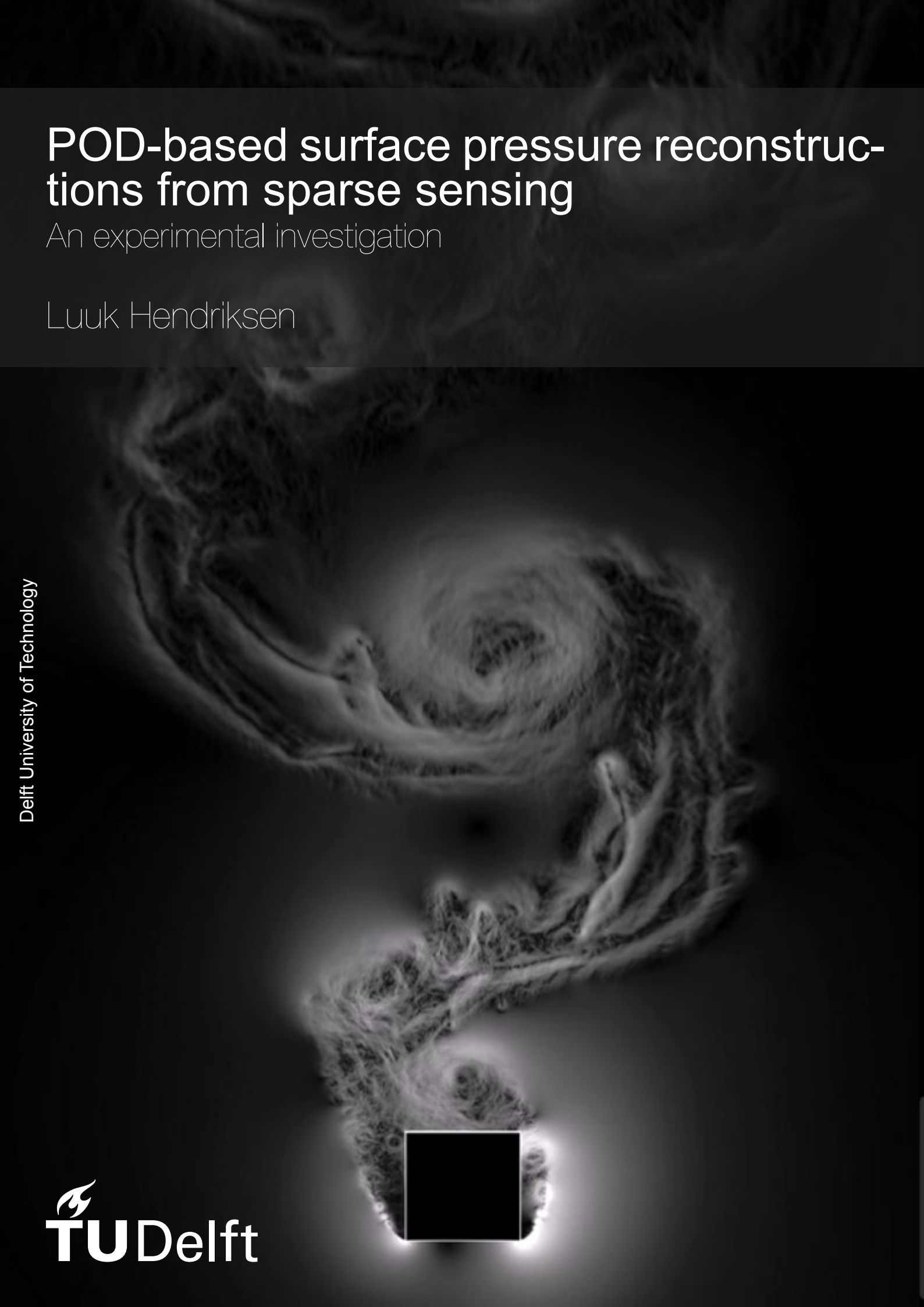
POD-based surface pressure reconstructions from sparse sensing

An experimental investigation

Luuk Hendriksen



Delft University of Technology

TU Delft

POD-based surface pressure reconstructions from sparse sensing

An experimental investigation

by

Luuk Hendriksen

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday December 20, 2022 at 1:00 PM.

Student number: 4654412
Project duration: January 14, 2022 – November 17, 2022
Thesis committee:

Dr. Andrea Sciacchitano,	TU Delft, Thesis supervisor
Prof. Dr. Fulvio Scarano,	TU Delft, Committee chair
Dr. Daniele Ragni,	TU Delft, External member
Dr. Anh Khoa Doan,	TU Delft, Additional member

Cover: Instantaneous pressure gradient magnitude of flow around a square cylinder at $Re = 22,000$. DNS study by Trias et al. (2015)

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Preface

Finalization of this report marks the culmination of a roughly ten-month-long master thesis project as part of the aerodynamics master at the Delft University of Technology. With overall feelings of happiness and fulfillment I look back at my, thus far, five year long university 'career' including the bachelor. I am grateful to have had the possibility to study in Delft and work with people from all over the world. With some ups and downs along the way, I can not help to feel a slight sense of relief but also pride as the day of my graduation approaches, an achievement I could not have envisioned five years ago.

First of all, I would like to express my sincere gratitude to Dr. Andrea Sciacchitano. As my daily thesis supervisor, he has provided me with this opportunity and has been a great source of information and advice throughout the project. His practical attitude and way of providing input and feedback has left me with a feeling of control and understanding of the direction for the thesis project. Not only in a purely academic sense has working with him been a pleasant experience, his interest and participation in thesis projects during the role as supervisor is encouraging. Also his patience is greatly appreciated when an hour long planned progress meeting for which I hoped to 'keep it short' approaches two hours instead. I hope his patience will last with a four-year-long PhD upcoming.

I would also like to thank the friends I made during both the bachelor and master. Coen and Yannick, who I have worked and laughed with and could rely upon from day 1 of my studies but who have since temporarily diverged into different countries for their studies. Of course, Kaj is also part of this list to which I would like to give special thanks. We have followed an identical path at the TU Delft up until this very day which I would not have been able to keep up with without your support and collaboration. During my thesis work you have dedicated a lot of your own time to helping me; from experiments in the weekend at the OJF to simulations on my laptop ("let's (not) trust the CFD results"). Because of this I could almost consider you as a second supervisor. Also thanks to Siddharth, Matthijs and Tom who I have met more recently during my master. Together, I feel like we were able to create a pleasant and productive atmosphere at the HSL basement which encouraged me to travel to Delft everyday instead of sitting in front of my laptop from home, allowing me to break with the boring 'work from home' COVID rut I was getting accustomed to.

Last but certainly not least, I would like to thank my family. My parents, from whom I have always felt an unconditional support and trust in everything that I do. My studies at the TU Delft has required an investment from you which exceeds any monetary value for which I am eternally grateful. My brother, who despite being different in almost every way I can always relate to and can always be myself around. Additionally, I would like to thank my grandparents in particular. Your interest in my studies in Delft has always been a source of motivation and together with the aforementioned people you are the reason that I come back to Rhenen every weekend to relax.

Luuk Hendriksen
Delft, December 2022

Abstract

A commonly used method for the determination of aerodynamic forces is through the integration of the surface pressure distribution over an object. The accuracy of this approach is limited by the ability of discrete measurements to approximate a continuous surface pressure distribution and thus typically requires a large number of surface pressure transducers to be used. This increases complexity, cost, weight, measurement latency and requires physical access all over the object. Reducing the required number of sensors might be desired in general however is especially useful for practical applications outside of the controlled environment of a wind tunnel.

When prior knowledge of a system is available in the form of 'training data', a modal decomposition method; Proper Orthogonal Decomposition (POD), can be used to identify patterns in the data that optimally characterize the behaviour of the system in a low-dimensional representation. By placing sparse surface pressure sensors such that only the most dominant features are sampled, full state surface pressure distributions can be reconstructed using an extension to the POD named Gappy POD (GPOD). The aim of this thesis is to assess the viability of using Gappy POD for the reconstruction of sparsely sampled experimental surface pressure distributions and obtain drag estimates on a typical bluff body being a square cylinder at varying angle of attack.

Training data has been obtained experimentally in the form of surface pressure measurements around the square cylinder perimeter using 32 pressure taps. This number of taps was found to be sufficient for accurate determination of the drag by comparison with force balance measurements. POD analysis of the pressure distributions decomposes the data into orthogonal modes ordered in terms of their energetic 'relevance' in representing the data at all angles of attack. The dominant modes are used in combination with a QR factorization including column pivoting algorithm (QR-CP) for the identification of optimal sparse sensor locations, suited for sparse surface pressure reconstructions.

Sparse surface pressure measurements on the optimal sensor locations are used to reconstruct pressure distributions outside of the training. Using GPOD, pressure distributions with up to 85 % of their data missing (5 out of 32 taps remaining) are able to be reconstructed with a Root Mean Squared Error in the order of 10^{-2} in the pressure coefficient (C_p) and corresponding drag coefficients (C_D) reach the same level of accuracy. In an attempt to reduce the reliance on wind tunnel measurements, uRANS simulations of the same square cylinder in settings similar to the wind tunnel campaign are performed and used for the determination of optimal sensor placement and training data. Experimental training data in combination with sensor locations determined through application of QR-CP on uRANS results shows the potential for sparse reconstructions with the same level of accuracy as mentioned before and greatly outperforms reconstructions based on 'non-tailored' sensor layouts. Additionally using

uRANS results as the basis for training eliminates the need for prior windtunnel campaigns and allows for direct reconstructions based on sparse measurements only. Unfortunately, uRANS simulations were not found to be accurate enough to provide training data for reconstructions as simple linear interpolation based on the sparse measurements was found to yield better results for a majority of the angle of attack range.

Overall, GPOD based on sparse surface pressure measurements on a bluff body is capable of producing accurate reconstructions and drag predictions with a reduced number sensors. The greedy QR-CP algorithm consistently provides sensor locations suited for reconstructions on a tailored basis while avoiding an unfeasible combinatorial sensor layout search. QR sensors based on CFD results are viable as an initial guess for experimental sensor placement but could not be shown to consistently provide reconstructions better than experimentally determined sensor layouts. Avoiding the need for experimental training through CFD simulations shows potential but was limited by the inaccuracy of the RANS simulations in the sparse GPOD reconstructions.

Contents

Preface iii

Abstract v

List of figures xvi

List of tables xvii

Nomenclature xix

1 Introduction 1

2 Literature and theoretical background 5
2.1 Sparse surface pressure and drag correlation . . . 5
2.2 Modal decomposition . . . 7
2.3 Snapshot POD . . . 9
2.4 POD for sparse sensing using tailored bases . . . 12
2.5 Gappy POD . . . 14
2.6 GPOD sensor placement . . . 16
2.7 Greedy QR-CP algorithm . . . 20
2.8 Reconstruction example . . . 23
2.8.1 Eigenvalues and modes from POD . . . 23
2.8.2 Sparse reconstruction using GPOD . . . 25
2.9 Square cylinder flow . . . 28
2.9.1 Flow regime classification . . . 28
2.9.2 Surface pressure distributions . . . 31
2.9.3 Drag coefficients . . . 32

3 Experimental setup and procedures 35
3.1 Wind tunnel . . . 35
3.2 Square cylinder test object . . . 36
3.3 Measurement equipment . . . 38
3.3.1 Pressure modules . . . 38
3.3.2 Force balance . . . 38
3.4 Experimental setup . . . 39
3.5 Data acquisition procedures . . . 41
3.5.1 Test matrix . . . 41
3.5.2 Angle of attack zeroing . . . 42
3.5.3 Force balance and surface pressure measurements . . . 42
3.5.4 Blockage velocity correction . . . 43

4 Square cylinder CFD 45
4.1 Numerical setup . . . 45
4.2 Example flowfields . . . 46
4.3 Surface pressure results . . . 47
4.4 Surface pressure POD analysis . . . 50

4.4.1 Eigenvalue spectrum . . . 50
4.4.2 Dominant eigenmodes . . . 51
4.5 Numerical surface pressure reconstruction using GPOD . . . 52
4.5.1 Reconstruction examples . . . 53
4.5.2 Pressure reconstruction accuracy . . . 54
4.6 Numerical drag estimation using GPOD . . . 56
4.6.1 Drag estimation examples . . . 56
4.6.2 Drag estimation accuracy . . . 57
4.7 CFD based sensor placement . . . 59
4.7.1 Tap distribution for pressure drag determination . . . 59
4.7.2 QR-CP sparse sensor locations . . . 60

5 Results and discussion 61
5.1 Surface pressure measurements . . . 61
5.1.1 Even and cosine spaced taps . . . 61
5.1.2 Opt V1 and Opt V2 taps . . . 65
5.1.3 Comparison to literature . . . 66
5.2 Drag measurements . . . 67
5.2.1 Force balance C_D versus α and comparison to literature . . . 67
5.2.2 Drag through force balance versus surface pressure integration 69
5.3 Surface pressure POD analysis . . . 70
5.3.1 Eigenvalue spectrum . . . 70
5.3.2 Dominant eigenmodes . . . 71
5.4 Experimental surface pressure reconstruction using GPOD . . . 72
5.4.1 Reconstruction examples . . . 72
5.4.2 Pressure reconstruction accuracy . . . 75
5.5 Experimental drag estimation using GPOD . . . 78
5.5.1 Drag estimation examples . . . 78
5.5.2 Drag estimation accuracy . . . 79
5.6 GPOD error sources . . . 81
5.7 GPOD versus linear interpolation . . . 83
5.7.1 Linear interpolation reconstruction examples . . . 83
5.7.2 Comparison of surface pressure reconstruction accuracy . . . 85
5.7.3 Comparison of drag estimation accuracy . . . 86
5.8 CFD based sensors GPOD reconstructions . . . 87
5.8.1 Opt V2 reconstruction examples . . . 87
5.8.2 Opt V2 pressure reconstruction accuracy . . . 90
5.8.3 Opt V2 drag estimation accuracy . . . 91
5.9 CFD trained experimental surface pressure reconstruction using GPOD 93
5.9.1 Reconstruction examples . . . 93
5.9.2 Pressure reconstruction accuracy . . . 96
5.9.3 Drag estimation accuracy . . . 98

6 Conclusions and recommendations 101
6.1 Conclusions . . . 101
6.2 Recommendations . . . 105
6.2.1 Changes to the experimental pressure tap layouts . . . 105

6.2.2 Improvement of training data 106
6.2.3 Oversampling 106

Bibliography **109**

A Cosine results **115**

B Experimental even spacing POD modes and coefficients **119**

C GPOD source code **129**

List of Figures

1.1 Surface pressure measurements by means of pressure taps and tubing connected to transducers/scanners on a building section (left) and cone shaped model (right). Instrumenting the models requires meticulous model preparation and space for the tubing and equipment in or outside of the model. Left figure courtesy of CPP (2022), right figure CSIR-NAL (2022). 1

1.2 Campaniform sensilla (effectively strain sensor) distributions on the wings of different insect species, indicated by dots. Sensor distributions are mathematically sparse but provide mechanosensory feedback for robust flight control through nonarbitrary placement. Adapted from Aiello et al. (2021). 2

2.1 Linear regression approach towards the reconstruction of loads from sparse surface pressure measurements Burelle et al. (2020) 6

2.2 Data points and direction of eigenvectors 7

2.3 Experimental setup used by Mohammed-Taifour and Weiss (2016) (left) and first mode of the horizontal velocity component (right) identifying correlation in the flow. Figures adapted from Weiss (2019) 8

2.4 POD based outlier detection for PIV (Wang et al., 2015) applied to channel flow (top) and isotropic flow (bottom). From left to right: original longitudinal velocity field, velocity field including 5 % synthetic outliers and POD outlier corrected velocity field. Adapted from Higham et al. (2016) 11

2.5 Fourier image compression, adapted from Manohar et al. (2018) 12

2.6 Steady state solution of laminar lid driven cavity flow obtained by Lorenzi et al. (2016) using POD-Garlerkin ROM (left) and benchmark solution from Botella and Peyret (1998) (right) 13

2.7 Example eigenfaces obtained from application of POD procedure to Yale database, figure adapted from Brunton and Kutz (2019). 14

2.8 Adaptive gappy POD for gappy PIV image reconstruction, applied to turbulent channel flow velocity field. Raben et al. (2012) 16

2.9 C_p distribution at different section locations η along the wing span of a 3D wing-body aircraft configuration at $\alpha = 7.1°$ and $M = 0.85$, adapted from Mifsud et al. (2019) 17

2.10 3D aircraft wing-body model indicating sections along which experimental data is available (left), resulting surface pressure distribution from GPOD; fused experimental and CFD data (right) at $\alpha = 7.9°$ and $M = 0.85$, adapted from Mifsud et al. (2019) 18

2.11 Gappy POD LNG carrier drag estimation results using 32 sensors and different sensor layouts: layout 1 left, layout 2 right. Adapted from Xing et al. (2022) 19

