various variables was investigated, and the obtained accuracy was shown to be of good quality, especially considering the short amount of time required to obtain the results. POD is known for its ability to identify coherent structures in the flow, which is used to decompose the data. These coherent structures were presented, showing that, within the context of a turbopump, it is not immediately possible to extract useful knowledge from this analysis.

A model built on the capability of POD in reconstructing flow-field information was constructed, tested, and used to perform optimization. Specifically, two models were built: one with two parameters being tuned and a second with three parameters being perturbed. Both models showed good quality, despite some inaccuracies characteristic of reduced order models, and demonstrated excellent computational performance, allowing for the extraction of information that would otherwise take 2000 times longer using a full-order model.

In particular, the model where three parameters were tuned was tested for its optimization capability and showed a good ability to identify optimal configurations. This highlights the capability of POD in building a reduced order model that can enhance the design process of centrifugal turbomachinery.

Chapter 5

Conclusions

This thesis investigated the application of Proper Orthogonal Decomposition (POD) for constructing Reduced Order Models (ROMs) of small-scale turbopumps, focusing on geometrical variations. The primary goal was to develop a computationally efficient method that leverages CFD simulations to create ROMs, thus reducing the need for extensive full-order CFD simulations in subsequent analyses.

The study demonstrated that while POD has a strong capability for identifying the main modes of flow fields and reconstructing flow field data with reduced computational resources, there are limitations. Applying POD to varying geometries posed significant challenges, particularly in maintaining consistent mesh topology across different configurations. Interpolation methods such as linear, cubic, Kriging, and Radial Basis Function (RBF) had varying degrees of success, with Kriging showing the best adaptability and accuracy.

The accuracy of the POD models was shown to be highly dependent on the quality and placement of CFD simulations within the design space. In cases where the design space was not optimally filled, the POD model's prediction accuracy decreased. This was particularly evident when comparing models built with different numbers of blades. The study confirmed that not only the quantity but also the strategic placement of simulations is crucial for building an effective ROM.

Despite these challenges, the research achieved significant reductions in computational time while maintaining acceptable levels of accuracy in performance prediction. The developed ROM was capable of predicting key performance indicators such as pump head and shaft power, and it was used successfully to optimize turbopump design by tuning specific parameters.

The work highlighted the importance of integrating POD with commercial software like `ANSYS` for efficient data extraction and adaptation. The innovative approach of using POD to reconstruct flow fields on specific surfaces provided valuable insights into the turbopump's performance characteristics.

Future work should focus on enhancing the robustness of the ROM by exploring more advanced interpolation methods and incorporating a greater number of high-quality CFD simulations. Additionally, integrating physics-based constraints into the POD model could further improve accuracy and reduce inconsistencies.

This strategy can be used to perform multi-point optimization of turbomachinery design. The model demonstrated its capability to handle changes in the number of blades for each design,

indicating that changes in operational conditions do not present significant difficulties. Moreover, performing smaller geometrical changes to the turbomachinery, which have a lower impact on mesh topology, can ensure higher model quality.

In conclusion, this study has demonstrated the potential of POD as a powerful tool for the design and optimization of small-scale turbopumps, offering a faster, more computationally efficient alternative to traditional CFD simulations. This work contributes to the broader field of turbomachinery by providing a framework for developing reduced order models that can significantly enhance the design process and overall performance of turbopumps.

Bibliography

[1] Hannah Ritchie. "Cars, planes, trains: where do CO2 emissions from transport come from?" In: *Our World in Data* (2020). https://ourworldindata.org/co2-emissions-from-transport.

[2] Michael A. Tamor and Ellen B. Stechel. "Electrification of transportation means a lot more than a lot more electric vehicles". In: *iScience* 25.6 (2022), p. 104376. ISSN: 2589-0042. DOI: https://doi.org/10.1016/j.isci.2022.104376. URL: https://www.sciencedirect.com/science/article/pii/S2589004222006472.

[3] Roberto Capata and Claudia Toro. "Feasibility analysis of a small-scale ORC energy recovery system for vehicular application". In: *Energy Conversion and Management* volume 86 (June 2014), pp. 1078–1090. DOI: 10.1016/j.enconman.2014.06.024.

[4] Hua Tian, Peng Liu, and Gequn Shu. "Challenges and opportunities of Rankine cycle for waste heat recovery from internal combustion engine". In: *Progress in Energy and Combustion Science* 84 (2021), p. 100906. ISSN: 0360-1285. DOI: https://doi.org/10.1016/j.pecs.2021.100906. URL: https://www.sciencedirect.com/science/article/pii/S0360128521000046.

[5] Fanxiao Meng et al. "Study of efficiency of a multistage centrifugal pump used in engine waste heat recovery application". In: *Applied Thermal Engineering* 110 (Jan. 2017). DOI: 10.1016/j.applthermaleng.2016.08.226.

[6] Sajjad Zakeralhoseini and Jürg Schiffmann. "Design, computational and experimental investigation of a small-scale turbopump for organic Rankine cycle systems". In: *Energy Conversion and Management* 287 (2023), p. 117073. ISSN: 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2023.117073. URL: https://www.sciencedirect.com/science/article/pii/S0196890423004193.

[7] Sajjad Zakeralhoseini and Jürg Schiffmann. "The influence of splitter blades and meridional profiles on the performance of small-scale turbopumps for ORC applications; analysis, neural network modeling and optimization". In: *Thermal Science and Engineering Progress* 39 (2023), p. 101734. ISSN: 2451-9049. DOI: https://doi.org/10.1016/j.tsep.2023.101734. URL: https://www.sciencedirect.com/science/article/pii/S2451904923000872.

[8] Sajjad Zakeralhoseini and Jürg Schiffmann. "Analysis and modeling of the tip leakage flow on the performance of small-scale turbopumps for ORC applications". In: *Applied Thermal Engineering* 217 (2022), p. 119160. ISSN: 1359-4311. DOI: https://doi.org/10.1016/j.applthermaleng.2022.119160. URL: https://www.sciencedirect.com/science/article/pii/S1359431122010912.

[9] Vittorio Michelassi and Julia Ling. "Challenges and opportunities for artificial intelligence and high-fidelity simulations in turbomachinery applications: A perspective". In: *Journal of the Global Power and Propulsion Society* (May 2021), pp. 1–14. DOI: 10.33737/jgpps/135173.

[10] Esra Sorguven, Sevil Incir, and Jonathan Highgate. "Understanding loss generation mechanisms in a centrifugal pump using large eddy simulation". In: *International Journal of Heat and Fluid Flow* 96 (2022), p. 108994. ISSN: 0142-727X. DOI: https://doi.org/10.1016/j.ijheatfluidflow.2022.108994. URL: https://www.sciencedirect.com/science/article/pii/S0142727X22000698.

[11] James Hammond et al. "Machine Learning Methods in CFD for Turbomachinery: A Review". In: *International Journal of Turbomachinery, Propulsion and Power* 7.2 (2022). ISSN: 2504-186X. URL: https://www.mdpi.com/2504-186X/7/2/16.

[12] JiaQi Luo et al. "Flow reconstructions and aerodynamic shape optimization of turbomachinery blades by POD-based hybrid models". In: *Science China Technological Sciences* 60.11 (Nov. 1, 2017), pp. 1658–1673. ISSN: 1869-1900. DOI: 10.1007/s11431-016-9093-y. URL: https://doi.org/10.1007/s11431-016-9093-y.

[13] Zhiguo Wei et al. "Fast prediction of the performance of the centrifugal pump based on reduced-order model". In: *Energy Reports* 9 (2023). Selected papers from 2022 7th International Conference on New Energy and Future Energy Systems, pp. 51–64. ISSN: 2352-4847. DOI: https://doi.org/10.1016/j.egyr.2023.02.068. URL: https://www.sciencedirect.com/science/article/pii/S2352484723002111.

[14] Simão Marques et al. "Nonintrusive Aerodynamic Shape Optimisation with a POD-DEIM Based Trust Region Method". In: *Aerospace* 10.5 (2023). ISSN: 2226-4310. DOI: 10.3390/aerospace10050470. URL: https://www.mdpi.com/2226-4310/10/5/470.

[15] G Berkooz, PJ Holmes, and John Lumley. "The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows". In: *Annual Review of Fluid Mechanics* 25 (Nov. 2003), pp. 539–575. DOI: 10.1146/annurev.fl.25.010193.002543.

[16] Jeffrey P. Thomas, Kenneth C. Hall, and Earl H. Dowell. "REDUCED-ORDER AEROELASTIC MODELING USING PROPER-ORTHOGONAL DECOMPOSTIONS". In: URL: https://api.semanticscholar.org/CorpusID:219529687.

[17] L. Sirovich and M. Kirby. "Low-dimensional procedure for the characterization of human faces". In: *J. Opt. Soc. Am. A* 4.3 (Mar. 1987), pp. 519–524. DOI: 10.1364/JOSAA.4.000519. URL: https://opg.optica.org/josaa/abstract.cfm?URI=josaa-4-3-519.

[18] T. Bui-Thanh, M. Damodaran, and K. Willcox. "Aerodynamic Data Reconstruction and Inverse Design Using Proper Orthogonal Decomposition". In: *AIAA Journal* 42.8 (2004), pp. 1505–1516. DOI: 10.2514/1.2159. eprint: https://doi.org/10.2514/1.2159. URL: https://doi.org/10.2514/1.2159.

[19] Emiliano Iuliano and Domenico Quagliarella. "Aerodynamic shape optimization via non-intrusive POD-based surrogate modelling". In: *2013 IEEE Congress on Evolutionary Computation*. 2013, pp. 1467–1474. DOI: 10.1109/CEC.2013.6557736.

[20] M. Ripepi et al. "Reduced-order models for aerodynamic applications, loads and MDO". In: *CEAS Aeronautical Journal* 9.1 (2018), pp. 171–193. ISSN: 1869-5590. DOI: 10.1007/s13272-018-0283-6.

[21] Yanhui DUAN et al. "Performance improvement of optimization solutions by POD-based data mining". In: *Chinese Journal of Aeronautics* 32.4 (2019), pp. 826–838. ISSN: 1000-9361. DOI: https://doi.org/10.1016/j.cja.2019.01.014. URL: https://www.sciencedirect.com/science/article/pii/S1000936119300627.

[22] Jiaqi Luo. "Design Optimization of the Last Stage of a 4.5-Stage Compressor Using a POD-based Hybrid Model". In: *Aerospace Science and Technology* 76 (Feb. 2018). DOI: 10.1016/j.ast.2018.01.043.

[23] Ren-Hui Zhang et al. "Inverse Method of Centrifugal Pump Impeller Based on Proper Orthogonal Decomposition (POD) Method". In: *Chinese Journal of Mechanical Engineering* 30.4 (2017), pp. 1025–1031. ISSN: 2192-8258. DOI: 10.1007/s10033-017-0137-x.