2.12 Reconstruction performance of gappy POD on test case using MCN and mode extrema sensor locations, adapted from Brunton and Kutz (2019) . . . 20
2.13 Definition of matrix Θ using 'sensor matrix' $\boldsymbol{C}$ . . . 21
2.14 Schematic illustration of QR factorization with column pivoting for the determination of $\boldsymbol{C}^T$ . . . 22
2.15 Relation between matrix volume and determinant . . . 22
2.16 Individual (a) and cumulative (b) energy capture of the eigenvalues corresponding to U-velocity cylinder flow data . . . 24
2.17 First and second spatial POD modes (left) and corresponding temporal coefficients (right) for cylinder flow . . . 25
2.18 Instantaneous DNS solution of horizontal velocity field (top left), rank 5 POD projection (top right), gappy POD reconstruction using first 5 POD bases + 5 random sensors (bottom left) and schematic gappy POD matrix $\boldsymbol{M}$ with its condition number κ (bottom right). . . . 26
2.19 Instantaneous DNS solution of horizontal velocity field (top left), rank 5 POD projection (top right), gappy POD reconstruction using first 5 POD bases + 5 QR sensors (bottom left) and schematic gappy POD matrix $\boldsymbol{M}$ with its condition number κ (bottom right). . . . 27
2.20 Reconstruction root mean square error, random and QR sensor placement . . . 27
2.21 Square cylinder angle of attack and face definitions . . . 28
2.22 Subcritical- (left), superciritcal- (middle) and wedge- (right) flow patterns. For information regarding indicated locations, see Huang et al. (2010) . . . 29
2.23 Subcritircal flow patterns as obtained by van Oudheusden et al. (2008) at $Re = 20,000$ . . . 30
2.24 Flow pattern for $\alpha = 30°$ (left) and $\alpha = 45°$ (right, background colours follow from vorticity) at $Re = 20,000$, adapted from Roosenboom (2005) 30
2.25 Square cylinder surface pressure distributions at various angles of attack, adapted from Igarashi (1984) . . . 31
2.26 Drag coefficient versus angle of attack for square cylinder flow . . . 32

3.1 Schematic view of the Open Jet Facility at the Delft University of technology (Delft University of Technology, 2022) (a) and velocity profile in the shear layer at 10 cmfromtheoutlet(b) . . . 36
3.2 Side view drawing including main dimensions (in mm) of the square cylinder test object (left) and removable pressure tap plate (right). . . . 36
3.3 Cosine sampling for tap locations skewed towards square cylinder edges 37
3.4 Tap locations for different plate sets, including local coordinate system and angle/face definitions. . . . 38
3.5 Schematic of the experimental setup including main components and dimensions (in mm) . . . 40

4.1 Schematic CFD setup . . . 46
4.2 uRANS time averaged velocity fields including streamlines at angles of attack in both the sub- and supercritical flow regime. . . . 47

4.3 uRANS time averaged pressure distributions for $\alpha = 0°$ to $20°$ (a) and $\alpha = 25°$ to $45°$ (b) . . . 48
4.4 uRANS pressure distribution comparison to experimental results at $\alpha = 0°$ (a) and uRANS drag curve comparison to experimental results (b) . . 49
4.5 uRANS pressure distribution comparison at $\alpha = 45°$ to experimental and uRANS results from Mueller (2012) (left) and flowfield visualization at $\alpha = 45°$ of the uRANS results from Mueller (2012) (right) . . . 49
4.6 Individual (a) and cumulative (b) energy capture of the eigenvalues corresponding to CFD pressure distributions . . . 51
4.7 First and second spatial eigenmodes (left), first and second α-coefficients (right) for CFD square cylinder case . . . 52
4.8 CFD GPOD reconstruction at $\alpha = 12.5°$, $n_s = K = 10$ . . . 53
4.9 CFD GPOD reconstruction at $\alpha = 42.5°$, $n_s = K = 10$ . . . 53
4.10 CFD GPOD pressure distribution reconstruction (top) and projection (bottom) accuracy for every combination of $n_s = K$ and α . . . 54
4.11 CFD GPOD surface pressure reconstruction accuracy for increasing number of modes and sensors . . . 55
4.12 CFD GPOD drag curve reconstruction using $n_s = K = 10$ . . . 56
4.13 CFD GPOD drag curve reconstruction accuracy for increasing number of modes and sensors . . . 58
4.14 RMSE in the CFD C_D for even or cosine spaced pressure taps . . . 59

5.1 Experimental time averaged pressure distributions for $\alpha = 0°$ to $20°$ (a) and $\alpha = 25°$ to $45°$ (b), evenly spaced taps. Convention for angles ψ and α shown in bottom left of (b) . . . 62
5.2 Experimental surface pressure distributions around the critical angle of attack, evenly spaced taps . . . 63
5.3 Experimental time averaged pressure distributions for $\alpha = 0°$ to $20°$ (a) and $\alpha = 25°$ to $45°$ (b), cosine spaced taps . . . 64
5.4 Experimental time averaged pressure tap measurements for $\alpha = 0°$ to $\alpha = 45°$, Opt V1 taps . . . 65
5.5 Experimental time averaged pressure tap measurements for $\alpha = 0°$ to $\alpha = 45°$, Opt V2 taps . . . 65
5.6 Current experiment pressure distribution comparison to available experimental result at $\alpha = 0°$ (a) and $\alpha = 12.5°$ . . . 66
5.7 Drag coefficient versus angle of attack as measured by force balance during experimental campaign. Including reference results from literature 68
5.8 Integrated pressure drag compared to balance drag for even (a) and cosine (b) spaced taps. Root Mean Squared Difference (RMSD) between pressure tap and balance values reported in the box at the top left. . . . 69
5.9 Individual (a) and cumulative (b) energy capture of the eigenvalues corresponding to experimental pressure distributions . . . 70
5.10 First and second spatial POD mode (left), first and second α-coefficients (right) from experimental results . . . 71
5.11 Experimental GPOD reconstruction at $\alpha = 12.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$ . . . 73

5.12 Experimental GPOD reconstruction at $\alpha = 42.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$ 73
5.13 Experimental GPOD reconstruction at $\alpha = 32.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$ 74
5.14 Experimental GPOD pressure distribution reconstruction (top) and projection (bottom) accuracy for every combination of $n_s = K$ and α 76
5.15 Experimental GPOD surface pressure reconstruction accuracy for increasing number of modes and sensors 77
5.16 Sensor locations determined by QR algorithm for each value of $n_s = K$ 77
5.17 Experimental GPOD drag curve reconstruction. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$ 78
5.18 Experimental GPOD drag curve reconstruction accuracy for increasing number of modes and sensors 80
5.19 Schematic RMSE C_p plots representing sources of error: overall training inadequacy (left), sensor placement (middle) and training data skewness (right) 81
5.20 Experimental GPOD and LERP reconstruction at $\alpha = 17.5°$, $n_s = K = 5$. 83
5.21 Experimental GPOD and LERP reconstruction at $\alpha = 17.5°$, $n_s = K = 10$. 84
5.22 Experimental GPOD and linear interpolation reconstruction accuracy . . 85
5.23 Experimental GPOD and linear interpolation drag curve reconstruction accuracy 86
5.24 Opt V2 GPOD reconstruction at α = $25°$, $n_s = K = 10$ 88
5.25 Opt V2 GPOD reconstruction at α = $27.5°$, $n_s = K = 10$ 88
5.26 Opt V2 based reconstruction accuracy w.r.t. reference pressure distributions 90
5.27 Opt V2 GPOD drag curve reconstruction using $n_s = K = 10$ 92
5.28 CFD trained GPOD reconstruction using experimental Opt V2 sensor measurements compared to cosine reference- and CFD pressure distribution at $\alpha = 0°$, $n_s = K = 10$. 94
5.29 CFD trained GPOD reconstruction using experimental Opt V2 sensor measurements compared to cosine reference- and CFD pressure distribution at $\alpha = 12.5°$, $n_s = K = 10$. 95
5.30 CFD trained GPOD reconstruction using experimental Opt V2 sensor measurements compared to cosine reference- and CFD pressure distribution at $\alpha = 45°$, $n_s = K = 10$. 96
5.31 RMSE C_p to cosine reference achieved by CFD trained GPOD reconstructions using experimental Opt V2 sensor measurements. Also included: RMSE C_p of the CFD surface pressure distributions to the same cosine reference as well as the RMSE C_p achieved by linear interpolation (LERP) using the same Opt V2 sensors. 97
5.32 C_D versus α curve obtained by CFD trained GPOD reconstructions using experimental Opt V2 sensor measurements. Also included: CFD drag curve, drag curve obtained by linear interpolation (LERP) using the same Opt V2 sensors and the reference drag curve obtained through experimental balance measurements. 98

A.1 Experimental GPOD reconstruction at $\alpha = 12.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$, using cosine spacing pressure distributions. . . 115
A.2 Experimental GPOD reconstruction at $\alpha = 42.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$, using cosine spacing pressure distributions. . . 115
A.3 Experimental GPOD pressure distribution reconstruction (top) and projection (bottom) accuracy for every combination of $n_s = K$ and α, using cosine spacing pressure distributions. 116
A.4 Experimental GPOD surface pressure reconstruction accuracy for increasing sensor budget, using cosine spacing pressure distributions. . . 116
A.5 Sensor locations determined by QR algorithm for each value of $n_s = K$, using cosine spacing pressure distributions. 117
A.6 Experimental GPOD drag curve reconstruction. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$, using cosine spacing pressure distributions. . . 117
A.7 Experimental GPOD drag curve reconstruction accuracy for increasing sensor budget, using cosine spacing pressure distributions. 117

B.1 Spatial POD mode and α-coefficients 1 from pressure distributions using even tap spacing 119
B.2 Spatial POD mode and α-coefficients 2 from pressure distributions using even tap spacing 119
B.3 Spatial POD mode and α-coefficients 3 from pressure distributions using even tap spacing 120
B.4 Spatial POD mode and α-coefficients 4 from pressure distributions using even tap spacing 120
B.5 Spatial POD mode and α-coefficients 5 from pressure distributions using even tap spacing 121
B.6 Spatial POD mode and α-coefficients 6 from pressure distributions using even tap spacing 121
B.7 Spatial POD mode and α-coefficients 7 from pressure distributions using even tap spacing 122
B.8 Spatial POD mode and α-coefficients 8 from pressure distributions using even tap spacing 122
B.9 Spatial POD mode and α-coefficients 9 from pressure distributions using even tap spacing 123
B.10 Spatial POD mode and α-coefficients 10 from pressure distributions using even tap spacing 123
B.11 Spatial POD mode and α-coefficients 11 from pressure distributions using even tap spacing 124
B.12 Spatial POD mode and α-coefficients 12 from pressure distributions using even tap spacing 124
B.13 Spatial POD mode and α-coefficients 13 from pressure distributions using even tap spacing 125
B.14 Spatial POD mode and α-coefficients 14 from pressure distributions using even tap spacing 125
B.15 Spatial POD mode and α-coefficients 15 from pressure distributions using even tap spacing 126

B.16 Spatial POD mode and α-coefficients 16 from pressure distributions using even tap spacing 126
B.17 Spatial POD mode and α-coefficients 17 from pressure distributions using even tap spacing 127
B.18 Spatial POD mode and α-coefficients 18 from pressure distributions using even tap spacing 127
B.19 Spatial POD mode and α-coefficients 19 from pressure distributions using even tap spacing 128

List of Tables

3.1 OJF external balance load- and moment- ranges 39
3.2 OJF external balance accuracy in percentage of applied maximum calibration load 39
3.3 Condensed test matrix 41
3.4 Measured drag coefficients used for determination of the end plates drag contribution 43

4.1 CFD setup parameters 46
4.2 Opt V1 and Opt V2 sensor locations given by sensor angle ψ. 60

6.1 Summarized conclusion on the experimental GPOD reconstruction results depending on the source of the sensor layout and POD modes. RMSE C_p and RMSE C_D at $n_s = K = 10$ ($69\,\%$ gappiness) stated in last columns for reference. 104

Nomenclature

Abbreviations

Abbreviation	Definition
CAD	Computer-Aided Design
CFD	Computational Fluid Dynamics
CPU	Central Processing Unit
DEIM	Discrete Empirical Interpolation Method
DNS	Direct Numerical Simulation
FDM	Fused Deposition Modeling
FOM	Full Order Model
GPOD	Gappy Proper Orthogonal Decomposition
JHTDB	John Hopkins Turbulence Database
JPEG	Joint Photographic Experts Group
KL	Karhunen-Loève
LERP	Linear Interpolation
LES	Large Eddy Simulation
LNG	Liquefied Natural Gas
MCN	Minimum Condition Number
OJF	Open Jet Facility
PIV	Particle Image Velocimetry
PLA	Polylactic Acid
POD	Proper Orthogonal Decomposition
POD-OC	Proper Orthogonal Decomposition based Outlier Detection
QR-CP, QDEIM	QR factorization with column pivoting for sensors locations
RANS	Reynolds-Averaged Navier-Stokes
RMSE	Root Mean Squared Error
ROM	Reduced Order Model
SST	Sea Surface Temperature, Shear Stress Transport
SVD	Singular Value Decomposition
TSB	Turbulent Separation Bubble
uRANS	Unsteady RANS
UAV	Unmanned Aerial Vehicle

Symbols

Symbol	Definition
a	POD coefficient
A	Square cylinder reference area, $H \times L$

Symbol	Definition
b	GPOD coefficient
c	Speed of sound
C_D	Drag coefficient
C_p	Pressure coefficient
E_k	Turbulent kinetic/modal energy capture
$E_{k,cumulative}$	Cumulative turbulent kinetic/modal energy capture
f	Frequency
f_s	Vortex shedding frequency
$F_{x,y,z}$	Force balance axial, side, vertical force
H	Square cylinder height
k	Turbulent Kinetic Energy
K	Number of used POD modes
l	Pressure tap tube length
L	Square cylinder side length
m	Number of training snapshots
M	Mach number
$M_{x,y,z}$	Force balance rolling, pitching, yawing moment
n	Total number of datapoints per snapshot
n_f	Number of pressure taps per cylinder face
n_s	Number of sensors/sparse datapoints
N	Total number of pressure taps on square cylinder
p_actual	Pressure at location of tap
p_measured	Pressure measured by transducer
P_5	5th percentile
q	Dynamic pressure
r_{ii}	Diagonal entry in matrix $\boldsymbol{R}$
R^2	Coefficient of determination
Re	Reynolds number
S	Area of parallelogram
St	Strouhal number
t	Time
V_∞	Freestream flow velocity
x_x	Arbitrary variable
x, y, z	Coordinate system directions
y^+	Dimensionless wall distance
α	Angle of attack/incidence
α_{crit}	Critical angle of attack
β	Offset to global wind tunnel axis system
γ	Angle between sides parallelogram
δ	Indication of sensor placement optimality
δ_{99}	Boundary layer thickness
Δ	Indication of training data skewness
Δ_0	Indication of training data inadequacy
$\Delta F_{x,y,z}$	Force balance axial, side, vertical force accuracy

Symbol	Definition
$\Delta M_{x,y,z}$	Force balance rolling, pitching, yawing moment accuracy
Δt	Timestep size
$\Delta\alpha$	Angle of attack interval
$\Delta\chi$	Cosine spacing angle interval
ϵ	Error
η	Non-dimensional distance from wing root
κ	Condition number
λ	Eigenvalue
μ	Average
ρ	Density
χ	Cosine spacing angle
ψ	Sensor angle
ω	Specific turbulence dissipation
$\boldsymbol{a}$	POD coefficient vector
$\boldsymbol{a}_{1:K}$	Truncated POD coefficient vector
$\boldsymbol{A}$	POD coefficient matrix
$\boldsymbol{A}_{1:K}$	Truncated POD coefficient matrix
$\boldsymbol{b}$	GPOD coefficient vector
$\boldsymbol{C}$	Covariance matrix, sampling matrix
$\boldsymbol{D}$	Column permutation matrix
$\boldsymbol{f}$	Sparse POD modes and $\boldsymbol{w}$ inner product vector
$\boldsymbol{I}$	Identity matrix
$\boldsymbol{M}$	Sparse POD mode inner product matrix
$\boldsymbol{Q}$	Unitary matrix QR factorization
$\boldsymbol{R}$	Upper triangular matrix QR factorization
$\boldsymbol{u}$	Measurement row vector
$\tilde{\boldsymbol{u}}$	GPOD approximation of $\boldsymbol{u}$
$\boldsymbol{U}$	Measurement matrix
$\boldsymbol{v}$	Eigenvector
$\boldsymbol{V}$	Eigenvector matrix
$\boldsymbol{w}$	Sparsely sampled $\boldsymbol{u}$
$\boldsymbol{X}$	Arbitrary matrix
$\boldsymbol{\Theta}$	Sparse POD basis matrix
$\boldsymbol{\Lambda}$	Eigenvalue matrix
$\boldsymbol{\Phi}$	POD basis matrix
$\boldsymbol{\Phi}_{1:K}$	Truncated POD basis matrix

1

Introduction

Identification of aerodynamic parameters for an object of interest is important from a structural, control and overall aerodynamic efficiency point of view. Extensive wind tunnel testing campaigns are dedicated to providing insights into aerodynamic parameters such as the experienced loads under varying conditions. A commonly used method for the determination of both steady and unsteady aerodynamic loads is to measure the surface pressure distribution around an object, which additionally provides valuable insight into the behaviour of the flow (e.g. transition/separation). Measuring surface pressure for aerodynamic load determination is viable by its very principle but in practice limited in accuracy by the ability to approximate a continuous pressure distribution through discrete measurements (He and Williams, 2022). Typically, installing a large number of pressure sensors/taps is required which increases the experimental setup complexity, cost, measurement latency and weight (see Figure 1.1). This hinders the feasibility of the measurement technique for use outside of a wind tunnel and especially for objects with limited physical access and strict size, weight or cost requirements, such as small UAV's.