[24] K. Willcox, J. Peraire, and J. White. "An Arnoldi approach for generation of reduced-order models for turbomachinery". In: *Computers & Fluids* 31.3 (2002), pp. 369–389. ISSN: 0045-7930. DOI: https://doi.org/10.1016/S0045-7930(01)00046-9. URL: https://www.sciencedirect.com/science/article/pii/S0045793001000469.

[25] David J. Lucia, Philip S. Beran, and Walter A. Silva. "Reduced-order modeling: new approaches for computational physics". In: *Progress in Aerospace Sciences* 40.1 (2004), pp. 51–117. ISSN: 0376-0421. DOI: https://doi.org/10.1016/j.paerosci.2003.12.001. URL: https://www.sciencedirect.com/science/article/pii/S0376042103001131.

[26] *POD-Driven Adaptive Sampling for Efficient Surrogate Modeling and its Application to Supersonic Turbine Optimization*. Vol. Volume 2B: Turbomachinery. Turbo Expo: Power for Land, Sea, and Air. June 2014, V02BT45A023. DOI: 10.1115/GT2014-27229. eprint: https://asmedigitalcollection.asme.org/GT/proceedings-pdf/GT2014/45615/V02BT45A023/4228977/v02bt45a023-gt2014-27229.pdf. URL: https://doi.org/10.1115/GT2014-27229.

[27] Julien Weiss. "A Tutorial on the Proper Orthogonal Decomposition". In: June 2019. DOI: 10.2514/6.2019-3333.

[28] Ivo Couckuyt, Tom Dhaene, and Piet Demeester. "ooDACE Toolbox: A Flexible Object-Oriented Kriging Implementation". In: *Journal of Machine Learning Research* 15 (Oct. 2014), pp. 3183–3186.

[29] Ivo Couckuyt et al. "Blind Kriging: Implementation and performance analysis". In: *Advances in Engineering Software* 49 (July 2012). DOI: 10.1016/j.advengsoft.2012.03.002.

[30] Fasheng Sun, Yuhui Yin, and Min-Qian Liu. "Construction of nested space-filling designs using difference matrices". In: *Journal of Statistical Planning and Inference* 143 (Jan. 2013), pp. 160–166. DOI: 10.1016/j.jspi.2012.06.021.

[31] Muhamad Safi'i. *Advanced Engineering Mathematics by Erwin E.* Sept. 2020. ISBN: ISBN 978-0-470-45836-5.

Appendix A

Extra plots

A.1 Surface mesh interpolation

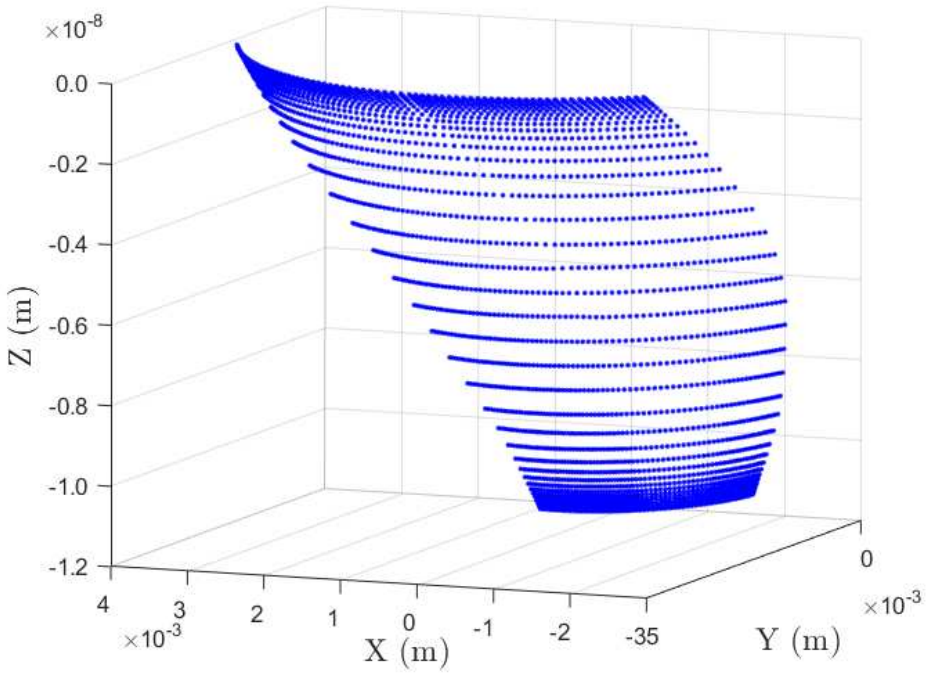


Figure A.1: Original mesh at the inlet in 3D ($m'_{le,sp} = 0.5$, $\psi_{sp} = 0.25$).

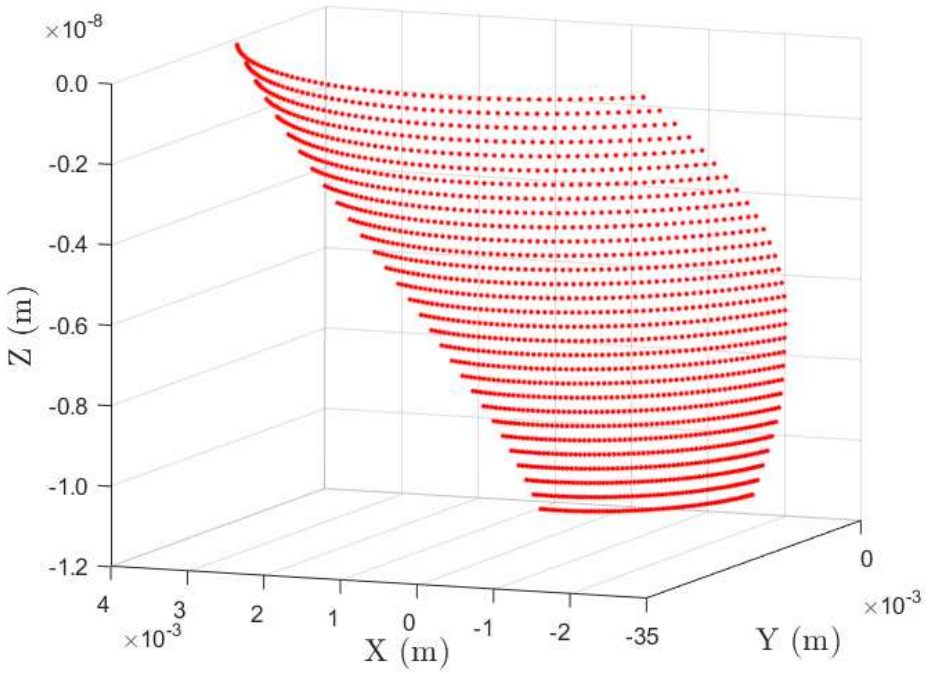


Figure A.2: Interpolated mesh at the inlet in 3D ($m'_{le,sp} = 0.5$, $\psi_{sp} = 0.25$).

A.2 Blade to blade plane initial POD

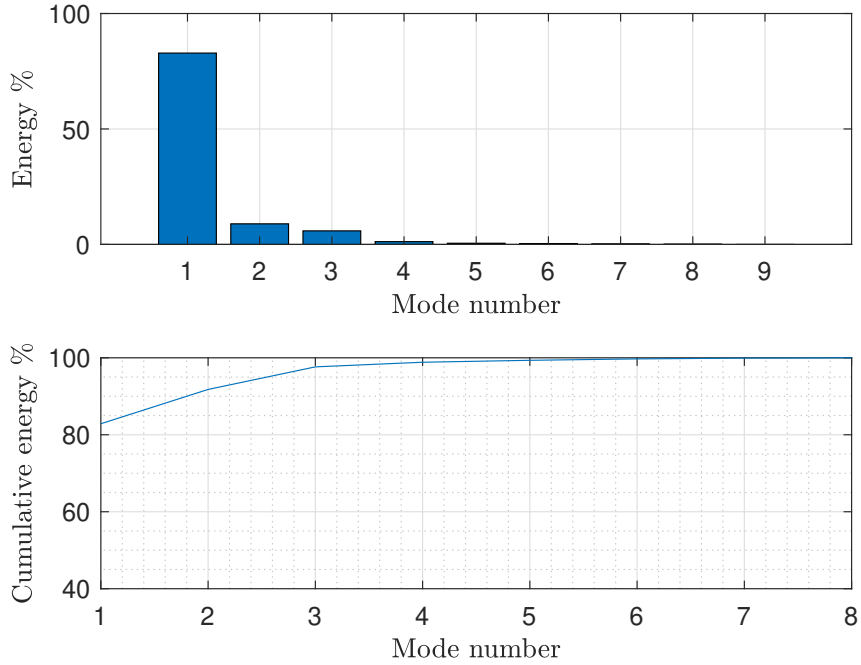


Figure A.3: Energy distribution of POD eigenmode for the pressure field in the blade to blade surface. In the upper subplot of the figure, the energy distribution across each mode is depicted. The lower subplot illustrates the cumulative energy.