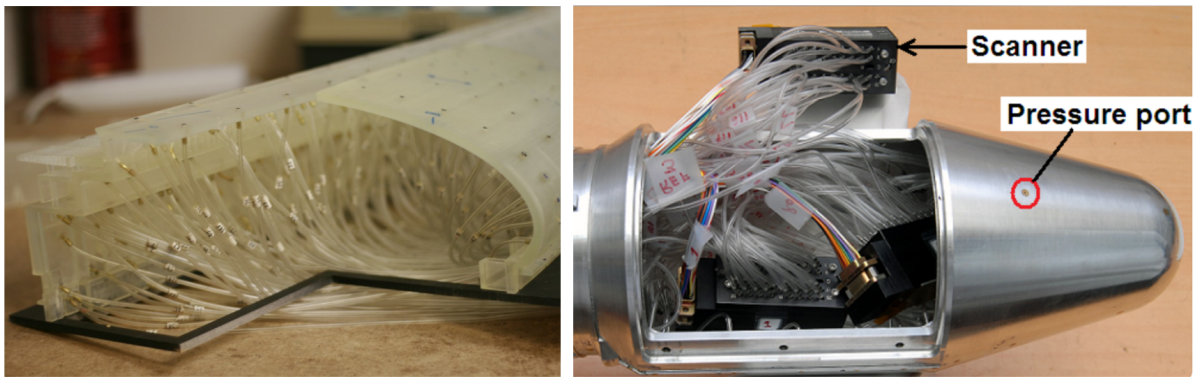


Figure 1.1: Surface pressure measurements by means of pressure taps and tubing connected to transducers/scanners on a building section (left) and cone shaped model (right). Instrumenting the models requires meticulous model preparation and space for the tubing and equipment in or outside of the model. Left figure courtesy of CPP (2022), right figure CSIR-NAL (2022).

Inspiration to overcome the requirement of collecting vast amounts of measurements for determination of aerodynamic parameters often originates from observations in

nature. Birds, fish and insects are able to harness complex fluid phenomena to optimize their thrust, lift or drag accordingly despite their limited computational resources and sparsely distributed noisy sensors (see Figure 1.2). According to Bright et al. (2013) this implies the existence of low-dimensional but informative structures in flow fields which can be sampled sparsely instead to determine the state of a system. Mohren et al. (2018) build upon this idea with neural-inspired sensors for efficient flow classification and Iacobello et al. (2022) use a data-driven approach for the determination of unsteady loads using sparse surface pressure measurements. Data-driven approaches rely on prior training datasets for the identification of underlying low-dimensional structures to be encoded into a reduced order representation or model (ROM).

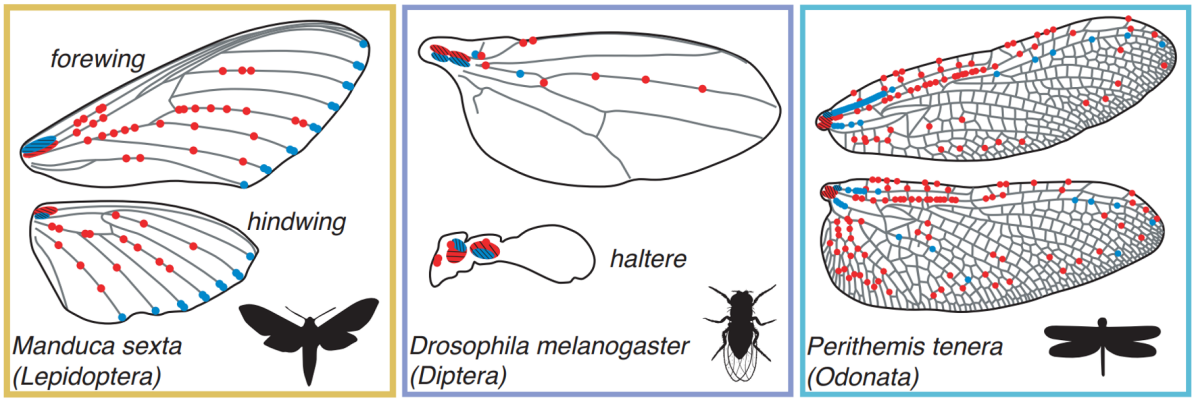


Figure 1.2: Campaniform sensilla (effectively strain sensor) distributions on the wings of different insect species, indicated by dots. Sensor distributions are mathematically sparse but provide mechanosensory feedback for robust flight control through nonarbitrary placement. Adapted from Aiello et al. (2021).

Fortunately, many seemingly complex signals or systems indeed show a data compressible nature in some appropriate coordinate system or basis (Manohar et al., 2018). When information about a system is known in the form of training data, the problem of identifying a suited set of bases for a low-dimensional representation is known as sparse sensing (Kutz et al., 2016). This often involves the use of modal decomposition methods as described by Taira et al. (2017), amongst which is the Proper Orthogonal Decomposition (POD). The POD was introduced to the field of fluid dynamics by Lumley (1967). It is a popular data analysis and reduced order modelling technique because of its ability to decompose an ensemble of data 'snapshots' into modes which can be used for the identification of coherent structures. Also, the orthogonality and energetic ranking of the modes allows for a truncated set of modes to approximate data optimally in the L^2-norm (Chatterjee, 2000).

The ability of POD to provide an efficient 'library' of modes trained for a specific application also finds use in reconstructing full state approximations from sparse measurements instead. This procedure is known as Gappy POD (GPOD) and was introduced by Everson and Sirovich (1995) for the reconstruction of marred or gappy human face images. GPOD has successfully been used for reconstructing complete flowfields and

surface pressure distributions with only sparse surface pressure measurements available as illustrated in Bui-Thanh et al. (2004) and Willcox (2006). In the work by Willcox, unsteady airfoil pressure distributions are reconstructed meaning that the parameter space consists of time. GPOD can however be applied in any other parameter space as long as suitable training is provided. Mifsud et al. (2019) reconstruct complete surface pressure distributions around a three-dimensional wing-body aircraft configuration at multiple angles of attack from sparse surface pressure measurements on the wing alone. POD, in combination with its gappy extension, is therefore demonstrated to be a viable method for the determination of full state pressure distributions and aerodynamic forces based on sparse surface pressure measurements instead. This has the potential to reduce the required number of physical sensors at the cost of a one-time high resolution training dataset to be generated prior, either using CFD or experiments. Many research papers making use of GPOD in this context however tend to test the potential of GPOD on aerodynamic bodies within parameter spaces characterized by predictable fluid behaviour, such as linear aerodynamics. Also, with only sparse sensors being used to characterize a full state pressure distribution, optimal positioning of the sensors is critical. This was realized already by Cohen et al. (2003) for the reconstruction of flowfields around a cylinder but is often not considered in later experiments, such as by Mifsud et al., despite the possibility of providing more robust reconstructions with fewer sensors (Manohar et al., 2018). Therefore, sensor placement is considered in combination with GPOD on a bluff-body with the following objective:

”To assess the viability of using Gappy Proper Orthogonal Decomposition as a method to obtain surface pressure distributions and drag estimates based on experimentally obtained sparse surface pressure measurements on a 2D square cylinder”

The steps undertaken to reach the objective are described in this thesis report. The structure of the report is as follows: In Chapter 2 related research topics will be briefly discussed and a theoretical background will be provided. The core of the thesis focuses on the combination of POD, GPOD and a sensor placement algorithm for which mathematical descriptions are provided and a basic sparse reconstruction example as illustration. This chapter concludes with a description of typical square cylinder flow characteristics and flow topology. Chapter 3 contains a description of the experimental setup and procedures. The setup includes the square cylinder model, the measurement equipment and external support structures. Data acquisition procedures include a summary of the test matrix, parameters used for the force balance and surface pressure measurement equipment and a short description of calibration and correction procedures. A numerical counterpart to the experimental measurement campaign is setup as well and its results are discussed in Chapter 4. The underlying goal of this chapter is to eventually provide a source of training data without the requirement of a wind tunnel. All results based on the experimental measurements from the wind tunnel campaign are presented in Chapter 5. GPOD is applied on sparse experimental surface pressure measurements with purely experimental training data and in two steps CFD results are introduced into the procedure to decrease the reliance on wind tunnel measurements. Finally in Chapter 6, the conclusions which can be drawn from the

results obtained in Chapter 5 are stated and recommendations to improve or extend upon the work are provided.

2

Literature and theoretical background

In the following chapter, several topics related to the mathematical and experimental background of predicting drag and surface pressure using sparse measurements will be discussed. Some attempts described in literature are described first which highlights the connection with reduced order modeling and sparse sampling of which some applications and useful mathematical frameworks will therefore be discussed in more detail as well. A practical yet simple object for a sparse drag determination procedure to be demonstrated on experimentally is found in a square cylinder. Square cylinder flow is described extensively throughout literature of which a summarized version is provided at the end of this chapter

2.1. Sparse surface pressure and drag correlation

Relating the variation of aerodynamic loads to fluctuations in the measured surface pressure seems like a logical approach, especially for pressure drag dominated objects known as bluff bodies. By its very principle, integration of the spatial pressure distribution on a bluff body yields a close approximation of the net forces acting on an object (He and Williams, 2022). The accuracy of this approach in practice is often limited by the finite number of pressure sensors available; however even using sparse surface pressure measurements a 'signature' of the pressure distribution remains. An attempt to exploit this signature with the goal of identifying its correlation with spatio-temporal aerodynamic loads is described in the work by Burelle et al. (2020). In their work, a non-slender delta wing equipped with sparsely distributed pressure sensors is used for simultaneous measurements of the surface pressure and aerodynamic forces under the effect of gusts. These measurements allow for the determination of pressure port specific coefficients to be used in a linear combination of pressure measurements which approximate the desired aerodynamic loads. A graphical summary of the method is provided by Burelle et al. which is shown below in Figure 2.1.

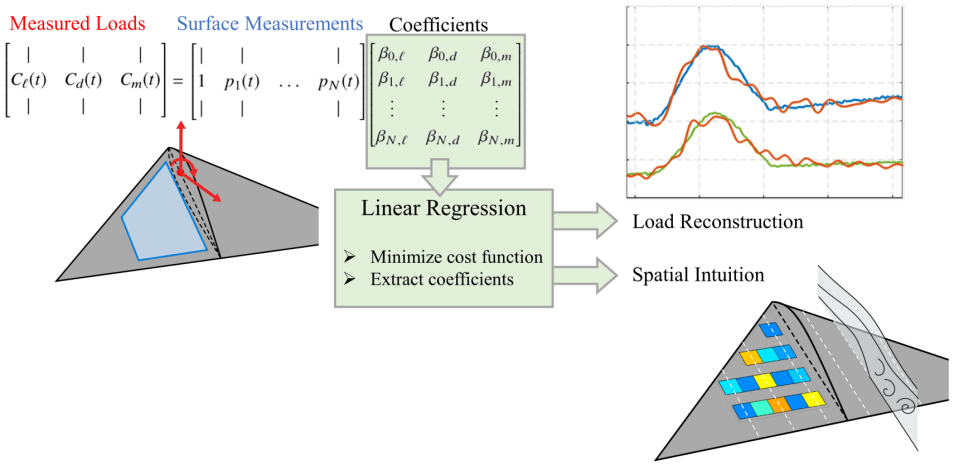


Figure 2.1: Linear regression approach towards the reconstruction of loads from sparse surface pressure measurements Burelle et al. (2020)

In the model shown above, the temporal variation of the aerodynamic loads is accounted for in the pressure measurements. The coefficients therefore represent the spatial aspect of the chosen sensor distribution and can be used to identify critical regions on the body through their relative magnitude. This relatively simple model was used to provide sets of coefficients suited for drag determination at $\alpha = 10°, 20°, 30°$ and across these angles of attack. At a low angle of attack ($10°$ in this case), the delta wing drag is expected to not be dominated by normal pressure components which is confirmed by overall poor drag predictions which highlights an obvious limitation of pressure based drag determination. At $\alpha = 20°$ and $30°$ especially, drag determination accuracy increases with Root Mean Squared Errors (RMSE) in the order of 2 % of the actual drag coefficients and R^2 values above 0.8 for the angle specific models. Aggregate models usable across the angles of attack instead showed reduced accuracy as would be expected however remained remarkably capable of determining qualitative drag while serving as a more robust model.

With their work, Burelle et al. fundamentally test the possibility of measuring pressure signatures tied to the experienced loads using sparse sensors. The linear regression approach is a relatively simple model and as no particular reasoning behind the used sensor locations is provided, the model might not be optimal for the desired purpose meaning similar drag prediction accuracy could be possible using fewer sensors. Another paper by Liu et al. (2021), describing their findings when correlating surface pressure and drag of an Ahmed body including active flow control, incorporates the expected importance of the pressure sensor locations for a more efficient model. Despite the more specific application oriented research of Liu et al. compared to Burelle et al., both use a linear regression approach to obtain drag in terms of a linear combination of constant coefficients and varying pressure measurements. Liu et al. sequentially add or remove a predictor variable (pressure sensor) based on its statistical significance in predicting variations in the drag. Despite 19 pressure sensors being available on the backside of the Ahmed body, this approach showed that only two of those are sufficient to accurately predict variations in the drag. This shows how a more efficient and less complex model can be obtained by considering sensor

locations, something which will be discussed in more detail later.

2.2. Modal decomposition

The regression model approaches used in the above mentioned research are data-driven in the sense that useful patterns are recognized in the drag variation from provided data only without additional input or user knowledge. As mentioned by Burelle et al. however, generic sensor locations in combination with the model might lead to undesired correlation between measurements. Especially with the aim of increasing sparsity in the drag determination, this might yield a sub-optimal model not focused on measuring the most relevant features of the pressure distribution. Instead, a model based on a set of orthogonal bases might be better suited to capture dominant features of a system and provide more informative sensor locations, possibly reducing the number of required sensors. The benefits of using such orthogonal bases derived from the data become clear and physically interpretable in a 2D case example which will be provided following the explanation by Weiss (2019).

Say a variable is measured through time at two locations (e.g. velocities within a wake or pressures on an airfoil surface similar to Burelle et al.). The measurements at location 1 can be denoted x_1 and at location 2 x_2. Assuming the temporal mean is removed, the data can be visualized in a 2D plot as shown in Figure 2.2 and stored in $\boldsymbol{X} \in \mathbb{R}^{m \times 2}$, m being the number of times x_1 and x_2 are measured. Generally, the shape of an ellipse emerges meaning that the measurements of x_1 and x_2 are correlated. This correlation can be presented in the form of covariance matrix $\boldsymbol{C} \in \mathbb{R}^{2 \times 2}$:

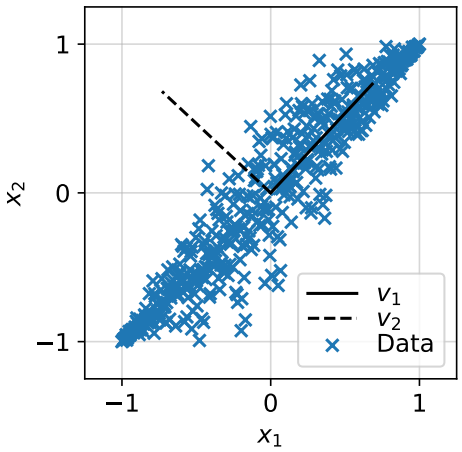


Figure 2.2: Data points and direction of eigenvectors

$$\boldsymbol{C} = \frac{1}{m-1}\boldsymbol{X}^T\boldsymbol{X} \tag{2.1}$$

$\boldsymbol{C}$ is symmetric with diagonal elements being the variance of x_1 and x_2 and the off-diagonal elements being the covariance between x_1 and x_2. The variances in $\boldsymbol{C}$ are measured along the 'natural axes' being the shown horizontal and vertical axes for the variance of x_1 and x_2 respectively, these natural axes thus serve as modes of variation. This choice of modes is arbitrary, however maximum variance is obtained along the major axis of the ellipse and all remaining variance can be accounted for along the minor axis of the ellipse. The eigenvector corresponding to the largest eigenvalue of $\boldsymbol{C}$ can be shown to point in the direction of maximum variance, therefore in the direction of the major axis. Since $\boldsymbol{C}$ is symmetric, its eigenvectors are orthogonal, as are the principle axes of the ellipse. The second eigenvector of $\boldsymbol{C}$ thus points in the direction of the minor axis. The eigenvectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ of $\boldsymbol{C}$ are included in Figure 2.2 and clearly point in the expected directions. $\boldsymbol{C}$ can be diagonalized as follows:

$$\boldsymbol{C} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{T} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} \tag{2.2}$$

with corresponding eigenvalues λ. The data in $\boldsymbol{X}$ can be projected onto the eigenvectors as:

$$\boldsymbol{X'} = \boldsymbol{X}\boldsymbol{V} \tag{2.3}$$

It can be shown that the covariance matrix of $\boldsymbol{X'}$ will be $\boldsymbol{\Lambda}$. This shows how for the projected data, correlation between x_1 and x_2 is not given by their covariance as $\boldsymbol{\Lambda}$ is diagonal but instead is implicit in the directions of the eigenvectors. Also, the eigenvalues can be seen to measure the variance along these eigenvector directions. Note that Equation 2.2 corresponds to an eigendecomposition which provides the desired set of orthogonal bases and the corresponding eigenvalues, often ranked in order of magnitude.

The orthogonality and ranking of the modes allows for efficient data decomposition for identification of dominant features. It is therefore not surprising that the eigendecomposition is used extensively as the underlying framework for data analysis throughout various modal decomposition methods such as the Singular Value Decomposition (SVD) and related Proper Orthogonal Decomposition (POD), all discussed in Taira et al. (2017) for fluid dynamics related use. The POD was first introduced into the field of fluid dynamics by Lumley (1967) and is very similar to the eigendecomposition of the covariance matrix as discussed. The original idea was to decompose fluid flows into orthogonal modes and coefficients, each mode linked to a coherent structure which can serve as a useful tool in the analysis of data. Berkooz et al. (1993) describes the use of POD and its properties for turbulent flows in general. Mohammed-Taifour and Weiss (2016) performed POD analysis on a more specific case; the identification of dominant characteristics of unsteady turbulent separation bubbles (TSB), for which PIV measurements were used. A schematic of their experimental setup together with the first found POD mode of the horizontal velocity component is shown in Figure 2.3. The high modal activity region in the first mode identifies the extent of the turbulent separation bubble. POD modes visualize correlation within the flow, mode 1 therefore illustrates the global fluctuation of the complete TSB.

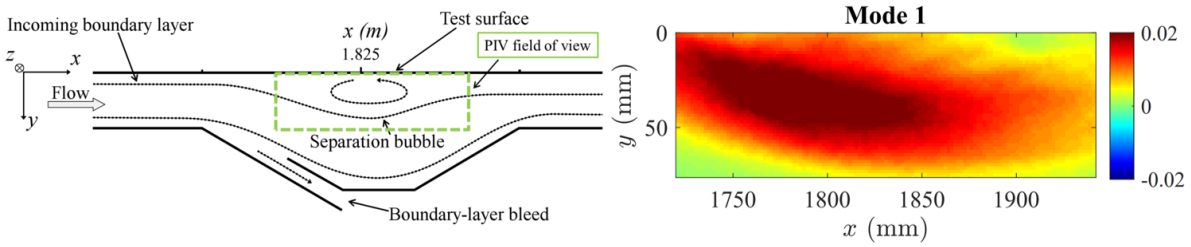


Figure 2.3: Experimental setup used by Mohammed-Taifour and Weiss (2016) (left) and first mode of the horizontal velocity component (right) identifying correlation in the flow. Figures adapted from Weiss (2019)

2.3. Snapshot POD

The snapshot POD is a variation on the direct POD introduced by Sirovich (1987) which is especially suited for datasets that are much larger in the number of datapoints n compared to the number of instants these datapoints are measured m. Such datasets for which $n \gg m$ are commonly encountered in the field of fluid dynamics in the form of PIV measurements or CFD simulation results for example and often capture the state of a system at distinct time instants, as if snapshots were taken of the flow. The theory required to understand and implement the snapshot POD method will be explained below following the information provided by Weiss (2019) and Brunton and Kutz (2019). Similar to the tutorial by Weiss, the theory will be explained by assuming a dataset in the form of PIV measurements throughout time; however could also be applied to datasets of different origin, even when not sampled throughout time (Bui-Thanh et al., 2004) which will actually be the focus in the remainder of this report.

Let's assume PIV measurements are taken on a two-dimensional domain with a resolution of $n_x \times n_y$. Each PIV 'snapshot' therefore contains $n_x \times n_y = n$ measurements of the velocity which can be ordered in a row vector $\boldsymbol{u} \in \mathbb{R}^{1\times n}$. In total m of such snapshots are taken throughout time which can be stacked vertically to form matrix $\boldsymbol{U} \in \mathbb{R}^{m\times n}$. The temporal mean over all snapshots is assumed to be subtracted from each individual snapshot in the construction of $\boldsymbol{U}$, although this is not strictly required. The goal of the snapshot POD is to represent every snapshot $\boldsymbol{u}$ stored in $\boldsymbol{U}$ as a linear combination of temporal coefficients $a(t)$ and spatial modes $\boldsymbol{\Phi}_k(\boldsymbol{x})$ such that:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \sum_{k=1}^{m} a_k(t)(\boldsymbol{\Phi}_k(\boldsymbol{x}))^T \tag{2.4}$$

When such a decomposition is found for all m snapshots of $\boldsymbol{u}$, the coefficients $a(t)$ can be stored in $\boldsymbol{A} \in \mathbb{R}^{m\times m}$ and the corresponding modes in $\boldsymbol{\Phi} \in \mathbb{R}^{n\times m}$ such that Equation 2.4 can be represented in matrix form:

$$\boldsymbol{U} = \boldsymbol{A}\boldsymbol{\Phi}^T \tag{2.5}$$

With:

$$\boldsymbol{A} = \begin{bmatrix} a_1(t_1) & \cdots & a_m(t_1) \\ \vdots & \ddots & \vdots \\ a_1(t_m) & \cdots & a_m(t_m) \end{bmatrix}, \quad \boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_1(\boldsymbol{x}) & \cdots & \boldsymbol{\Phi}_m(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_m(x_1) \\ \vdots & \ddots & \vdots \\ \phi_1(x_n) & \cdots & \phi_m(x_n) \end{bmatrix} \tag{2.6}$$

Obtaining the above shown matrices for this decomposition starts by computing the covariance matrix $\boldsymbol{C}$ as:

$$\boldsymbol{C} = \frac{1}{m-1}\boldsymbol{U}\boldsymbol{U}^T \tag{2.7}$$

Since $\boldsymbol{U}$ has dimension $m \times n$, matrix $\boldsymbol{C} \in \mathbb{R}^{m\times m}$. Note that the matrix multiplication of $\boldsymbol{U}$ with its transpose results in $\boldsymbol{C}$ being symmetric. Eigenvalues λ_k of this symmetric matrix $\boldsymbol{C}$ are real, distinct and can be ordered in descending order. Since the original

data in matrix $\boldsymbol{U}$ was of unit m s^{-1} it follows that the eigenvalues of matrix $\boldsymbol{C}$ have units m^2 s^{-2}. Ordering the found eigenvalues in descending order therefore maximizes the kinetic energy content captured by the first K ($1 \leq K \leq m$) eigenvalues. The P in POD thus refers to proper or optimal capture of the dataset's energy content within the first K eigenvalues and modes.

When ordered following their corresponding eigenvalues, the eigenvectors $\boldsymbol{v}_k$ of $\boldsymbol{C}$ together are denoted by the aforementioned matrix $\boldsymbol{A}$. Since all eigenvalues are distinct, the eigenvectors or columns in $\boldsymbol{A}$ are orthogonal and even orthonormal. The spatial modes in $\boldsymbol{\Phi}$, also referred to as the POD modes or POD bases can be obtained by projecting the data in $\boldsymbol{U}$ onto the eigenvectors in $\boldsymbol{A}$ as shown below:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_m \end{bmatrix}, \quad \boldsymbol{\Phi} = \boldsymbol{U}^T\boldsymbol{A} \tag{2.8}$$

At this point, the POD modes in $\boldsymbol{\Phi}$ are orthogonal and the desired decomposition of the dataset in the form of Equation 2.5 has been realized:

$$\boldsymbol{\Phi} = \boldsymbol{U}^T\boldsymbol{A} \rightarrow \boldsymbol{U}^T = \boldsymbol{\Phi}\boldsymbol{A}^{-1} = \boldsymbol{\Phi}\boldsymbol{A}^T \rightarrow \boldsymbol{U} = \boldsymbol{A}\boldsymbol{\Phi}^T \quad \text{using} \quad \boldsymbol{A}^{-1} = \boldsymbol{A}^T \tag{2.9}$$

Different from the direct POD method however, $\boldsymbol{A}$ is orthonormal while $\boldsymbol{\Phi}$ is only orthogonal in its columns. For both the temporal coefficients and orthogonal modes obtained using the snapshot POD to match the direct POD (signs might be reversed), the columns of $\boldsymbol{\Phi}$ should be orthonormalized instead and $\boldsymbol{A}$ scaled accordingly:

$$\boldsymbol{\Phi} = \begin{bmatrix} \frac{\boldsymbol{\Phi}_1}{||\boldsymbol{\Phi}_1||} & \cdots & \frac{\boldsymbol{\Phi}_m}{||\boldsymbol{\Phi}_m||} \end{bmatrix} \rightarrow \boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Phi} \tag{2.10}$$

Throughout literature, the terms orthogonality and orthonormality are often used indistinguishably however the orthonormality (and hence also orthogonality) of the spatial bases/columns in $\boldsymbol{\Phi}$ is what is referred to by the O in POD. Through the scaling of $\boldsymbol{A}$, it loses its orthonormality however the columns keep orthogonality since $\boldsymbol{A}^T\boldsymbol{A} = (m-1)\boldsymbol{\Lambda}$ with:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix} \tag{2.11}$$

On top of being able to recognize coherent structures in the spatial distribution of the POD modes, the power of the proper orthogonal decomposition lies in the ability to construct energy optimal lower rank approximations using the obtained coefficients and bases. Instead of using all m coefficients and modes in reproducing $\boldsymbol{u}$ at a specific time instant as shown in Equation 2.4, selecting instead $K < m$ to capture the desired amount of energy can be used to approximate the snapshot $\boldsymbol{u}$ instead. Such a lower rank approximation will be referred to as a POD projection at rank K and can be denoted as follows for a single snapshot- and full dataset lower rank projection respectively:

$$\boldsymbol{u}(\boldsymbol{x},t) \approx \sum_{k=1}^{K} a_k(t)(\boldsymbol{\Phi}_k(\boldsymbol{x}))^T, \quad \boldsymbol{U} \approx \boldsymbol{A}_{1:K}(\boldsymbol{\Phi}_{1:K})^T \quad \text{with} \quad 1 \leq K \leq m \tag{2.12}$$

The notation $\boldsymbol{X}_{1:K}$ means that only a subset of the columns, being column 1 up to and including K, are used in the matrix multiplications. The steps in Equation 2.10 are not strictly necessary as projections with or without these actions performed yield identical results. Note that the optimality of the POD bases translates into a POD projection providing the best approximation of $\boldsymbol{u}$ in a least squares sense for a given value of K.

A fluid mechanics application making use of the energy optimality of POD projections is described by Wang et al. (2015). In their work, Wang et al. propose a POD based outlier correction (POD-OC) but also detection method for PIV images. Their method was implemented and evaluated by Higham et al. (2016) through application on channel flow and isotropic flow cases from the John Hopkins Turbulence Database (JHTDB) by including artificial outliers. Some of the results are shown in Figure 2.4.

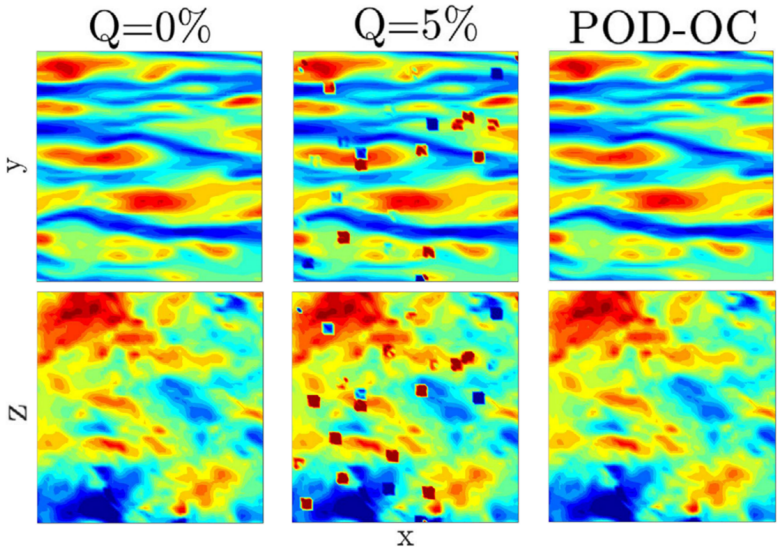


Figure 2.4: POD based outlier detection for PIV (Wang et al., 2015) applied to channel flow (top) and isotropic flow (bottom). From left to right: original longitudinal velocity field, velocity field including 5 % synthetic outliers and POD outlier corrected velocity field. Adapted from Higham et al. (2016)

The method relies on the assumption that outliers end up in the higher order modes of a Proper Orthogonal Decomposition, something which appears reasonable to expect as a result of the incoherent nature of outliers. The detection of outliers is then performed by comparing flow snapshots to low-order projections and computing statistics of the differences observed. Detected outliers can subsequently be replaced by such lower rank projections to retrieve an approximation of the original flowfield. Using POD thus enables information of the complete datasets to be used in the correction of outliers unlike methods such as linear, bi-linear or spline interpolation that rely on local information only.

2.4. POD for sparse sensing using tailored bases

The energy optimality of the POD modes allows POD projections as in Equation 2.12 to efficiently represent complex systems in reduced dimensions. This inherently assumes the existence of such a lower dimensional representation and raises the question if such 'compressible signals' are actually common. For the research presented by Burelle et al. and Liu et al., compressibility in the surface pressure distribution, referred to as the 'pressure signature', seems to have been found for their specific applications in the context of drag determination. Observations in nature of animals such as birds and insects harnessing complex unsteady fluid phenomena despite their limited sensory capabilities and 'computational resources' led to the conjecture of the existence of low-dimensional structures capable of being sparsely measured by Bright et al. (2013). Manohar et al. (2018) illustrates the highly compressible nature of many audio signals and images using an image compression example as shown below in Figure 2.5

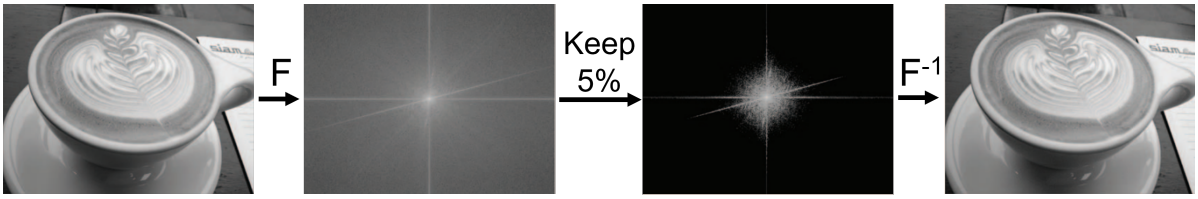


Figure 2.5: Fourier image compression, adapted from Manohar et al. (2018)

The image compression example is performed using the 2D Fourier transform. As can be seen, neglecting $95\,\%$ of the smallest Fourier coefficients before inverse Fourier transforming does not significantly affect image quality while reducing the required memory. The Fourier modes are a universal basis in the sense that many natural images have a sparse representation in them (this is the working principle for JPEG compression). This is useful for the compression or reconstruction of datasets with unknown contents. For more specific applications however where the dataset's contents is known (e.g. pressure distributions or PIV images of an airfoil) the POD provides a 'tailored' set of basis instead, capturing the dominant features of the specific use case (Manohar et al., 2018).

POD modes and coefficients can thus effectively be 'trained' to contain information for a specific use case. This is often used in combination with Garlerkin projection (Quarteroni and Rozza, 2014) for example where high-dimensional systems are projected onto a low-dimensional subspace identified by POD. This high dimensional system could be in the form of non-linear partial differential equations such as the Navier-Stokes equations used for POD-Galerkin projection in Lorenzi et al. (2016). In this paper, both laminar and turbulent cavity flow are simulated using an expensive Full Order Model (FOM, pimpleFoam) and using a POD-Galerkin projection obtained Reduced Order Model (ROM). Lorenzi et al. compare instantaneous results of the ROM to the FOM but also compare the steady state solution of the laminar ROM case to the benchmark cavity flow results by Botella and Peyret (1998). The steady state vorticity distributions for this comparison are shown in Figure 2.6 where it can clearly be seen that the POD based ROM is in almost perfect agreement with results from literature.

Note that the online procedure of the ROM simulation (actual computation of the solution) took only 5 s to simulate 100 s of cavity flow while the FOM would require 2590 cpu-hours for the same.

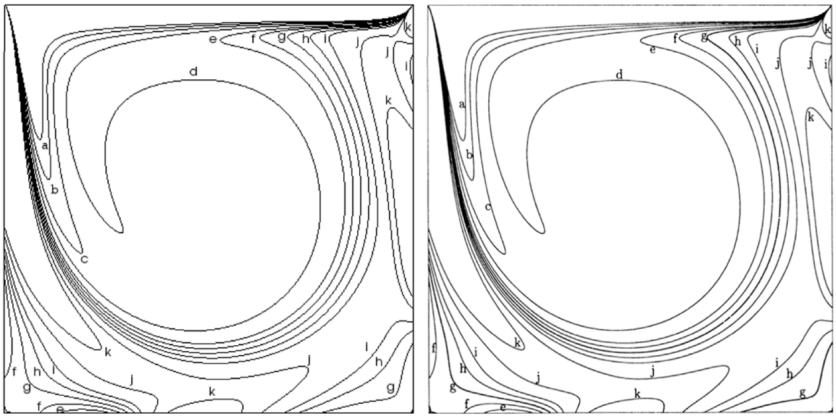


Figure 2.6: Steady state solution of laminar lid driven cavity flow obtained by Lorenzi et al. (2016) using POD-Garlerkin ROM (left) and benchmark solution from Botella and Peyret (1998) (right)

Both the Fourier image example and POD-Galerkin projection use the bases to represent full-state information sparsely. This is however exactly the opposite of what is aimed for when approximating for example the drag coefficient or pressure distribution using only a small number of pressure sensors; instead of compressing data into a sparse representation using generic or tailored bases, the goal is to sparsely sample a signal and infer an approximation of the full-state. When it is known what type of full-state is being approximated, this inversion of the compression paradigm (Manohar et al., 2018) can leverage information obtained from training data stored in a set of tailored bases, something the POD is able to provide given said training data is appropriate.

This idea of using tailored POD bases, obtained through training, to reconstruct a data set from only sparse measurements of the data set has been successfully tested first by Everson and Sirovich (1995). In their paper named the ”Karhunen-Loève (KL) procedure for gappy data” (KL being another name under which POD is known), a set of images containing human faces was used as training data. The orthogonal modes produced by the POD procedure were called eigenfaces due to their link with eigenvectors/values and clear resemblance to human faces as can be seen in Figure 2.7. The method used by Everson and Sirovich was later referred to as ’gappy POD’ and is described in more detail in the following Section

Figure 2.7: Example eigenfaces obtained from application of POD procedure to Yale database, figure adapted from Brunton and Kutz (2019).