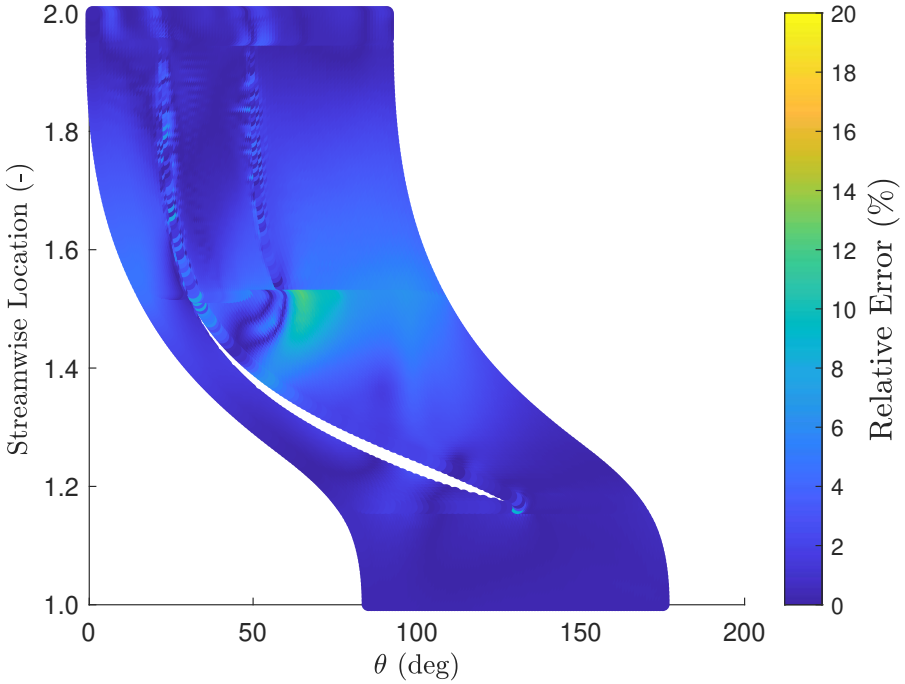
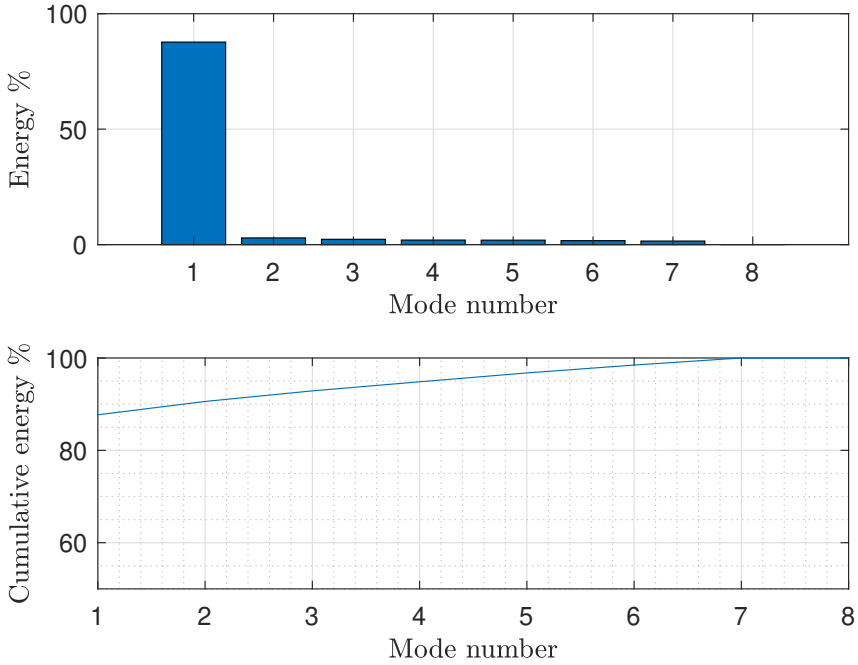
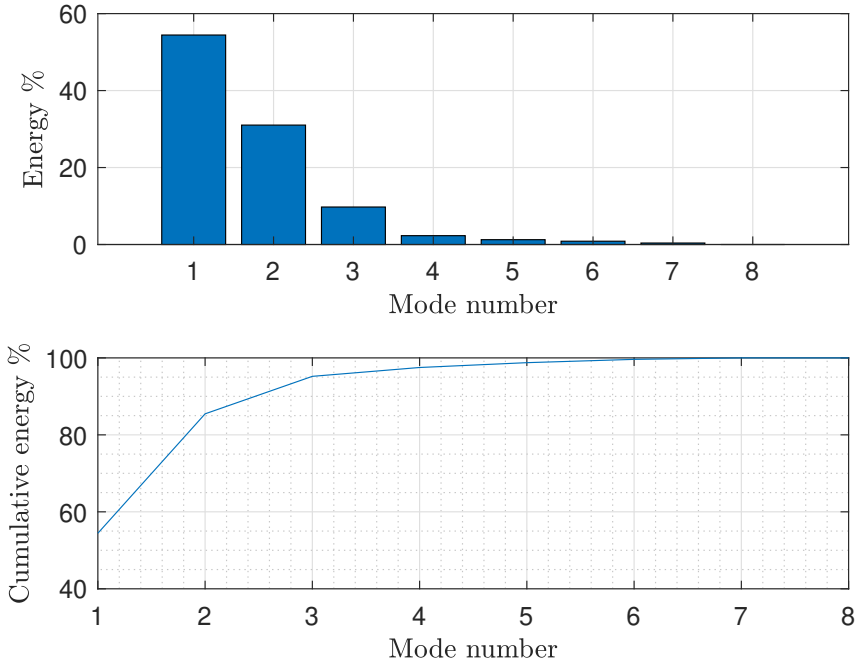


Figure A.4: Relative Error between CFD results and POD results of the blade to blade surface at $S_N = 0.5$ for the pressure field ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

A.3 Outlet and inlet initial POD

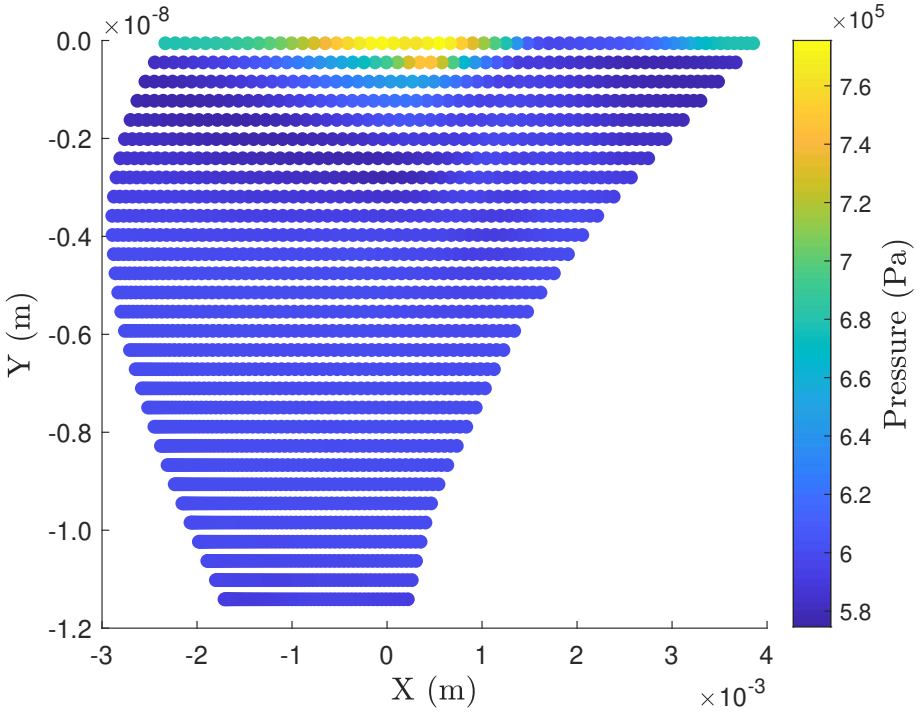


(a) Energy distribution of POD eigenmode for the pressure field in the inlet surface.

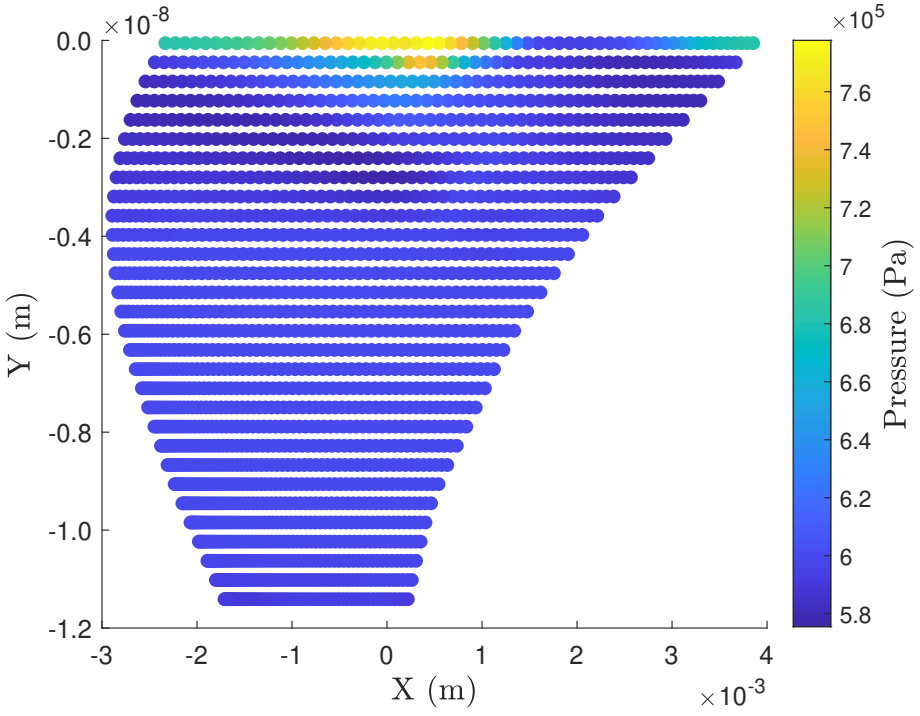


(b) Energy distribution of POD eigenmode for the pressure field in the outlet surface.

Figure A.5: For each subfigure the upper subplot depicts the energy distribution across each mode is depicted. The lower subplot illustrates the cumulative energy.

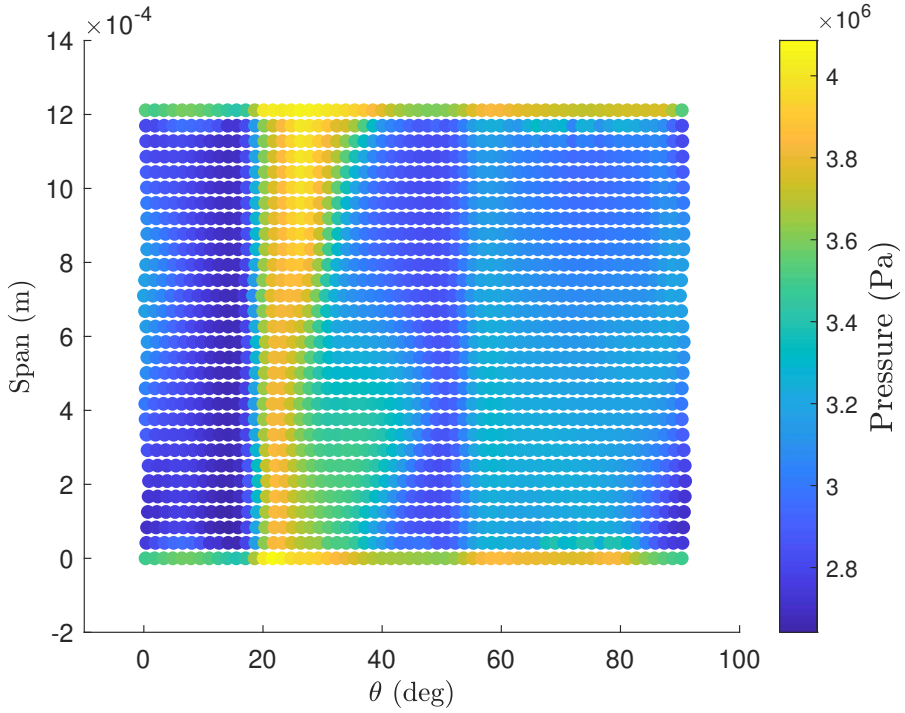


(a) CFD results for the total pressure at the inlet

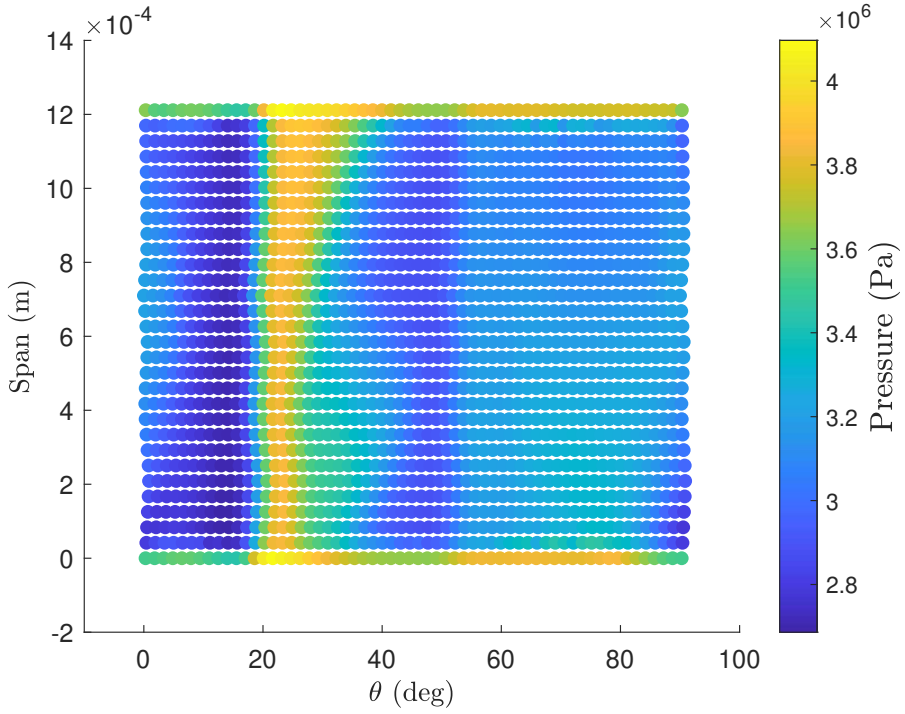


(b) POD results for the total pressure at the inlet

Figure A.6: Visual comparison of the CFD and POD results ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).