2.5. Gappy POD

Instead of obtaining reduced order representations of data using proper orthogonal decomposition, Everson and Sirovich (1995) inverted the usual POD application and described a procedure of using a tailored set of POD bases (eigenfaces) to restore marred or 'gappy' snapshots to their full state. The procedure relies on using a set of bases being representative for a specific class from which the gappy snapshot to be reconstructed originates. The properties of the POD make its bases a good choice to represent such a 'library' of information and can be obtained following the procedure described in Section 2.3. For applications in the field of fluid dynamics, the tailored POD bases could be obtained from e.g. PIV snapshots $\boldsymbol{u}$ of a particular flow case and stored in $\boldsymbol{\Phi}$ (using the same nomenclature as used in Section 2.3). The goal of the gappy POD would then be to approximate a complete snapshot $\boldsymbol{u}$ of size n, as a linear combination of K coefficients b and modes stored in $\boldsymbol{\Phi}$ while only having access to $n_s < n$ sparse measurements from $\boldsymbol{u}$:

$$\boldsymbol{u} \approx \tilde{\boldsymbol{u}} = \sum_{k=1}^{K} b_k \boldsymbol{\Phi}_k^T = \boldsymbol{b}(\boldsymbol{\Phi}_{1:K})^T \tag{2.13}$$

With $\boldsymbol{b} = [b_1\ b_2 \cdots b_K]$. The sparse samples from $\boldsymbol{u}$ are denoted by $\boldsymbol{w}$ with dimension n_s. If $\tilde{\boldsymbol{u}}_{[\boldsymbol{n_s}]}$ denotes those entries of $\tilde{\boldsymbol{u}}$ supported by the sparse n_s measurements also present in $\boldsymbol{w}$, then it would make sense to minimize the difference between $\boldsymbol{w}$ and $\tilde{\boldsymbol{u}}_{[\boldsymbol{n_s}]}$ in a least squares sense, thus minimizing the error defined as:

$$\boldsymbol{\epsilon} = ||\boldsymbol{w} - \tilde{\boldsymbol{u}}_{[\boldsymbol{n_s}]}||^2 \tag{2.14}$$

Since the support over the sparse measurements can only be accomplished through the spatial POD bases in $\boldsymbol{\Phi}_{1:K} \rightarrow \tilde{\boldsymbol{u}}_{[\boldsymbol{n_s}]} = \boldsymbol{b}(\boldsymbol{\Phi}_{1:K})^T_{[n_s]} = \boldsymbol{b}\boldsymbol{\Theta}^T$. Note that thus far vectors $\boldsymbol{u}$ and $\tilde{\boldsymbol{u}}_{[\boldsymbol{n_s}]}$ have been assumed to be row vectors. If instead represented as column vectors by means of taking the transpose, the error can be rewritten as:

$$\boldsymbol{\epsilon} = ||\boldsymbol{w}^T - \tilde{\boldsymbol{u}}^{\boldsymbol{T}}_{[\boldsymbol{n_s}]}||^2 = ||\boldsymbol{w}^T - \boldsymbol{\Theta}\boldsymbol{b}^T||^2 \tag{2.15}$$

The solution for $\boldsymbol{b}^T$ which minimizes this error is given by solving a linear system of the form:

$$\boldsymbol{M}\boldsymbol{b}^T = \boldsymbol{f} \rightarrow \boldsymbol{\Theta}^T\boldsymbol{\Theta}\boldsymbol{b}^T = \boldsymbol{\Theta}^T\boldsymbol{w}^T \tag{2.16}$$

Since $\boldsymbol{\Theta} \in \mathbb{R}^{n_s \times K}$ matrix $\boldsymbol{M} \in \mathbb{R}^{K \times K}$. By comparing Equation 2.13 with Equation 2.4, it is clear that the approximation sign present in Equation 2.13 is there because of errors introduced in the reconstruction which originate from two aspects. First, similar to obtaining a lower rank POD projection as described in Equation 2.12, K is generally taken to be smaller than m as to use only the most energetic POD modes present. In principle one could use $K = m$, however for the GPOD procedure described above to function, $n_s \geq K$ which means that increasing the number of POD bases involved in the reconstruction increases the required samples from $\boldsymbol{u}$, decreasing sparsity. Note that $n_s = K$ yields $\boldsymbol{\epsilon} = 0$ resulting in the provided sparse measurements being interpolated exactly. Secondly, the coefficients b_k in the sparse case are only approximations of the coefficients a_k for the non-sparse snapshot POD described in Section 2.3. This can be recognized by applying the GPOD procedure in an extreme case where $n_s = n$. For this case, $\boldsymbol{w} = \boldsymbol{u}$ and $\boldsymbol{\Theta}$ reverts back to being equal to $\boldsymbol{\Phi}_{1:K}$ which has K orthogonal columns. Equation 2.16 can thus be written as:

$$(\boldsymbol{\Phi}_{1:K})^T \boldsymbol{\Phi}_{1:K} \boldsymbol{b}^T = (\boldsymbol{\Phi}_{1:K})^T \boldsymbol{u}^T \rightarrow \boldsymbol{b}^T = (\boldsymbol{\Phi}_{1:K})^T \boldsymbol{u}^T \rightarrow \boldsymbol{b} = \boldsymbol{u} \boldsymbol{\Phi}_{1:K} \text{ with } (\boldsymbol{\Phi}_{1:K})^T \boldsymbol{\Phi}_{1:K} = \boldsymbol{I} \tag{2.17}$$

Comparing this result to Equation 2.10 shows how for $n_s = n$ the coefficients obtained through GPOD are equal to those from the snapshot POD method, $\boldsymbol{b} = \boldsymbol{a}_{1:K}$ or in matrix format $\boldsymbol{B} = \boldsymbol{A}_{1:K}$ with $\boldsymbol{B}$ obtained through stacking all computed row vectors $\boldsymbol{b}$ corresponding to different snapshot reconstructions. For any $n_s < n$, the orthogonality of the used POD bases in the GPOD reconstructions is generally destroyed, hence $\boldsymbol{\Theta}^T \boldsymbol{\Theta} = \boldsymbol{M} \neq \boldsymbol{I}$ which affects the solution for $\boldsymbol{b}^T$ resulting in $\boldsymbol{b} \neq \boldsymbol{a}_{1:K}$ and an additional 'error due to gappiness'.

Everson and Sirovich even extended the gappy POD algorithm for the case where the training snapshots themselves are gappy. Instead of a single application of the snapshot POD algorithm, it is applied iteratively to converge towards a final set of POD bases required for the gappy reconstruction procedure. For POD to work with an incomplete snapshot set however, the locations of missing data in each snapshot should be spread throughout the snapshot ensemble as to never contain a complete 'black spot'. This renders the method impractical for most sparse sensing applications with fixed sensor positions however has been used successfully in the reconstruction of experimental measurements containing gaps in the data as demonstrated by Murray and Ukeiley (2007) and Raben et al. (2012) on PIV measurements for example. A turbulent channel flow PIV image reconstruction from Raben et al. is shown in Figure 2.8, who state to have improved upon the 'standard' iterative gappy POD by Everson and Sirovich with their introduction of 'adaptive gappy POD'.

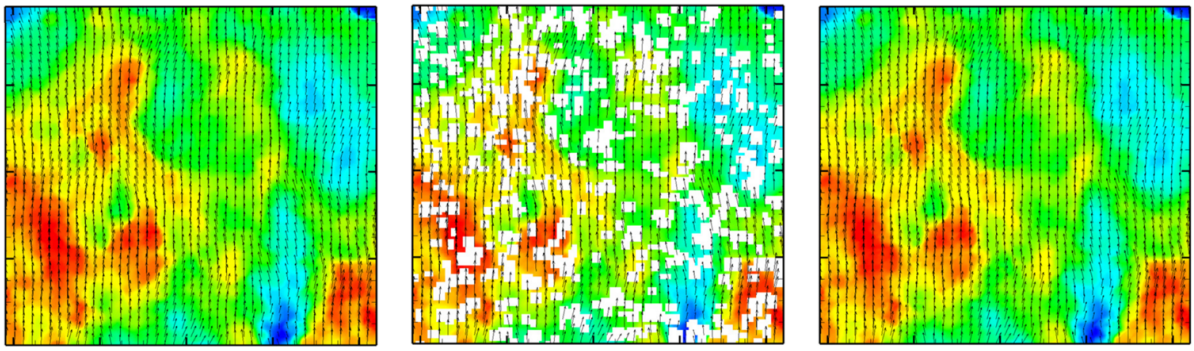


Figure 2.8: Adaptive gappy POD for gappy PIV image reconstruction, applied to turbulent channel flow velocity field. Raben et al. (2012)

Similar in essence to the standard iterative gappy POD, adaptive gappy POD iteratively reconstructs missing velocity vectors using a lower rank POD projection originating from the full snapshot ensemble. The optimal rank of this projection might vary for each spatial location however and is therefore found for each reconstructed vector individually, again iteratively. This yields visually indistinguishable reconstructions from the original in this particular case as can be seen however introduces significant additional computational cost.

2.6. GPOD sensor placement

Two important aspects to take into consideration during the gappy POD procedure are the locations at which the n_s sparse measurements are taken (sensor locations) and what snapshots to provide as training. The sensor locations determine what is sampled in $\boldsymbol{w}$ and the POD modes and therefore what is incorporated into $\boldsymbol{\Theta}$ while the provided training snapshots directly affect the POD modes themselves hence $\boldsymbol{\Theta}$ as well. From Equation 2.16 it is clear that this will affect the solution for $\boldsymbol{b}$ which is used for the reconstruction.

Unlike what has been used in the examples before, gappy POD can in principle have its training data vary in any parameter (or even multiple) yielding POD coefficients varying with the chosen parameter(s). When training data is provided in the form of pressure distributions over an airfoil at different angles of attack for example, matrix $\boldsymbol{A}$ from Section 2.3 will contain coefficients depending on the angle of attack instead. Gappy POD effectively provides a data-driven interpolation tool and just like regular interpolation, the interpolation points should sample the underlying function appropriately. Sticking with the airfoil training data example; the angles of attack chosen to provide a training snapshot on should thus capture the flow physics to be encountered during sparse measurements for accurate pressure distribution reconstructions. For gappy POD use in the field of fluid dynamics, training snapshots could be obtained by means of non-sparsely measured windtunnel campaigns. In case the accuracy of CFD simulations for the specific use case is expected to be sufficient however, numerical simulations could be used instead, possibly eliminating the need for highly resolved windtunnel tests to be performed.

Mifsud et al. (2019) illustrate such a use case in their attempt to enhance the determination of aerodynamic coefficients w.r.t. CFD by using experimentally measured surface pressures in combination with gappy POD and CFD training snapshots. In their work, Mifsud et al. apply gappy POD on a 3D wing-body aircraft configuration. At $Re = 3.3 \cdot 10^6$ and $M = 0.85$, seven RANS simulations from $\alpha = 0°$ to $12°$ were performed to act as training snapshots. Experimental balance and surface pressure measurements at $\alpha = 7.1°$ and $7.9°$ were performed as well using 381 pressure sensors. Using RANS as training and the experimental surface pressure measurements as gappy snapshots, the GPOD procedure was able to reconstruct surface pressure distributions over the model in its entirety. Figure 2.9 shows pressure distribution results at certain section cuts of the model's wing (η being a non-dimensional distance from the wing root).

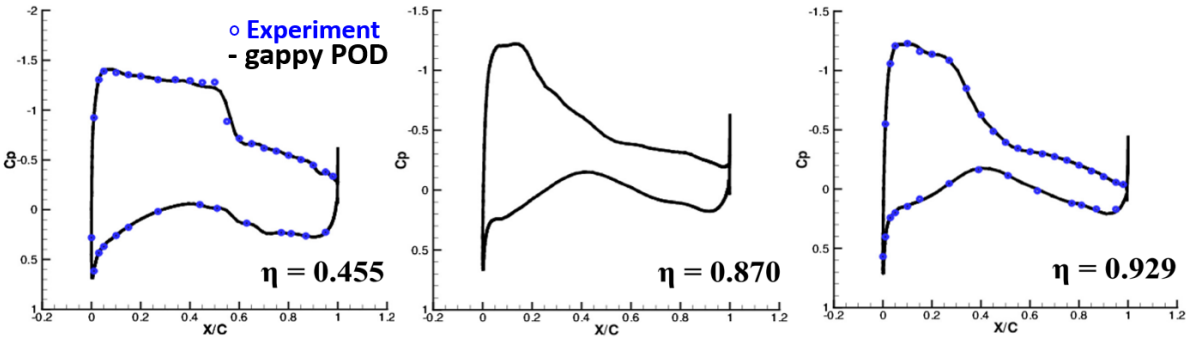


Figure 2.9: C_p distribution at different section locations η along the wing span of a 3D wing-body aircraft configuration at $\alpha = 7.1°$ and $M = 0.85$, adapted from Mifsud et al. (2019)

Even at sections along the wingspan where no experimental surface pressure measurements were taken (middle Figure 2.9), gappy POD is able to predict a practically continuous surface pressure distribution conform the RANS training while informed by the experimental measurements in the proximity. The experimental measurements performed at the sections indicated in the left of Figure 2.10 alone are therefore sufficient for the entire surface pressure reconstruction as shown in the right of Figure 2.10.

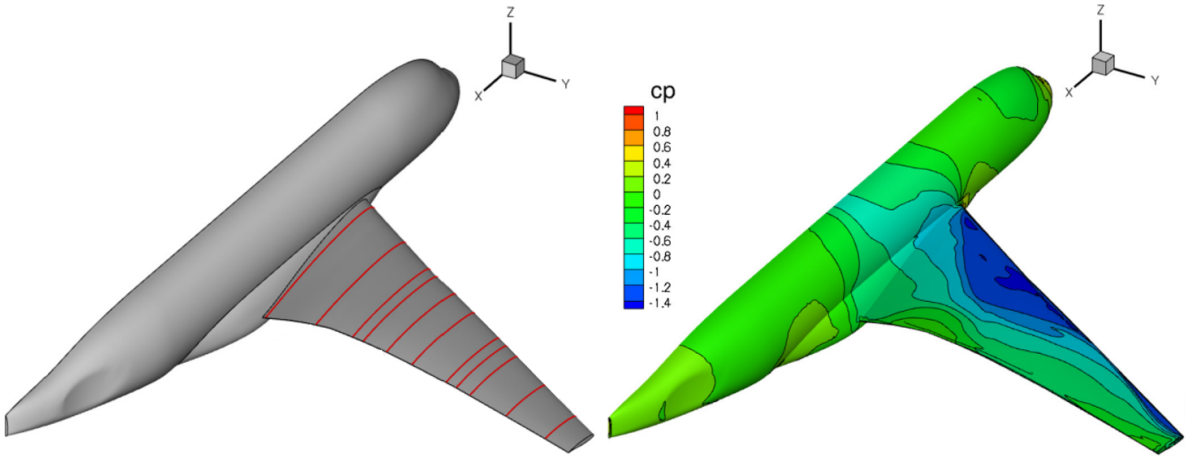


Figure 2.10: 3D aircraft wing-body model indicating sections along which experimental data is available (left), resulting surface pressure distribution from GPOD; fused experimental and CFD data (right) at $\alpha = 7.9°$ and $M = 0.85$, adapted from Mifsud et al. (2019)

Integration of the reconstructed surface pressure over the complete model surface resulted in both lift and drag coefficients closer to balance measurements compared to the CFD simulations alone. The research is thus able to achieve its goal of enhancing aerodynamic coefficients determination by fusion of numerical and experimental results however does not seem to aim for sparsity in the process considering the relatively large number of sensors. Leveraging the information in POD modes by means of smart sensor placement could possibly have allowed for a reduction in the number of sensors while still enhancing CFD results. This seems to have been overlooked as no reasoning behind both the number of sensors and their locations is provided.

The influence of sensor locations and sensor budget for gappy POD reconstructions is considered in the research by Xing et al. (2022). In their paper, reconstructions of the surface pressure on a LNG (Liquefied Natural Gas) carrier are performed for various angles of incidence. Training data is obtained from CFD simulations while sparse surface pressure measurements are obtained experimentally with a varying sensor budget. Using gappy POD, sparse surface pressure measurements are used to successfully recover the complete pressure distribution in accordance with the CFD simulations. By changing the number and layout of sparse measurements however, it was noticed that reconstructions almost indistinguishable from the CFD results are obtained at vastly different sensor numbers depending on the placement strategy. This is illustrated in Figure 2.11 which shows the drag estimations of gappy POD at various heading angles compared to CFD results. Both the drag curves on the left and right of Figure 2.11 are obtained using 32 sensors scattered across the LNG carrier surface however the sensor layout is changed from layout 1, shown on the top left, to layout 2 as shown on the top right. No physical reasoning behind the changed sensor locations between layouts is provided however unlike layout 1, layout 2 does not contain any sensors positioned on the carrier's symmetry plane which seems to have affected the gappy POD performance for the worse. Similar accuracy to layout 1 was found to be achievable without placing sensors on the symmetry plane, however this required 46 sensors instead.