(a) CFD results for the pressure field at the outlet.



(b) POD results for the pressure field at the outlet.

Figure A.7: Visual comparison of the CFD and POD results ($m'_{le,sp} = 0.35$, $\psi_{sp} = 0.40$).

A.4 Main and splitter blades

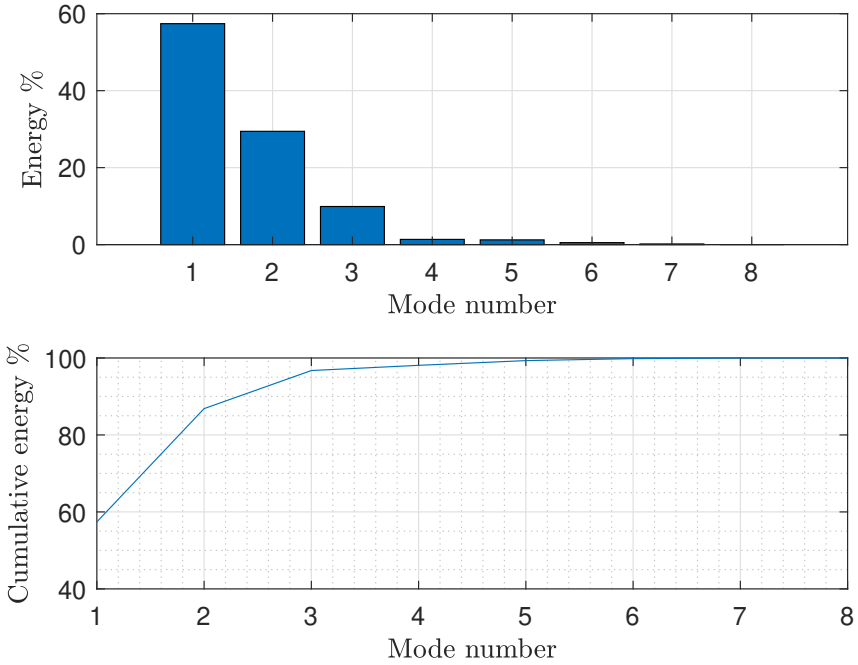


Figure A.8: Energy distribution of POD eigenmode for the velocity field on the main blade surface.

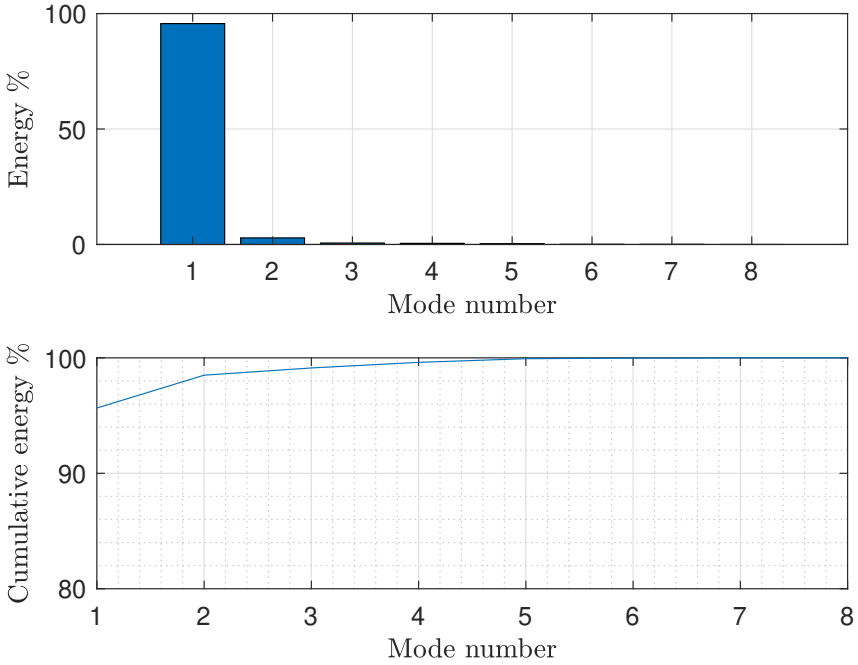


Figure A.9: Energy distribution of POD eigenmode for the velocity field on the splitter blade surface.

A.5 Modal visualization

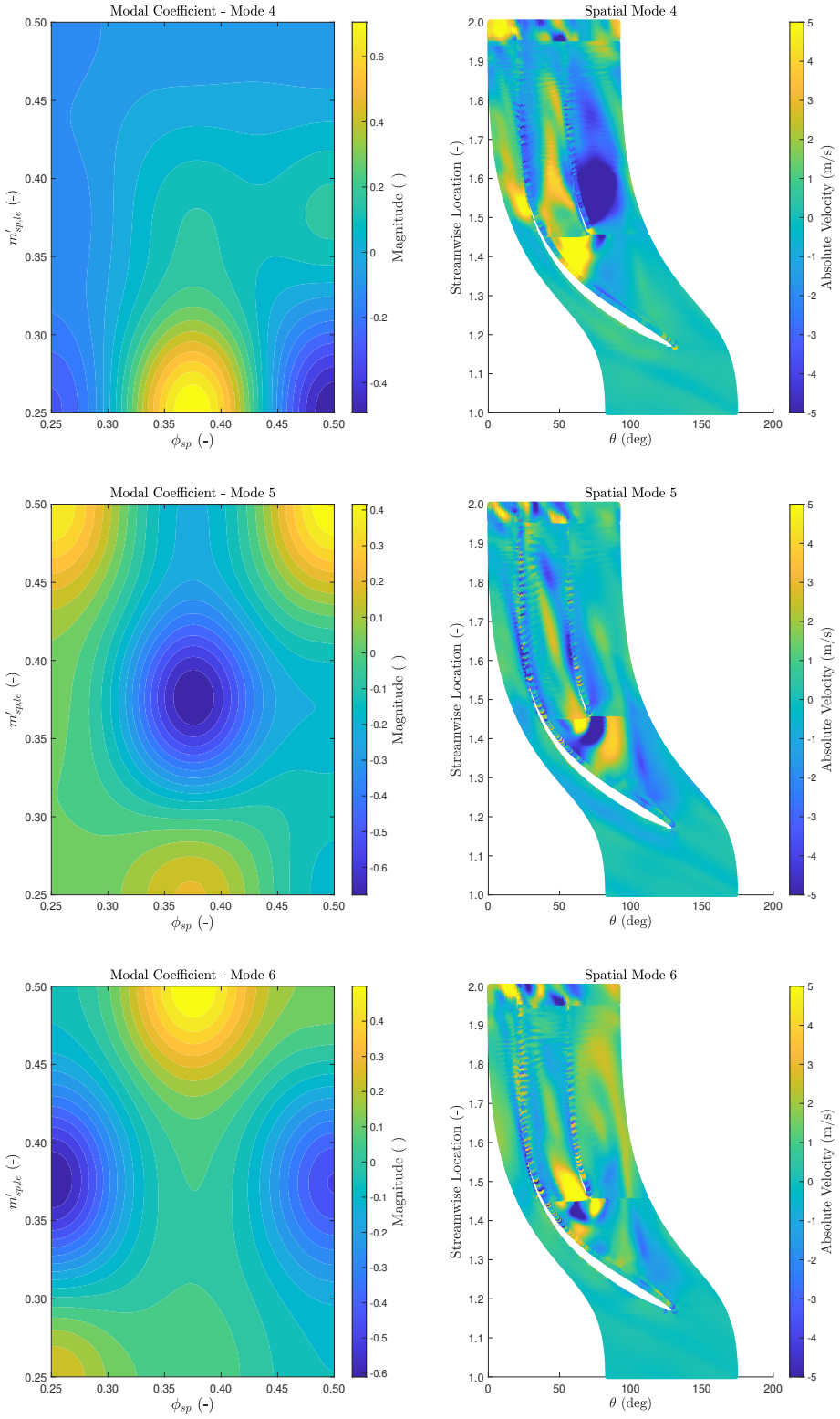


Figure A.10: Modes 4 to 6 obtained from the first POD model, of the blade to blade surface for the velocity field.

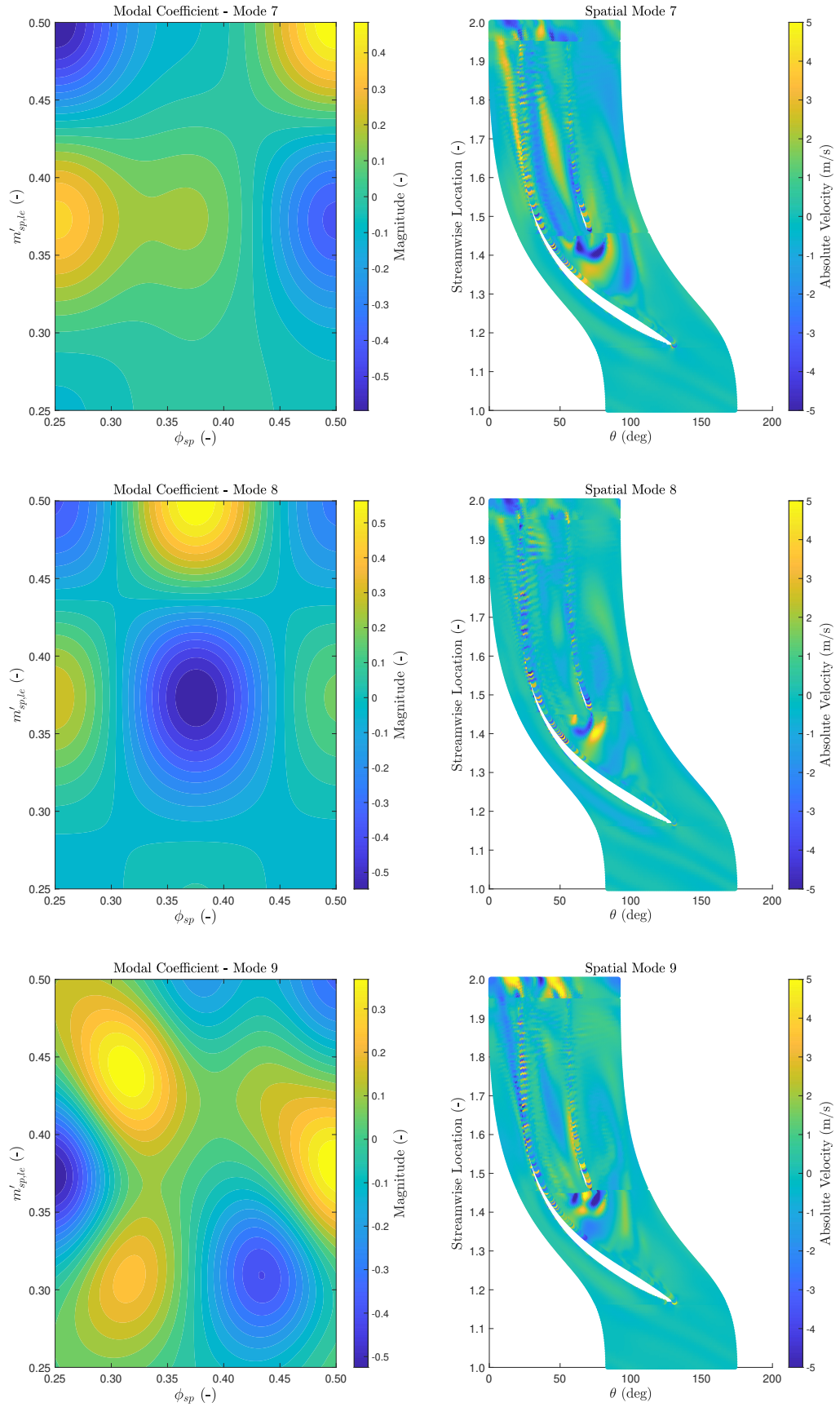
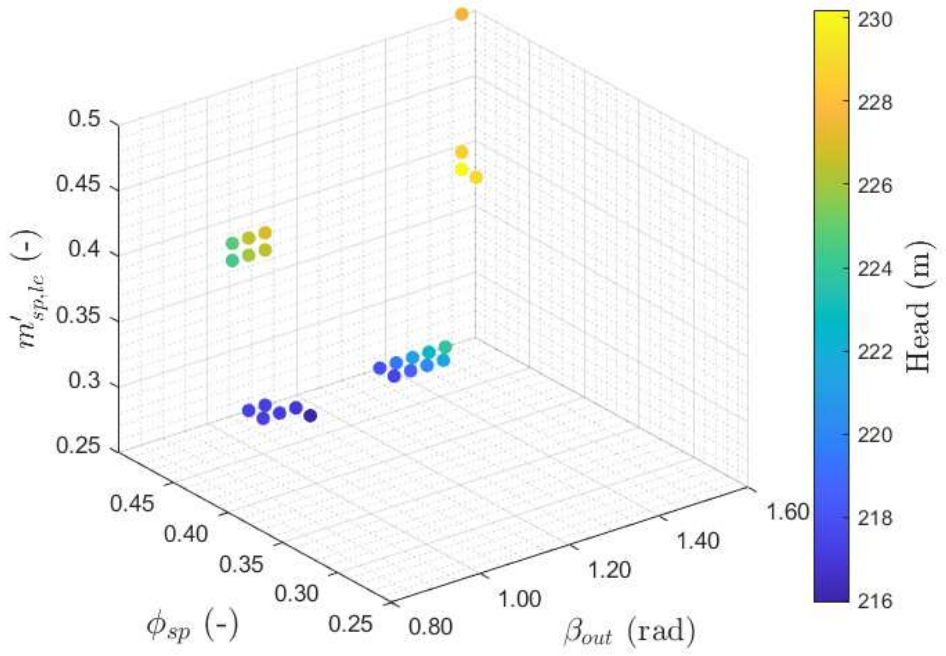


Figure A.11: Modes 7 to 9 obtained from the first POD model, of the blade to blade surface for the velocity field.

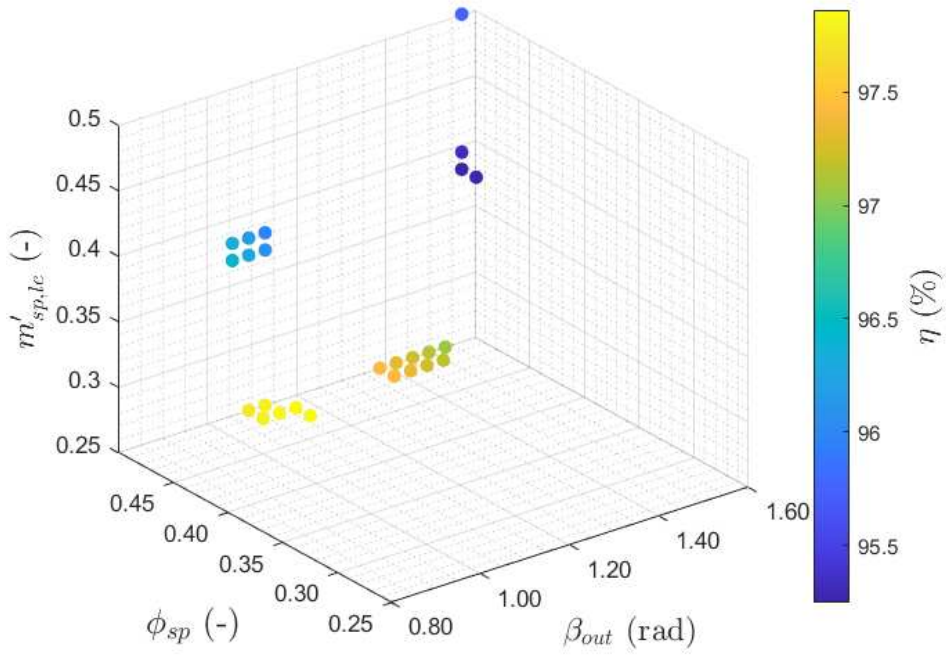
A.6 Four blades turbopump

Z	β_{out} (rad)	Tip clearance	ψ_{sp}	$m'_{le,sp}$
4	1.255	0.000	0.324	0.436
4	1.005	0.000	0.422	0.296
4	1.130	0.000	0.406	0.445
4	0.929	0.000	0.482	0.444
4	1.213	0.000	0.359	0.362
4	1.087	0.000	0.377	0.378
4	1.443	0.000	0.449	0.411
4	1.137	0.000	0.453	0.383
4	1.118	0.000	0.485	0.469
4	1.257	0.000	0.406	0.397
4	1.018	0.000	0.325	0.368
4	1.034	0.000	0.461	0.299
4	1.030	0.000	0.293	0.307
4	1.177	0.000	0.328	0.481
4	1.173	0.000	0.296	0.476
4	1.557	0.000	0.360	0.278
4	1.053	0.000	0.352	0.399
4	1.056	0.000	0.401	0.428
4	1.027	0.000	0.279	0.324
4	1.095	0.000	0.356	0.377

Table A.1: Turbopump test simulations parameters for a four blades configuration.



(a) Head predicition



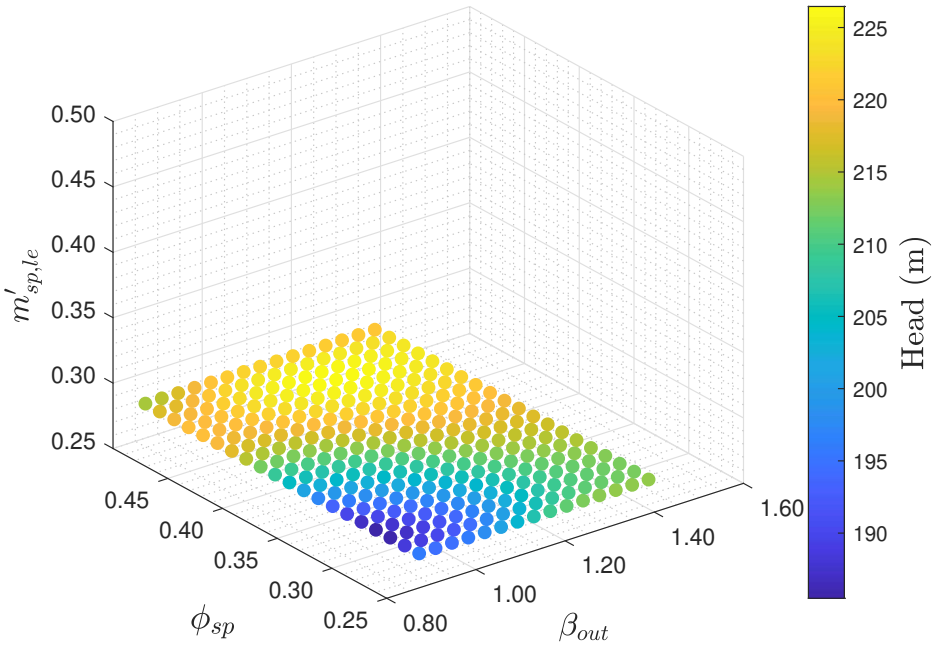
(b) Total efficiency prediction

Figure A.12: Non-dominated solution of the full population, identified with the Pareto front. Four blades turbopump.