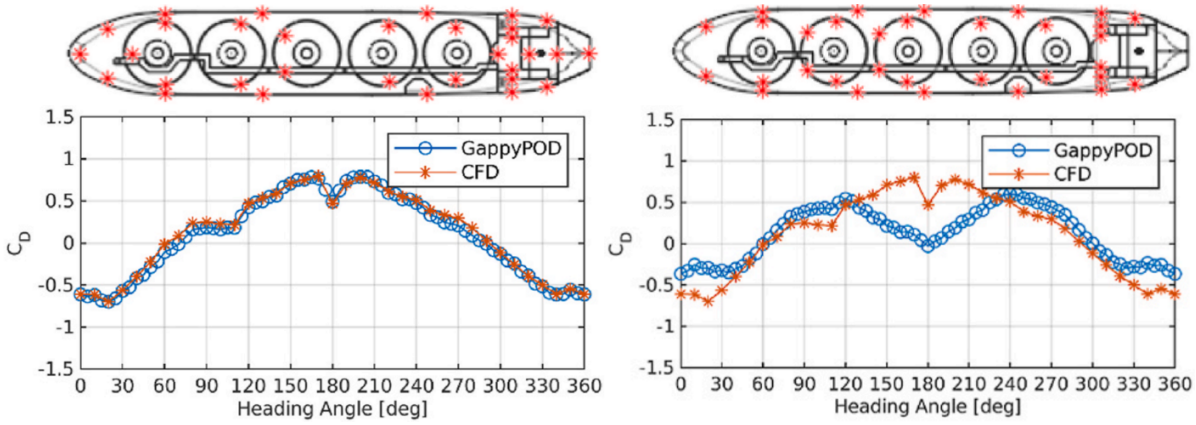


Figure 2.11: Gappy POD LNG carrier drag estimation results using 32 sensors and different sensor layouts: layout 1 left, layout 2 right. Adapted from Xing et al. (2022)

The findings by Xing et al. highlight the effect sensor locations have on reconstruction accuracy but are unfortunately not a result of a systematic sensor placement approach making use of the flow physics captured by the POD modes. This despite earlier findings by for example Cohen et al. (2003) and Yildirim et al. (2009) that the extrema of POD modes provide very good sensor positions for accurate reconstructions. A first algorithm for finding gappy POD sensor locations was described by Willcox (2006). Willcox reasoned that for accurate determination of the coefficients in $\boldsymbol{b}$ from Equation 2.16, matrix $\boldsymbol{M} = \boldsymbol{\Theta}^T\boldsymbol{\Theta}$ should be well-conditioned. This can be measured in terms of the condition number κ which indicates how sensitive a matrix inversion is to errors (Manohar et al., 2018). The condition number of $\boldsymbol{M}$ should be kept low, ideally as close to one as possible, and therefore Willcox proposed an algorithm to evaluate the condition number after placing a single sensor onto every possible location, picking the one resulting in the smallest $\kappa(\boldsymbol{M})$ and repeating for the remainder of the sensor budget. This is a greedy algorithm as it only considers the best sensor location in the current iteration for a certain objective while ignoring the effect this choice might have on future iterations of sensor placement. The condition number of a matrix $\boldsymbol{X}$ satisfying $\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{I}$ equals one and therefore minimizing the condition number can be seen as keeping the POD modes in $\boldsymbol{\Theta}$ as orthogonal as possible (orthogonality being lost in the POD modes due to the sparse sampling). Brunton and Kutz (2019) illustrate the performance of the 'minimum condition number' (MCN) algorithm compared to placing sensors on modal optima in the figure provided below. Note the difference in the notation of the POD bases; $\boldsymbol{\Psi}$ instead of $\boldsymbol{\Phi}$.

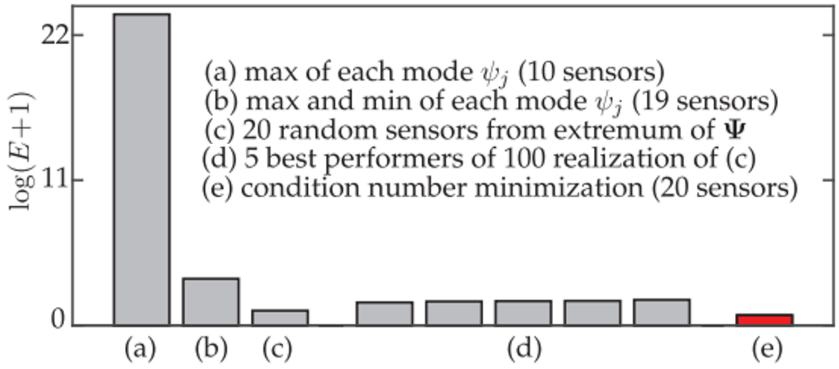


Figure 2.12: Reconstruction performance of gappy POD on test case using MCN and mode extrema sensor locations, adapted from Brunton and Kutz (2019)

Since the introduction of the systematic sensor placement procedure for GPOD reconstructions in the form of the MCN algorithm, several other sensor placement algorithms have been suggested specifically for use with POD bases. In the same paper in which the MCN approach is introduced by Willcox, already a modified version combining both MCN and POD optima in the sensor placement algorithm was found to perform better. Four sensor placement algorithms are compared in terms of reconstruction accuracy and have their working principles explained in the paper by Jayaraman et al. (2019). It includes the MCN algorithm, random sensor placement, QR factorization with column pivoting (QR, QR-CP or QDEIM (Drmač and Gugercin, 2016), will be explained in Section 2.7) and a Discrete Empirical Interpolation Method (DEIM (Chaturantabut and Sorensen, 2010)) for sensor placement. All algorithms are compared on a low rank case: 2D cylinder flow and a high rank case: global Sea Surface Temperature maps (SST). Low-rank referring to a fast decay of the problem specific eigenvalues making it 'easier' to reconstruct using a smaller number of sensors. Random sensor placement works very well for compressed sensing which includes reconstructions based on a generic set of bases such as in the Fourier transform example. For a tailored set of bases obtained through POD however, Jayaraman et al. found that random sensor placement generally results in relatively poor reconstructions. Surprisingly, the MCN approach was found not to perform better than random sampling for the cylinder flow case and was not even included in the SST case due to its exceptionally high computational expense for high-dimensional systems. DEIM was found to perform best in both cases followed closely by the QR algorithm. This conclusion is slightly different from what was found by Brunton and Kutz and Drmač and Gugercin who concluded that 'QDEIM may improve error performance over standard DEIM'. The performance of the sensor placement algorithms vary for each application however the QR decomposition with column pivoting is stated to provide a robust mathematical architecture for sensor placement which is easy to implement and will therefore be explained in more detail below.

2.7. Greedy QR-CP algorithm

Obtaining a full-state reconstruction from sparse measurements using gappy POD requires solving the linear system as shown in Equation 2.16 for $\boldsymbol{b}^T$, which through $\boldsymbol{\Theta}$ and $\boldsymbol{w}$ depend on the chosen sensor locations. Equation 2.16 presents a generic for-

mulation for the linear system encountered using gappy POD in the sense that matrix $\boldsymbol{\Theta}$ can both be square or non-square corresponding to using $n_s = K$ or $n_s > K$ respectively. For $n_s = K$, the number of sensors equals the number of modes used in the reconstruction. In this case, solving the least squares solution as in Equation 2.16 is redundant and $\boldsymbol{\Theta b}^T = \boldsymbol{w}^T$ could be solved instead. Using $n_s = K$ provides the most sparse reconstruction solution for a given number of modes K.

$\boldsymbol{\Theta}$ was defined in Section 2.5 to consist of the entries in the first K columns of $\boldsymbol{\Phi}$ supported by the n_s sparse measurements: $(\boldsymbol{\Phi}_{1:K})_{[n_s]}$. This follows the definition used by Everson and Sirovich; however it seems to obscure the importance of the sensor placement. Instead, an alternative but equivalent definition of $\boldsymbol{\Theta}$ is $\boldsymbol{C\Phi}_{1:K}$, where $\boldsymbol{C} \in \mathbb{R}^{K \times n_s}$ consists of K rows of the $n \times n$ identity matrix.

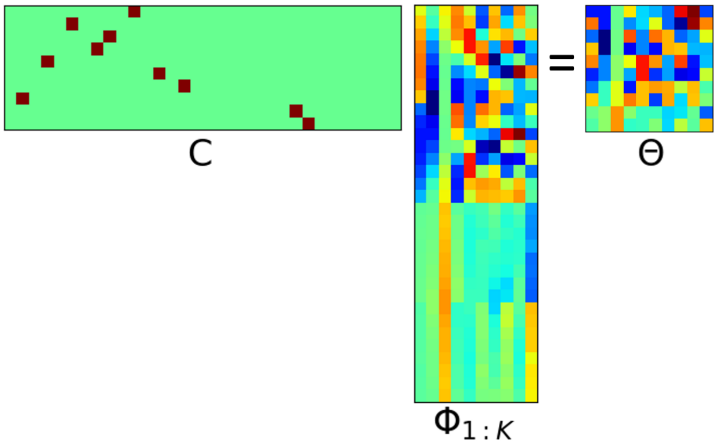


Figure 2.13: Definition of matrix $\boldsymbol{\Theta}$ using 'sensor matrix' $\boldsymbol{C}$

From the schematic in Figure 2.13 it can be seen how matrix $\boldsymbol{C}$ samples the spatial modes of $\boldsymbol{\Phi}_{1:K}$, effectively providing sensor locations in the form of the column indices at which $\boldsymbol{C}$ contains ones. For the specific case using $n_s = K$, the condition number of $\boldsymbol{\Theta}$ should thus be minimized as it requires inversion. An indirect bound on the condition number of $\boldsymbol{\Theta}$ can be achieved by optimizing its spectral contents through maximizing the product of its eigenvalues (Manohar et al., 2018). The reduced QR factorization, decomposes any real matrix $\boldsymbol{X}$ into a unitary matrix $\boldsymbol{Q}$ and upper triangular matrix $\boldsymbol{R}$ such that $\boldsymbol{X} = \boldsymbol{QR}$ (Trefethen and Bau, 1997). Therefore

$$|\det(\boldsymbol{X})| = |\det(\boldsymbol{R})| = |\prod_i r_{ii}| = |\prod_i \lambda_i| \tag{2.18}$$

where r_{ii} are the diagonal entries of $\boldsymbol{R}$ and λ_i eigenvalues. The diagonal entries of $\boldsymbol{R}$ are stored in no particular order and thus generally the product of eigenvalues is not maximized. If instead combined with column pivoting such that $\boldsymbol{XD} = \boldsymbol{QR}$, with $\boldsymbol{D}$ a square column permutation matrix, the sequence of r_{ii} can be controlled to be in decreasing order (Jayaraman et al., 2019). By realizing that $\det(\boldsymbol{\Theta}) = \det(\boldsymbol{\Theta}^T) = \det((\boldsymbol{\Phi}_{1:K})^T \boldsymbol{C}^T)$, it can be seen how QR factorization with column pivoting of $(\boldsymbol{\Phi}_{1:K})^T$, yielding $(\boldsymbol{\Phi}_{1:K})^T \boldsymbol{D} = \boldsymbol{QR}$, thus maximizes the product of eigenvalues of $\boldsymbol{\Theta}$ if $\boldsymbol{C}^T$ is chosen to be the first n_s columns of $\boldsymbol{D}$ (Williams et al., 2022). Such a QR factor-

ization with column pivoting for the identification of sensor locations is illustrated in Figure 2.14.

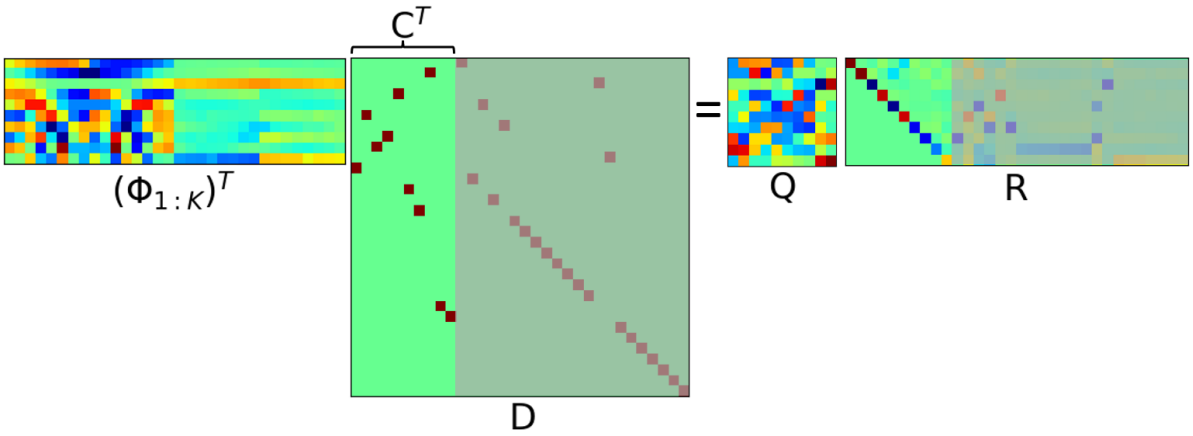


Figure 2.14: Schematic illustration of QR factorization with column pivoting for the determination of $\boldsymbol{C}^T$

Matrix $\boldsymbol{D}$ which reorders the columns of $(\boldsymbol{\Phi}_{1:K})^T$ such that r_{ii} is decreasing is obtained by sequentially selecting a new column of $(\boldsymbol{\Phi}_{1:K})^T$ with maximal two-norm after which the orthogonal projection of the remaining columns onto this pivot column is subtracted from the remaining columns (Manohar et al., 2018). This is a greedy approach which favors sensor locations that experience high modal activity across the available modes. Controlling the order of the diagonal entries in $\boldsymbol{R}$ to be decreasing maximizes the submatrix volume. Since matrix volume is equal to the absolute value of the determinant, decreasing $\kappa(\boldsymbol{\Theta})$ through optimization of its spectral contents and maximizing 2.18 preserves mode orthogonality. This can be recognized through physical interpretation of matrix volume depicted below in Figure 2.15 for a $\boldsymbol{\Theta} \in \mathbb{R}^{2\times 2}$ example.

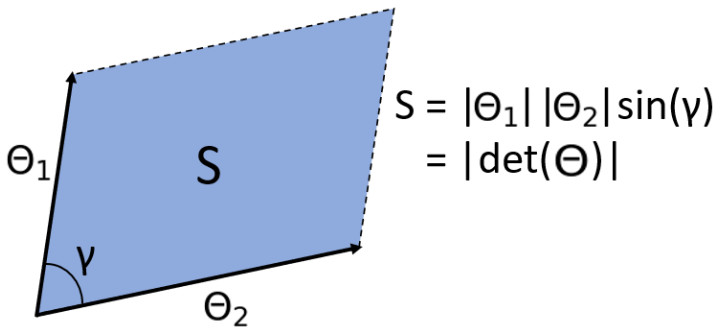


Figure 2.15: Relation between matrix volume and determinant

The absolute value of the determinant of $\boldsymbol{\Theta}$ is equal to the volume of the parallelepiped determined by either its columns or rows. In 2D, the columns of $\boldsymbol{\Theta}$ form a parallelogram with 'volume' or area S equal to $|\boldsymbol{\Theta}_1||\boldsymbol{\Theta}_2|\sin\gamma$. Maximization of 2.18 therefore maximizes S which can be realized through keeping the magnitude of the columns large or by preserving their orthogonality. Note that this assumes $\boldsymbol{\Theta}$ to be square.

The QR-CP algorithm thus provides nearly optimal sensor locations tailored to a specific basis while avoiding a combinatorial problem associated with direct maximization of 2.12 through a greedy approach. It is computationally efficient and simple to implement as the procedure is included in most scientific programming libraries (Brunton and Kutz, 2019). The above description focuses on the square $\boldsymbol{\Theta}$ case with $n_s = K$ however the QR factorization for sensor placement can also be used in the oversampled case where $n_s > K$. Instead of decreasing $\kappa(\boldsymbol{\Theta})$, it can be shown that this requires the condition number of $M = \boldsymbol{\Theta}^T\boldsymbol{\Theta}$ to be bounded through the QR factorization with column pivoting of $\boldsymbol{\Phi}_{1:K}(\boldsymbol{\Phi}_{1:K})^T$ (Jayaraman et al., 2019), as the gappy POD reconstruction effectively reverts back to the least squares problem from Equation 2.16.

2.8. Reconstruction example

The snapshot POD, gappy POD and QR-CP algorithm have been discussed individually however it is the combination of these which would allow for accurate sparse reconstructions on a tailored set of bases. To illustrate this, a reconstruction of circular cylinder flow at $Re = 100$ will be shown. The data consists of DNS flowfields of the velocity and vorticity at $m = 150$ time instants (snapshots) separated by $\Delta t = 0.02\,\text{s}$. Each flowfield is calculated on a grid of size $n_x \times n_y = 450 \times 200$ hence $n = 90,000$. At this Reynold's number, circular cylinder flow is characterized by laminar periodic vortex shedding. This dataset is obtained and made publicly available by Kutz et al. (2016).

2.8.1. Eigenvalues and modes from POD

An example of an instantaneous horizontal velocity component flowfield is shown on the top left of Figures. 2.18 and 2.19. After subtraction of the temporal mean, the ensemble of all horizontal velocity component snapshots is stored in a matrix for POD to applied on following the description provided in Section 2.3. This yields the matrix containing temporal coefficients $\boldsymbol{A} \in \mathbb{R}^{m\times m}$ and the matrix containing the orthogonal spatial modes $\boldsymbol{\Phi} \in \mathbb{R}^{n\times m}$, ordered corresponding to the magnitude of corresponding eigenvalues. These eigenvalues provide insight into the relative turbulent kinetic energy capture between modes for this specific problem and are shown in Figure 2.16.