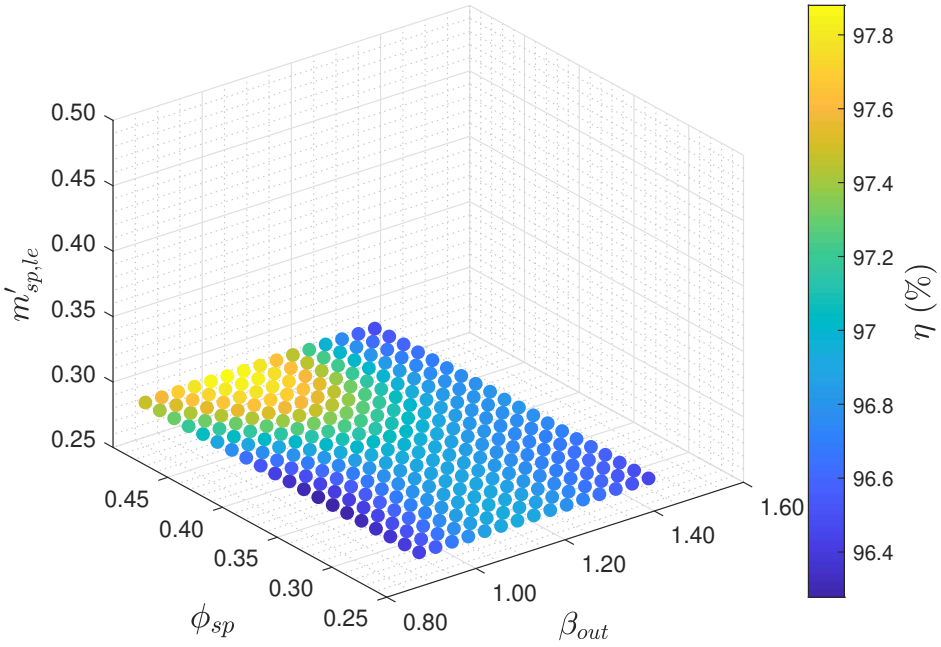
A.7 Five blades turbopump

Z	β_{out} (rad)	Tip clearance	ψ_{sp}	$m'_{le,sp}$
5	0.929	0.000	0.490	0.251
5	1.278	0.000	0.454	0.467
5	0.917	0.000	0.350	0.315
5	1.292	0.000	0.358	0.478
5	0.968	0.000	0.316	0.286
5	0.944	0.000	0.467	0.395
5	1.161	0.000	0.286	0.463
5	1.198	0.000	0.338	0.378
5	1.083	0.000	0.269	0.310
5	0.937	0.000	0.296	0.310
5	1.091	0.000	0.262	0.476
5	1.367	0.000	0.373	0.372
5	1.049	0.000	0.475	0.342
5	0.931	0.000	0.445	0.347
5	0.999	0.000	0.351	0.274
5	0.942	0.000	0.486	0.489
5	1.174	0.000	0.265	0.309
5	1.058	0.000	0.455	0.254
5	0.895	0.000	0.292	0.412
5	1.256	0.000	0.412	0.363

Table A.2: Turbopump test simulations parameters for a five blades configuration.

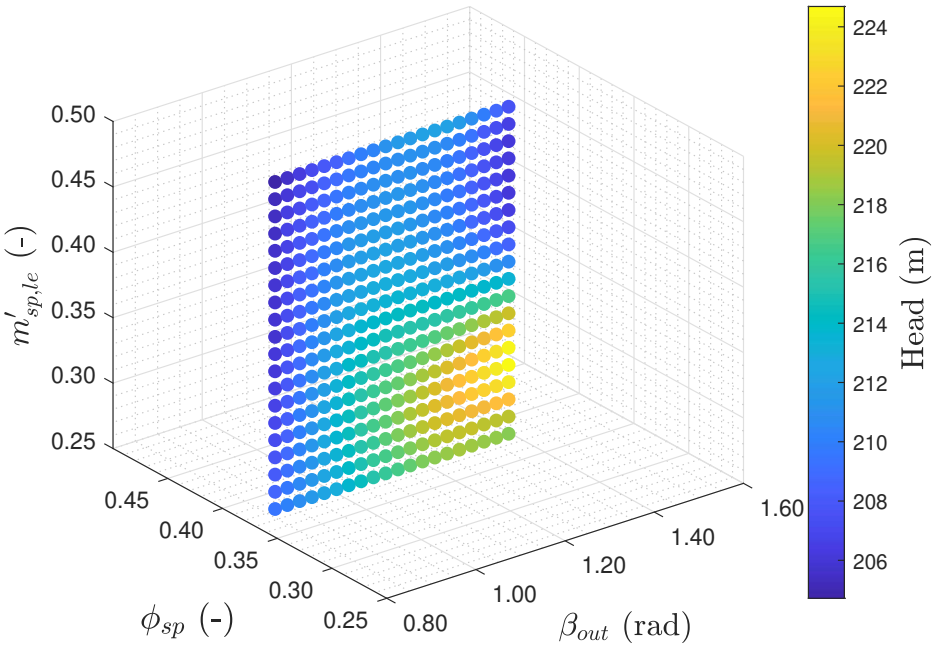


(a) Head prediction

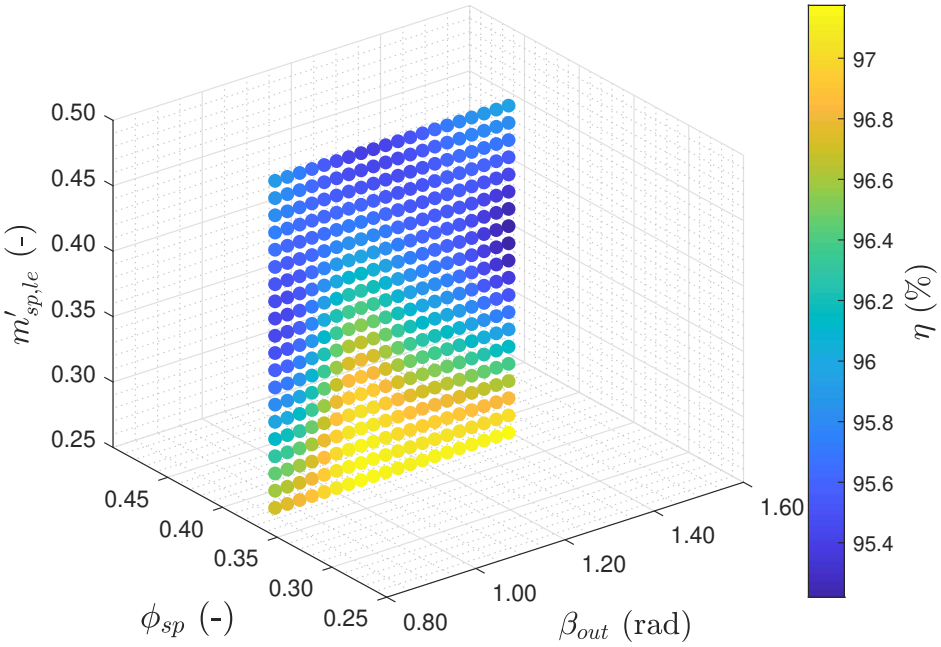


(b) Total efficiency prediction

Figure A.13: POD simulation for different splitter blade pitch position and outlet blade angle. Splitter blade leading edge position fixed to 0.27. Five blades turbopump.

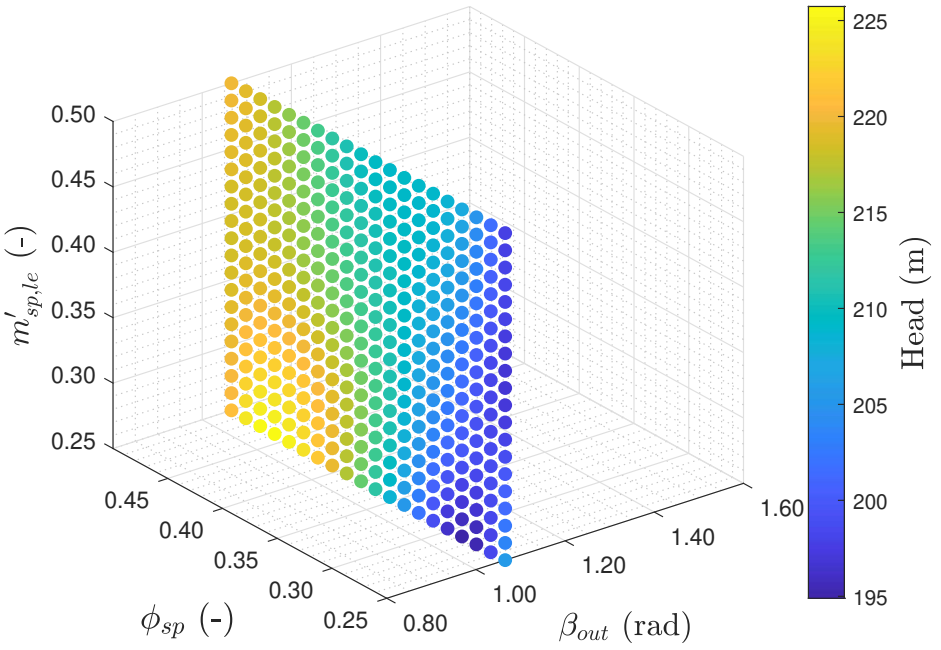


(a) Head prediction

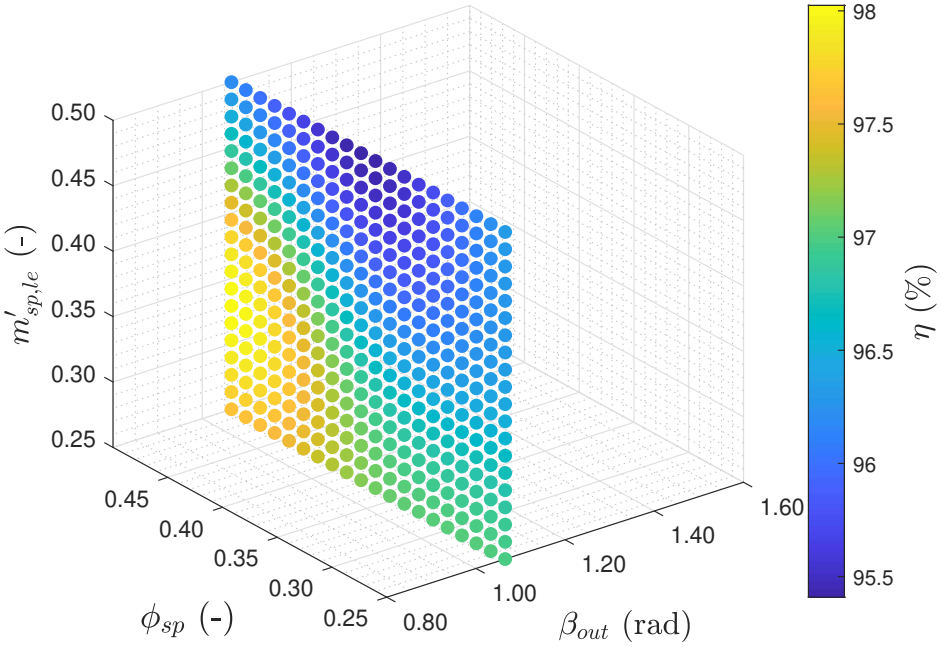


(b) Total efficiency prediction

Figure A.14: POD simulation for different splitter blade leading edge position and outlet blade angle. Splitter blade pitch position fixed to 0.38. Five blades turbopump.

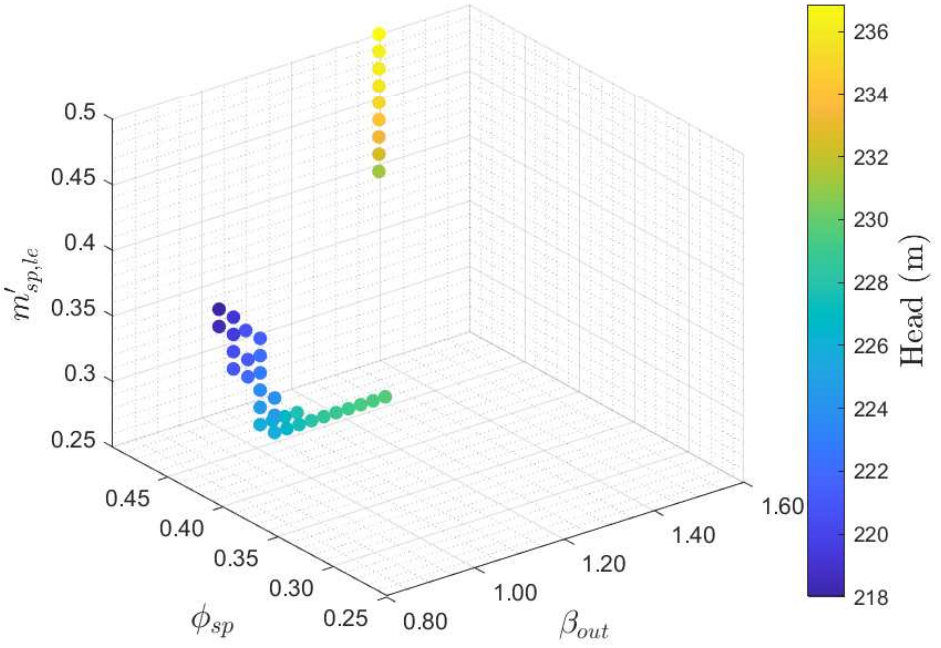


(a) Head prediction.

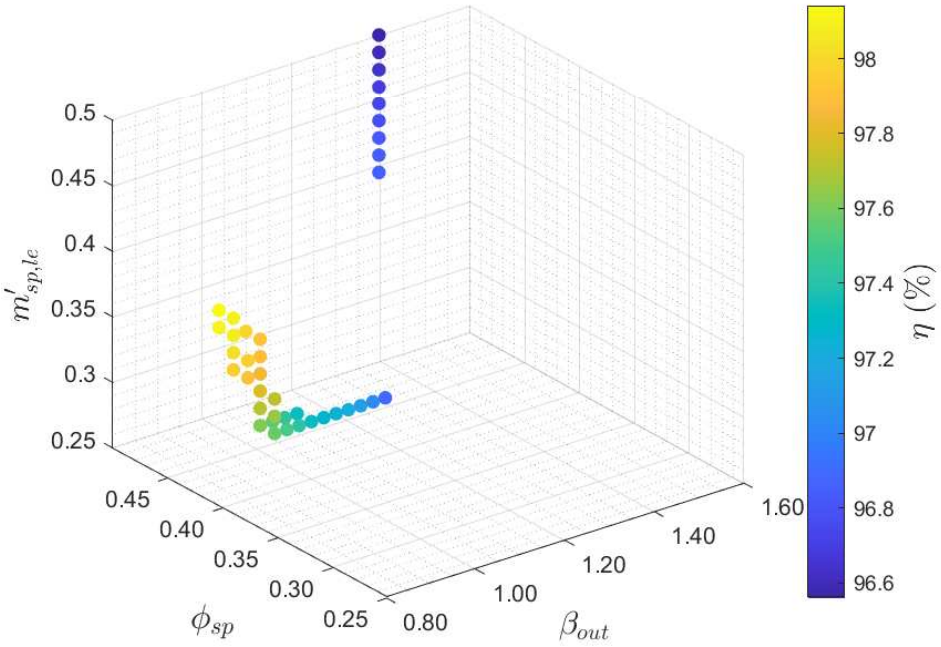


(b) Total efficiency prediction.

Figure A.15: POD simulation for different splitter blade leading edge position and pitch position. Blade outlet angle fixed to 60 deg. Five blades turbopump.



(a) Head prediction.



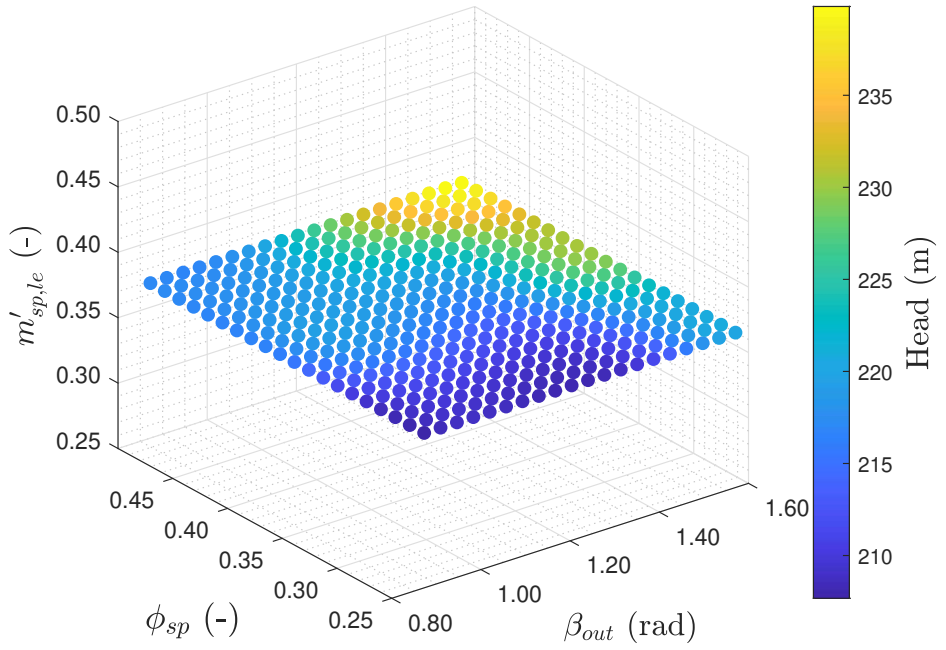
(b) Total efficiency prediction.

Figure A.16: Non-dominated solution of the full population, identified with the Pareto front. Five blades turbopump.

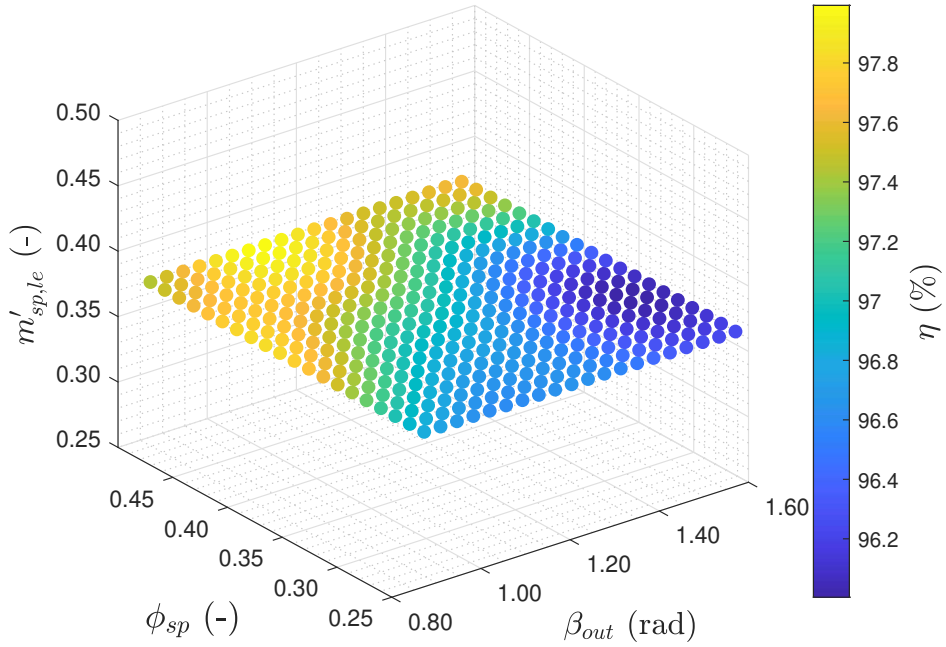
A.8 Six blades turbopump

Z	β_{out} (rad)	Tip clearance	ψ_{sp}	$m'_{le,sp}$
6	1.280	0.000	0.432	0.294
6	1.557	0.000	0.319	0.495
6	1.128	0.000	0.364	0.455
6	1.190	0.000	0.362	0.441
6	1.020	0.000	0.456	0.303
6	1.314	0.000	0.498	0.496
6	1.315	0.000	0.312	0.434
6	0.873	0.000	0.374	0.312
6	0.905	0.000	0.413	0.279
6	1.516	0.000	0.305	0.314
6	1.120	0.000	0.296	0.364
6	1.073	0.000	0.263	0.295
6	1.295	0.000	0.398	0.462
6	1.168	0.000	0.476	0.366
6	0.922	0.000	0.318	0.254
6	1.188	0.000	0.473	0.473
6	1.458	0.000	0.319	0.294
6	1.135	0.000	0.290	0.415
6	0.931	0.000	0.387	0.294
6	1.542	0.000	0.442	0.382

Table A.3: Turbopump test simulations parameters for a six blades configuration.

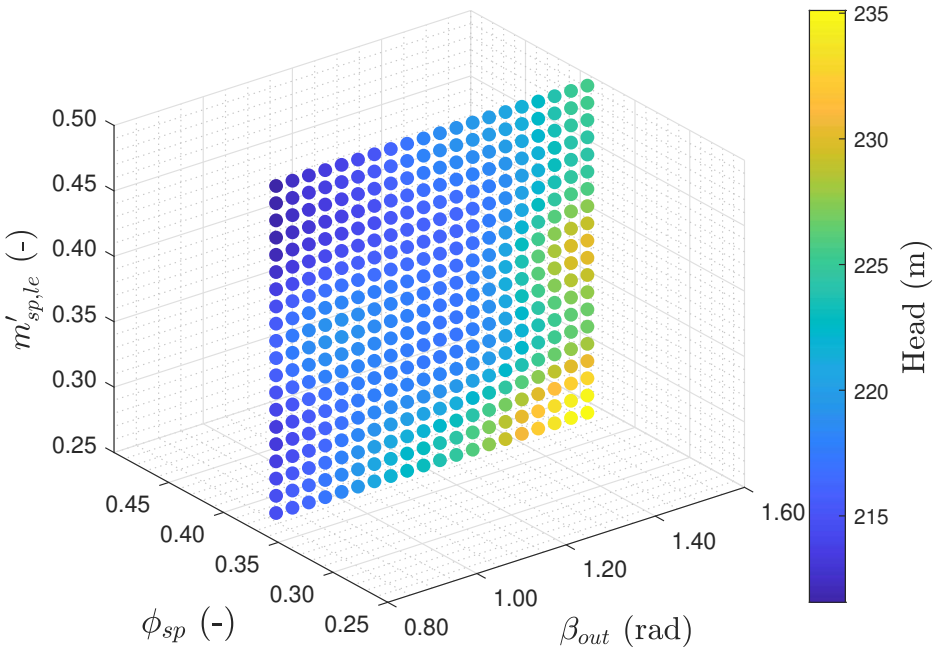


(a) Head prediction.