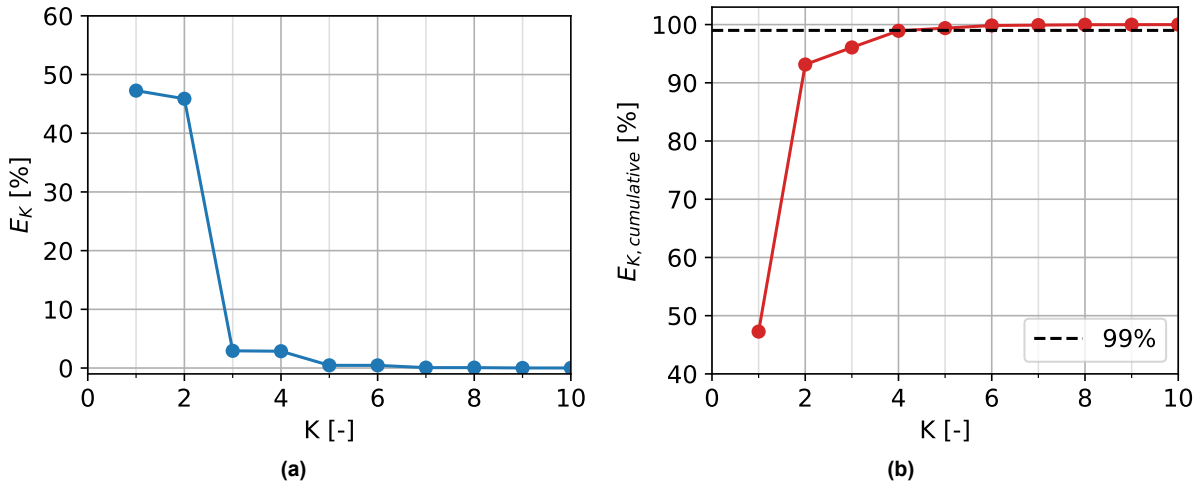


Figure 2.16: Individual (a) and cumulative (b) energy capture of the eigenvalues corresponding to U-velocity cylinder flow data

Figure 2.16a shows energy capture of individual eigenvalues and Figure 2.16b the cumulative energy capture for a given number of eigenvalues. Both are scaled by the total energy such that:

$$E_K = \frac{\lambda_K}{\sum_{i=1}^{m} \lambda_i}, \quad E_{K,cumulative} = \sum_{k=1}^{K} \frac{\lambda_k}{\sum_{i=1}^{m} \lambda_i} \tag{2.19}$$

The magnitude of the first few eigenvalues can be seen to decrease rapidly which is indicative of an inherently low rank system. Already at 5 out of the total of 150 eigenvalues, the cumulative energy capture exceeds $99\,\%$. The eigenvalues for this particular case come in pairs which is especially clear from Figure 2.16a. These pairs correspond to harmonics of the dominant vortex shedding frequency as will be clear from the spatial modes.

Each of the eigenvalues has its corresponding spatial mode and set of temporal coefficients. The first few spatial modes are therefore of major importance since they represent the underlying physics sampled throughout the training data to a large extent, which will be leveraged in gappy POD reconstructions. The two most dominant spatial modes or eigenmodes for the horizontal velocity component data are shown below together with their temporal coefficients in Figure 2.17.

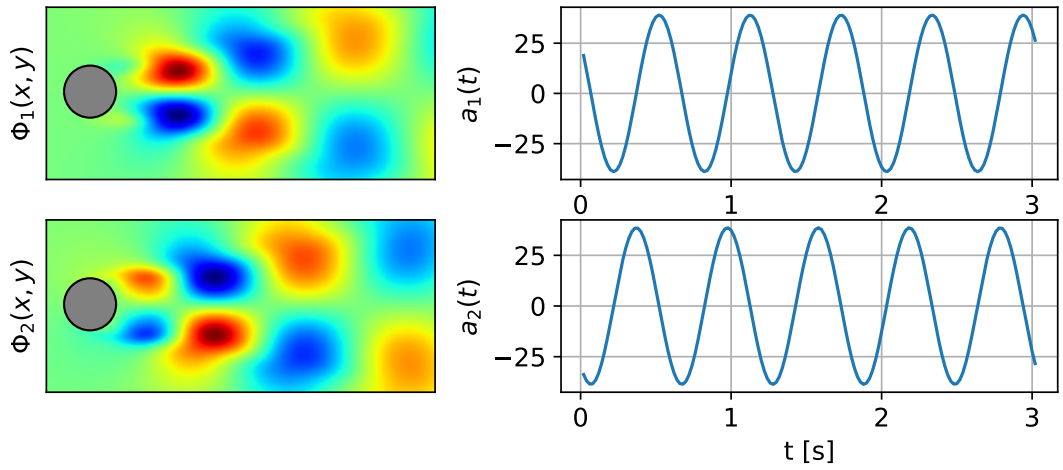


Figure 2.17: First and second spatial POD modes (left) and corresponding temporal coefficients (right) for cylinder flow

The periodic vortex shedding, typical to cylinder flow, is recognizable in the patterns of the first two spatial modes and the oscillations in the temporal modes. These modes and coefficients are useful for recognizing coherent structures and also show how a reduced order approximation of the data in the form of a POD projection begins to approximate the original data quite well at a significantly lower number of modes than available. A rank two POD projections would be obtained as a linear combination of the above shown modes and coefficients: $a_1(t)\Phi_1(x,y) + a_2(t)\Phi_2(x,y)$. Even though the the spatial modes are similar in shape, the phase difference in the temporal coefficients together with the offset in modal extrema between the eigenmodes allow for the rank two projection to capture the advection of vortical structures already. Increasing the number of modes and corresponding coefficients involved in the POD projection allows for additional phenomena to be represented which increase accuracy with respect to the original snapshots.

2.8.2. Sparse reconstruction using GPOD

The obtained set of eigenmodes, stored as the columns of Φ can be used as a library to infer full state reconstructions from sparse measurements using the gappy POD procedure discussed in Section 2.5. Given n_s measurements and K eigenmodes, GPOD will provide coefficients such that a linear combination of the eigenmodes will approximate the measurements optimally in a least squares sense. Increasing the number of modes generally increases reconstruction accuracy through increased energy capture in the additional modes however requires more sparse measurements to ensure a unique solution for the coefficients ($n_s \geq K$). A reconstruction of an instantaneous velocity field at $t = 2$ using five sensors and five POD modes is shown in Figure 2.18. Note that this is a snapshot originally included in the training data however this is not required.

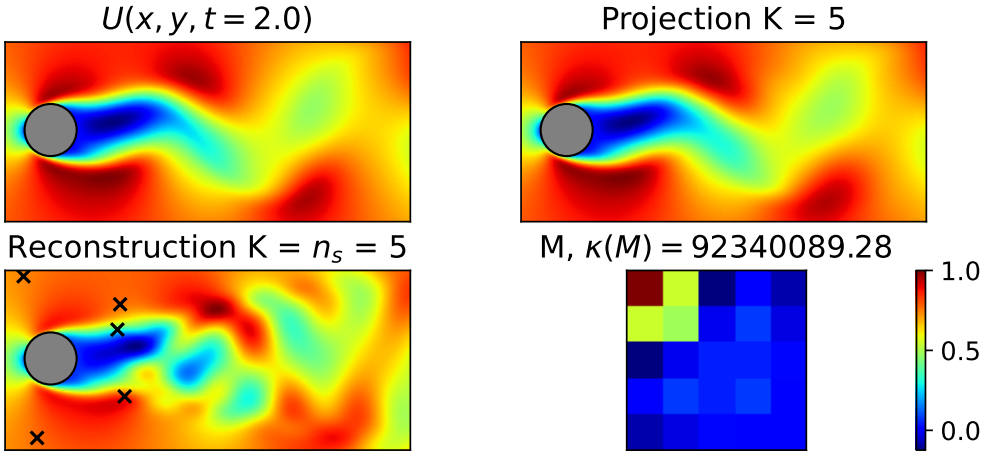


Figure 2.18: Instantaneous DNS solution of horizontal velocity field (top left), rank 5 POD projection (top right), gappy POD reconstruction using first 5 POD bases + 5 random sensors (bottom left) and schematic gappy POD matrix $\boldsymbol{M}$ with its condition number κ (bottom right).

The POD projection shown at the top right makes use of the same five POD modes as the GPOD reconstruction and can be seen as the best attainable by any sparse reconstruction. The projection is visually almost indistinguishable from the original as a result of using bases with a cumulative energy contents of over $99\,\%$. The same can not be said about the sparse reconstruction which shows significant discrepancies with respect to the original when using the sensor locations indicated by the crosses. These sensor locations are the result of a particularly bad random sensor placement realization. Two out of the five sensors are positioned in a region of almost zero modal activity across the modes near the left border of the domain. These essentially informative sensors result in an ill-conditioned reconstruction problem with a large condition number of matrix $\boldsymbol{M}$, see bottom right of Figure 2.18. The condition number of $\boldsymbol{M}$ serves as a proxy measure for mode orthogonality and reconstruction accuracy, which should be kept as close to one as possible through keeping $\boldsymbol{M}$ as close to identity as possible. Note that as mentioned in Section 2.7, inversion of $\boldsymbol{M}$ is not necessary for the specific case where $n_s = K$ for which only $\boldsymbol{\Theta}$ can be inverted instead. Matrix $\boldsymbol{M} = \boldsymbol{\Theta}^T\boldsymbol{\Theta}$ however clearly illustrates how reducing the condition number keeps modes in $\boldsymbol{\Theta}$ orthogonal by approaching identity.

To avoid an ill-conditioned $\boldsymbol{M}$ or $\boldsymbol{\Theta}$, the QR factorization including column pivoting as discussed in Section 2.7 can be used to provide nearly optimal sensor locations from the set of tailored POD bases. The reconstruction performed using such 'QR sensors' instead is shown below in Figure 2.19.

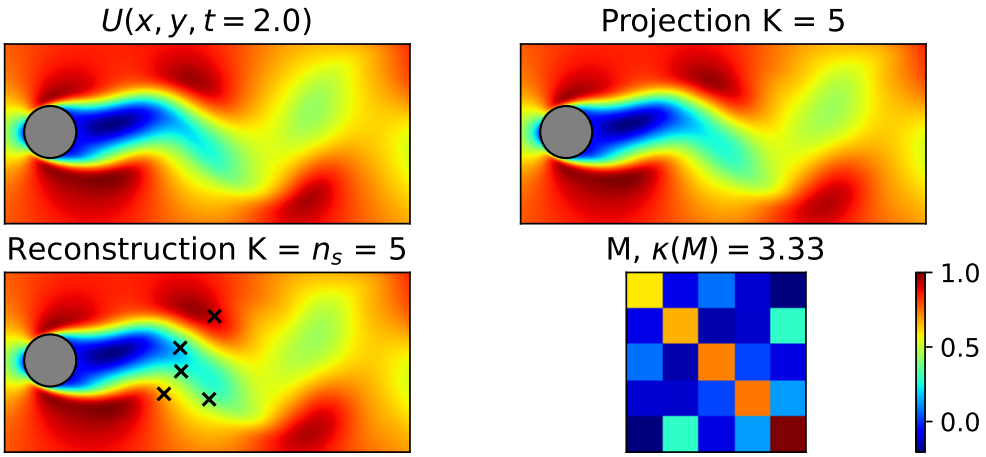


Figure 2.19: Instantaneous DNS solution of horizontal velocity field (top left), rank 5 POD projection (top right), gappy POD reconstruction using first 5 POD bases + 5 QR sensors (bottom left) and schematic gappy POD matrix M with its condition number κ (bottom right).

All of the five sensors are now placed within the wake of the cylinder where modal activity is generally high. The resulting reconstruction is almost identical to the projection which demonstrates the benefits of proper sensor positioning for reconstructions on a tailored set of bases. Different from the previous example, matrix M is clearly diagonally dominant and thus the modes involved in the reconstruction have been kept almost orthogonal over the sparse support of the sensors. This is further exemplified by the condition number of 3.33 being orders of magnitude smaller than the practically singular matrix shown previously. The realization of random sensor placement shown was a particularly bad example of random sensor placement for reconstructions. The average and 5th percentile (P_5) root mean square error (RMSE) in the horizontal velocity field for 100 realizations of random sensor placements is shown below in Figure 2.20 while varying sensor budget. It also includes the RMSE if instead the QR algorithm is used for sensor positioning. Note that the freestream velocity is $1\,\mathrm{m\,s^{-1}}$.

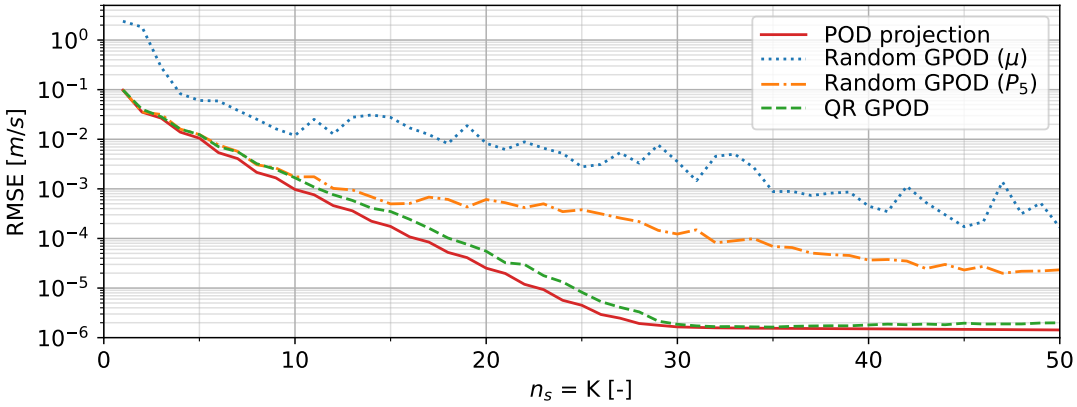


Figure 2.20: Reconstruction root mean square error, random and QR sensor placement

For a relatively small sensor budget, say $n_s < 10$, the 100 random realizations act as a brute force optimal sensor positioning search which is unfeasible for practical use. In this interval, the five best random realizations (5th percentile) are expected

to have identified a nearly optimal sensor configuration and virtually overlaps the QR GPOD curve. This shows how the QR factorization with column pivoting provides a systematic approach for finding almost optimal sensor locations. For larger n_s, 100 realizations is likely not enough for consistent brute force 'quasi optimal' sensor positioning and the 5th percentile curve diverges from the projection. The QR curve however stays close to the projection meaning that the accuracy loss in the sparse reconstruction due to 'gappiness' is kept low.

2.9. Square cylinder flow

Compared to the streamlined bodies as used in several of the referenced papers, bluff body aerodynamics characterized by separation and relatively large, unsteady wakes resulting in complex flow phenomena which are expected to increase the inherent rank of such systems, forming a challenge for POD reconstructions. Regardless, with pressure drag dominating the total experienced drag of bluff bodies, they lend themselves well for drag determination based on the pressure distributions. Sparse, surface pressure based, drag determination on bluff bodies was attempted in the earlier mentioned research by Liu et al. (2021) and Xing et al. (2022) however not while incorporating both the POD framework and the ideas for tailored sensor placement. Bluff body flow is of significant engineering interest, from wind engineering to (motor)sports-aerodynamics and thus a benchmark bluff body object in the form of a square cylinder will be used for sparse surface pressure distribution reconstructions and drag determination in an experimental setting. Square cylinder flow involves large separated regions, flow reattachment and vortex shedding depending on the angle of attack which have been studied extensively over many papers dedicated to the subject. The distinct pressure distributions at different angles of attack provide an experimental test case for sparse sensor placement and reconstructions on a more complex angle of attack basis compared to earlier used streamlined objects.

2.9.1. Flow regime classification

The angle of attack and geometric definitions used for square cylinder flow are shown in Figure 2.21.

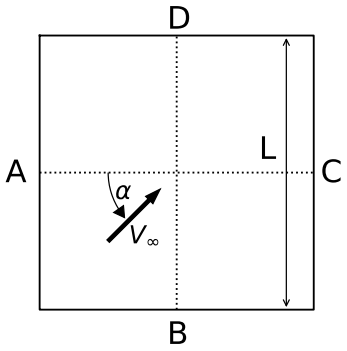


Figure 2.21: Square cylinder angle of attack and face definitions

The four faces are denoted A, B, C and D. At an angle of attack of $\alpha = 0°$, face A is orthogonal to the freestream velocity vector. At $\alpha = 90°$, face B is orthogonal to the

freestream velocity vector instead. The length of each of the four faces is denoted as L.

Within the angle of attack range $0° \leq \alpha \leq 45°$, the flow around a two-dimensional square cylinder is characterized by different flow topologies. These have been classified by both Igarashi (1984) and Huang et al. (2010) based on the observed time-averaged flow patterns. Igarashi recognized four flow patterns while Huang et al. merged two out of those four to obtain the following three with corresponding illustrations in Figure 2.22:

- Subcritical flow - $0° < \alpha < \alpha_{crit}$
- Supercritical flow - $\alpha_{crit} < \alpha < 45°$
- Wedge flow - $\alpha = 45°$

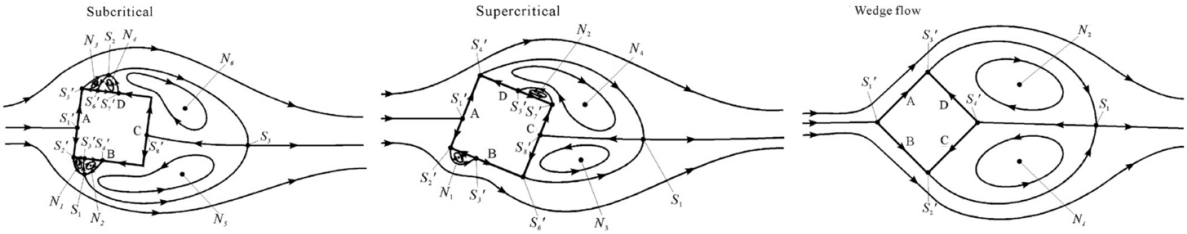


Figure 2.22: Subcritical- (left), superciritcal- (middle) and wedge- (right) flow patterns. For information regarding indicated locations, see Huang et al. (2010)

The subcritical flow regime extends until the critical angle of attack: α_{crit}. At the critical angle of attack, the fluctuating components of the pressure coefficient reaches a minimum at faces C and D while the average pressure reaches a maximum, typically resulting in minimum drag. The critical angle of attack is somewhat sensitive to the Reynolds number and freestream conditions ((Carassale et al., 2014),(Chen and Liu, 1999)) however generally lies within the range of $10° < \alpha < 15°$ as apparent from an extensive collection of experimental drag curves presented in Roosenboom (2005). Huang et al. experienced a critical angle of attack of $15°$. The subcritical regime is characterized by a forward stagnation point on face A and separation at both upstream corners of the cylinder. Faces B, C and D experience separated flow and the wake consists of two counter rotating recirculation regions. This description agrees well with flow visualizations obtained by van Oudheusden et al. (2008) for this flow regime which are shown in Figure 2.23.