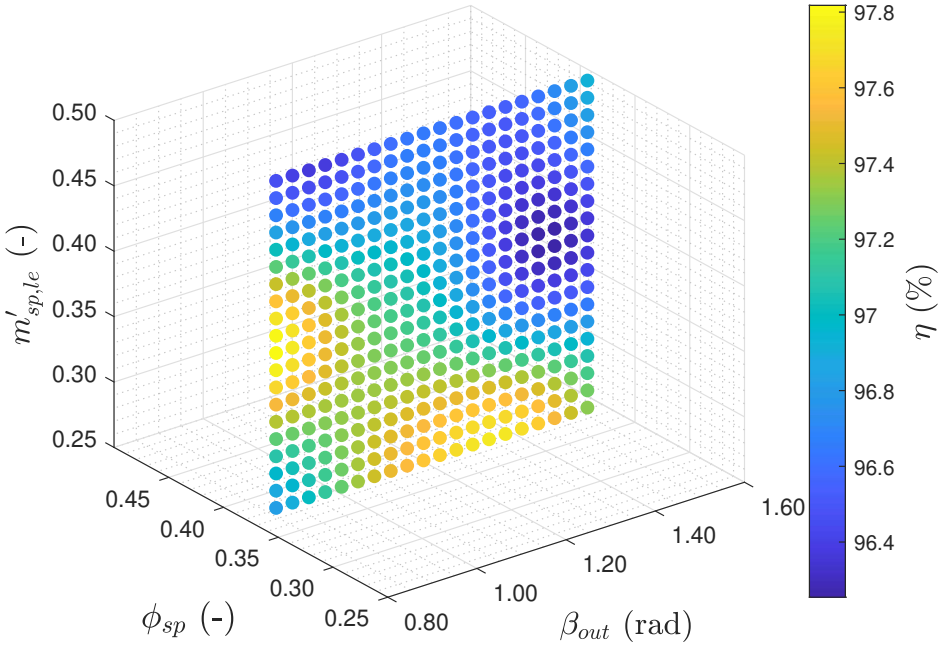


(b) Total efficiency prediction.

Figure A.17: POD simulation for different splitter blade pitch position and outlet blade angle. Splitter blade leading edge position fixed to 0.38. Six blades turbopump.

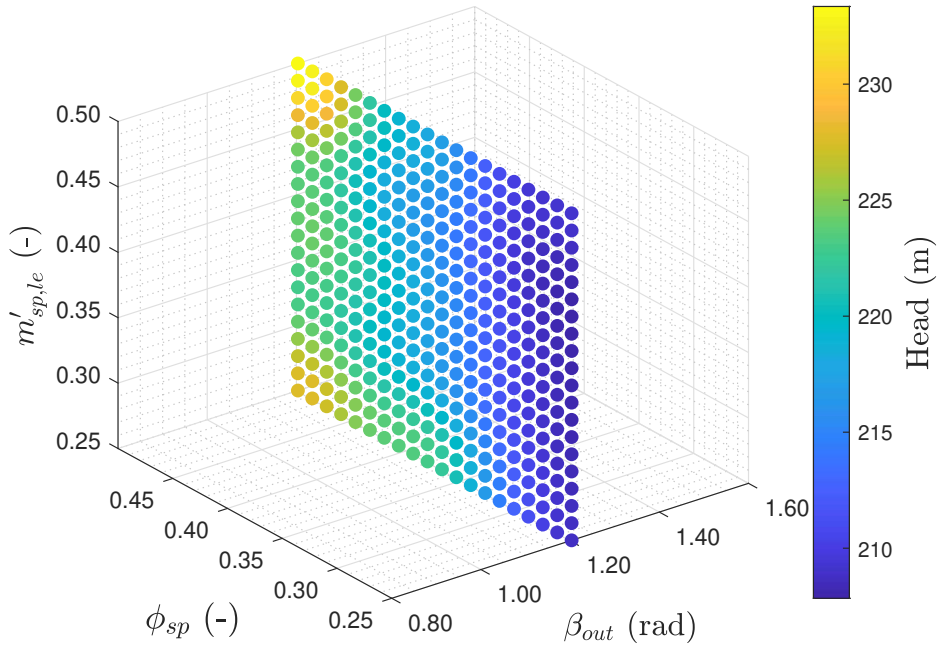


(a) Head prediction.

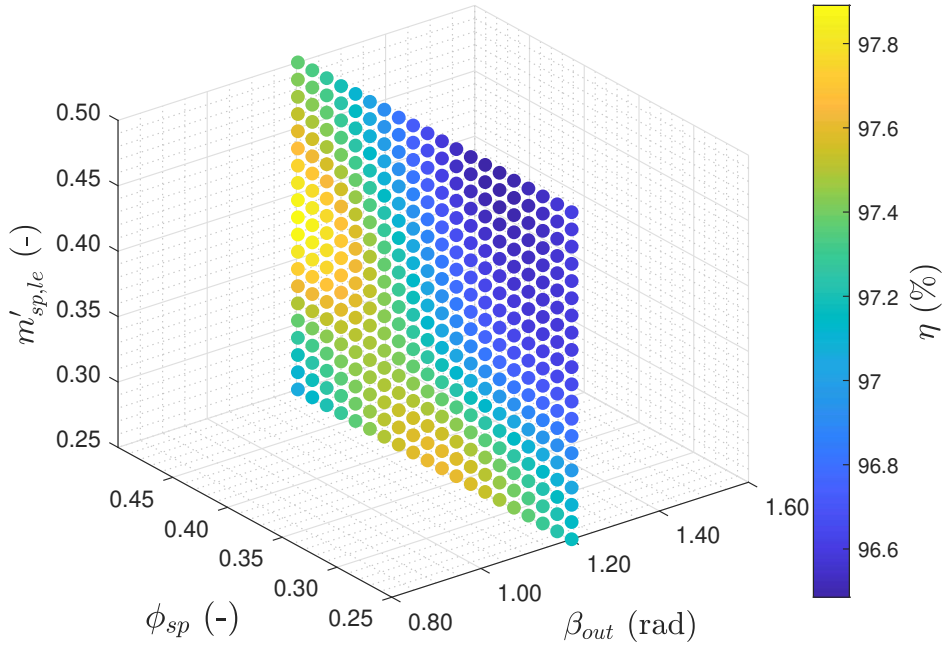


(b) Total efficiency prediction.

Figure A.18: POD simulation for different splitter blade leading edge position and outlet blade angle. Splitter blade pitch position fixed to 0.38. Six blades turbopump.

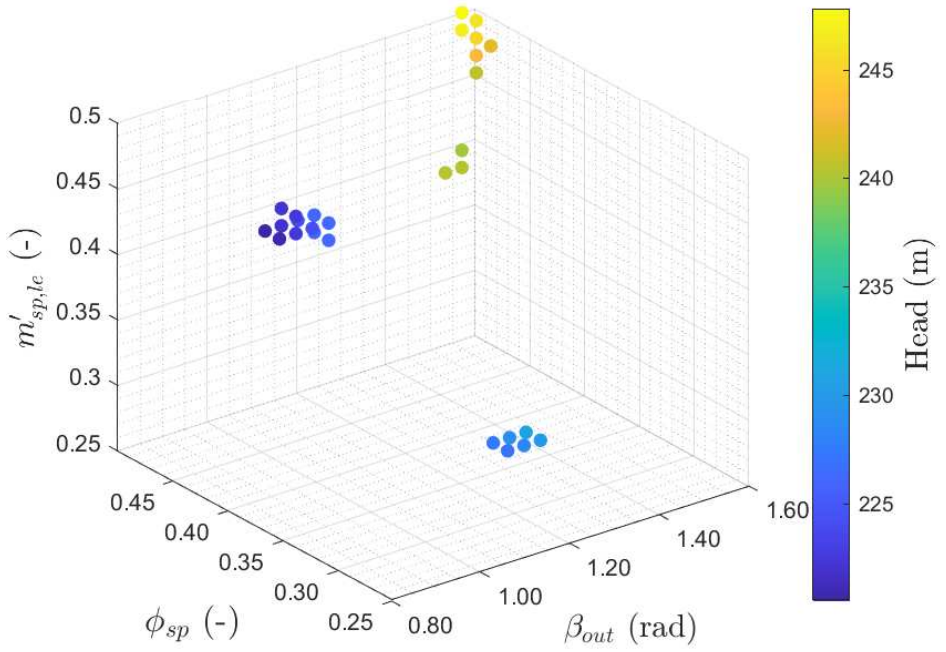


(a) Head prediction.

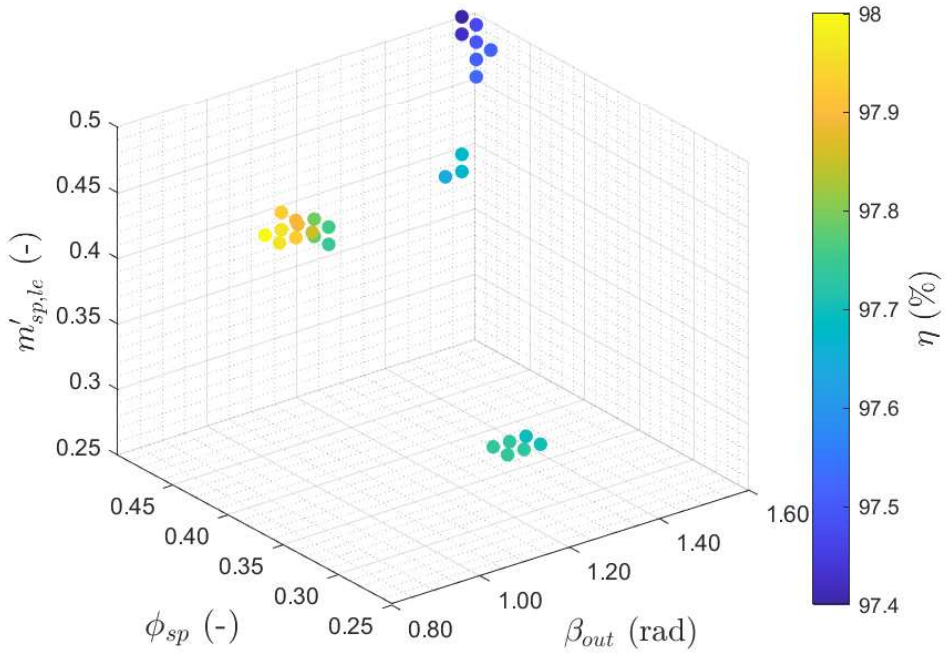


(b) Total efficiency prediction.

Figure A.19: POD simulation for different splitter blade leading edge position and pitch position. Blade outlet angle fixed to 70 deg. Six blades turbopump.



(a) Head prediction.



(b) Total efficiency prediction.

Figure A.20: Non-dominated solution of the full population, identified with the Pareto front. Six blades turbopump.

Università di Padova

Dipartimento di Ingegneria Industriale
CONTACTS Via Gradenigo, 6/a,
35131 Padova
Tlf. +45 4525 3031

segreteria.dii@unipd.it
https://dii.unipd.it/