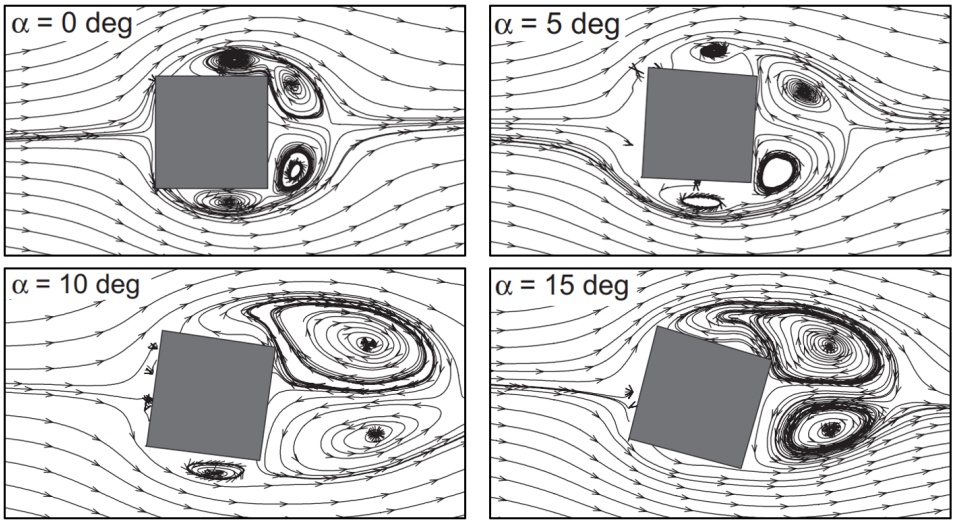


Figure 2.23: Subcritircal flow patterns as obtained by van Oudheusden et al. (2008) at $Re = 20,000$

Besides the large recirculation regions in the wake, two shallow ones are present on the upper and lower face of the cylinder for $\alpha = 0°$ and $5°$. Between $\alpha = 5°$ and $10°$ the bubble on face D merges with the wake. Between $\alpha = 10°$ and $15°$, the bubble on face B gets pinched of by the reattaching shear layer near the bottom downstream corner. This reattachment of the shear layer at the lower side of the cylinder coincides with the aforementioned critical angle of attack.

Roosenboom (2005) performed PIV measurements in both the subcritical and supercritical regime. The results of these experiments at $\alpha = 30°$ are shown on the left of Figure 2.24 and illustrate the main characteristics of the supercritical flow regime well.

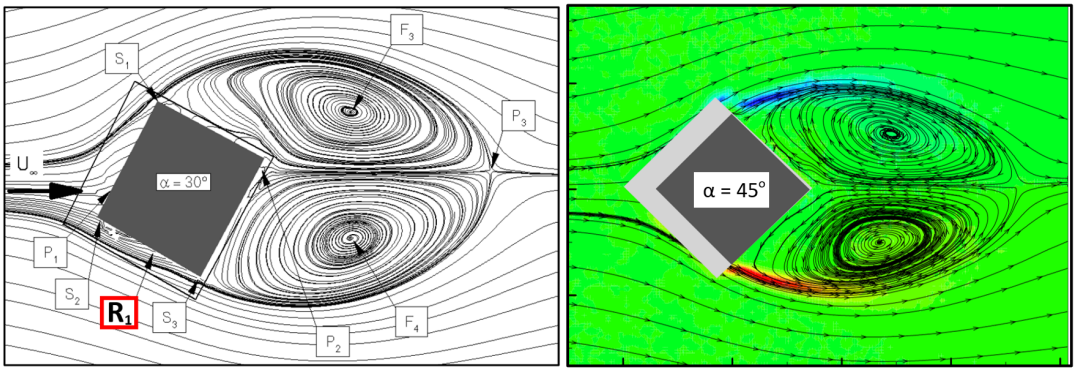


Figure 2.24: Flow pattern for $\alpha = 30°$ (left) and $\alpha = 45°$ (right, background colours follow from vorticity) at $Re = 20,000$, adapted from Roosenboom (2005)

In the supercritical angle of attack range, the forward stagnation point is still located at face A and the flow separates from the two upstream corners. Often a small recirculation bubble occurs at the downstream part of face D (Mueller, 2012), however this is not clearly visible in the left of Figure 2.24. The start of this flow regime was marked by the reattachment of the shear layer on the downstream corner of face B. From α_{crit} towards $\alpha = 45°$, the point of reattachment traverses face B in the upstream direction, thus decreasing the size of the lower recirculation bubble. This is clearly visible in

the position of the reattachment point indicated in red in Figure 2.24 being almost at the center of face B. While for the subcritical regime, the Strouhal number increases rapidly for increasing angle of attack until a maximum at $\alpha = \alpha_{crit}$ ($St \approx 0.18$), it remains roughly constant throughout the supercritical regime ($St \approx 0.17$, Chen and Liu (1999)).

The right of Figure 2.24 shows the flow patterns at $\alpha = 45°$ obtained by Roosenboom. Huang et al. classifies flow at exactly $\alpha = 45°$ as wedge type flow but others treat the boundary between the supercritical and wedge type flow regime more like a transition with Igarashi mentioning wedge-like flow from $\alpha = 35°$ onwards for example. According to Huang et al., the wedge flow type for square cylinders is characterized by bifurcation of the flow on the corner where faces A and B merge, resulting in a stagnation point on the 'nose' of the cylinder. The flow remains attached while traversing faces A and B however detaches where these faces merge with face D and C respectively. The two large recirculation regions are symmetric and engulf faces C and D such that no shallow secondary recirculation regions are expected on any of the cylinder faces.

2.9.2. Surface pressure distributions

The publication by Igarashi (1984) presents the pressure distribution around the square cylinder perimeter for the full angle of attack range from $\alpha = 0°$ to $45°$. These distributions, which are shown in Figure 2.25, conform well with the flow visualizations shown earlier but provide a different perspective on the aforementioned square cylinder flow regimes.

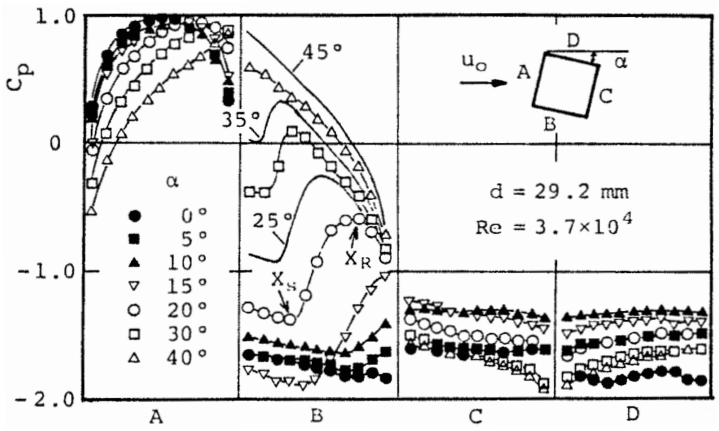


Figure 2.25: Square cylinder surface pressure distributions at various angles of attack, adapted from Igarashi (1984)

Except for $\alpha = 45°$, the flow stagnates on face A of the cylinder resulting in the pressure coefficient reaching a value of 1. In general for $\alpha < 45°$, face A can be seen as the pressure side with a relatively high surface pressure where the distribution skewes towards the corner between faces A and B as the angle of attack increases and the stagnation point travels in this direction.

Face B experiences fully separated, partially separated and fully attached flow depending on the flow regime. In the subcritical flow regime, the shear layer, which

detached from corner AB, does not reattach resulting in a shallow recirculation bubble and lower than freesteam surface pressures. Increasing the angle of attack, brings the shear layer closer to reattachment on corner BC leading to increased pressure on the downstream part of face B. Above α_{crit}, the shear layer reattaches on face B and the reattachment point traverses in upstream direction depending on the angle of attack. This shift in the reattachment point is clearly recognizable through the position of the strong pressure gradient on face B. At $\alpha = 45°$, the reattachment point reaches the upstream corner on face B leading to completely attached flow and a symmetric pressure distribution to face A.

The flow over faces C and D is separated from the surface across all angles of attack and therefore the pressure distribution shows relatively little variation. For $\alpha = 0°$, the symmetric recirculation regions result in backflow impinging on the face C surface raising the pressure with respect to faces B and D. Above $\alpha = 5°$, the shallow bubble on face D merges with the recirculating region in the wake and the pressure across faces C and D evens out. At the critical angle of attack, the pressure on faces C and D reaches a maximum. Within the the supercritical regime, a shallow recirculation bubble might appear near the corner between faces C and D, producing a local dip in the surface pressure. At $\alpha = 45°$, the pressure distribution on faces C and D is symmetric as expected and reaches a minimum at the corner between the faces.

2.9.3. Drag coefficients

Below, experimentally obtained drag coefficients of a square cylinder at incidence are reported in Figure 2.26. These results were obtained by Lee (1975), Tamura and Miyagi (1999) and Carassale et al. (2014). The Reynolds numbers at which the experiments have been performed are stated as well however the effect of the Reynolds number on the results will not be considered in this report (Reynold's number effects are discussed by Carassale et al. (2014) and Chen and Liu (1999)). Note that the drag coefficients in the figure below and throughout this report are defined with respect to the cylinder side length L instead of the projected area normal to the freestream.

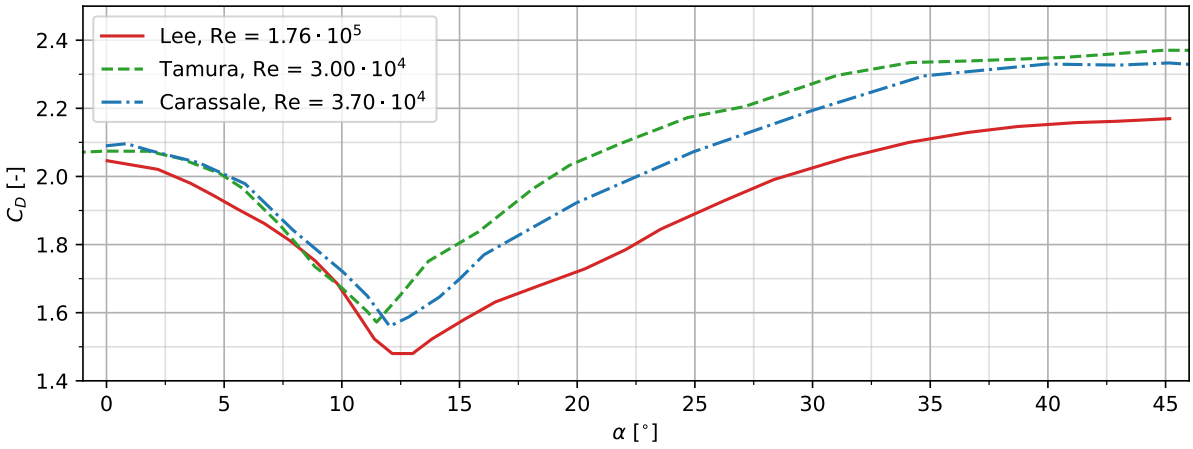


Figure 2.26: Drag coefficient versus angle of attack for square cylinder flow

With the total drag of a square cylinder being dominated by pressure drag, the be-

haviour of the drag curve at varying incidence angle correlates well with the pressure distributions shown in Figure 2.25. Within the subcritical regime, square cylinder flow is characterized by a negative slope in the drag. Lee postulated that this is caused by the vortex formation region moving downstream which increases the pressure on the downstream face. Without linking this directly to the drag, van Oudheusden et al. discussed and visualized this phenomena as the wakes in Figure 2.23 can be seen to elongate for increasing α within the subcritical range. The increasing pressure in the back of the cylinder reduces the net force in freestream direction and therefore decreases drag. Typically, at $\alpha = 0°$, a drag coefficient slightly above 2 is measured which drops roughly by 0.6 when reaching α_{crit}. At this critical angle of attack, the base pressure is at its maximum as a result of the wake elongation resulting in minimum drag. As mentioned before, α_{crit} typically occurs for $10° < \alpha < 15°$. Above the critical angle of attack, increasing α instead shortens the wake, decreasing base pressure and consequently increasing drag. A change in the sign of the drag curve slope is therefore present at the critical angle of attack which the shown curves can be seen to adhere to. The drag increases for increasing α until a maximum at $\alpha = 45°$. At $\alpha = 45°$, the situation is very sensitive to slight asymmetries and different freestream conditions (Mueller, 2012) as is evident from the variance in the reported drag coefficients close to this angle of attack (see for a more extensive collection of experimentally obtained drag curves also Roosenboom (2005)).

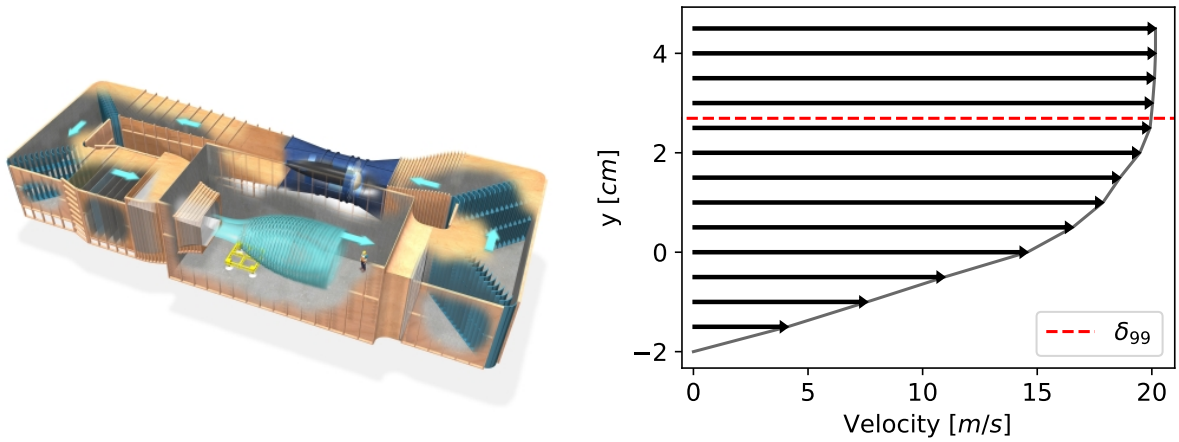
3

Experimental setup and procedures

In this chapter, the steps undertaken to be able to realize a wind tunnel setup capable of providing surface pressure and force balance measurements on a square cylinder model are described. Wind tunnel tests have been performed in the Open Jet Facility at the TU Delft of which information and an illustration are provided in Section 3.1. Dimensions and drawings of the used square cylinder model with variable pressure tap layouts are discussed in Section 3.2. The model is equipped with taps connected to pressure transducers/modules and is placed on a force balance. Specifications of this equipment are given in Section 3.3. Besides the model and measurement equipment, the experimental setup additionally consisted of a turntable and end plates to emulate two-dimensional flow. A description and drawing of how these aspects come together is provided in Section 3.4. Finally, in Section3.5 a summarized version of the test matrix used to survey the parameter space is presented. It also includes a description of the angle of attack calibration, flow and sampling parameters used during the measurements and used corrections for the influence of model end plates and an experienced blockage effect.

3.1. Wind tunnel

The experimental campaign consists of measurements on a square cylinder test object which is placed in the the Open Jet Facility (OJF) of TU Delft. The OJF is shown schematically in Figure 3.1a and is an open test-section, closed-loop type wind tunnel with an octogonal outlet of size $2.85\,\text{m} \times 2.85\,\text{m}$. It is driven by a $500\,\text{kW}$ fan capable of providing flow at a maximum velocity of $35\,\text{m}\,\text{s}^{-1}$ in the test Section. A $350\,\text{kW}$ radiator at the end of the test section is able to compensate for added heat to the flow which keeps the temperature constant during operation (Delft University of Technology, 2022). The turbulence intensity measured at $1\,\text{m}$ from the outlet is reported to be $0.5\,\%$ by Lignarolo (2016) and the shear layer at $10\,\text{cm}$ from the outlet was measured using a pitot tube to extend $2.7\,\text{cm}$ (δ_{99}) towards the inside of the outlet and expands $2\,\text{cm}$ outwards, see Figure 3.1b.



(a) (b)

Figure 3.1: Schematic view of the Open Jet Facility at the Delft University of technology (Delft University of Technology, 2022) (a) and velocity profile in the shear layer at 10 cmfromtheoutlet(b)

3.2. Square cylinder test object

With the goal of obtaining nominal two-dimensional surface pressure distributions depending on the angle of attack similar to Figure 2.25, an experimental setup is constructed in the OJF test section consisting of a square cylinder test object and external support structures. In this section, information on the test object itself is provided.

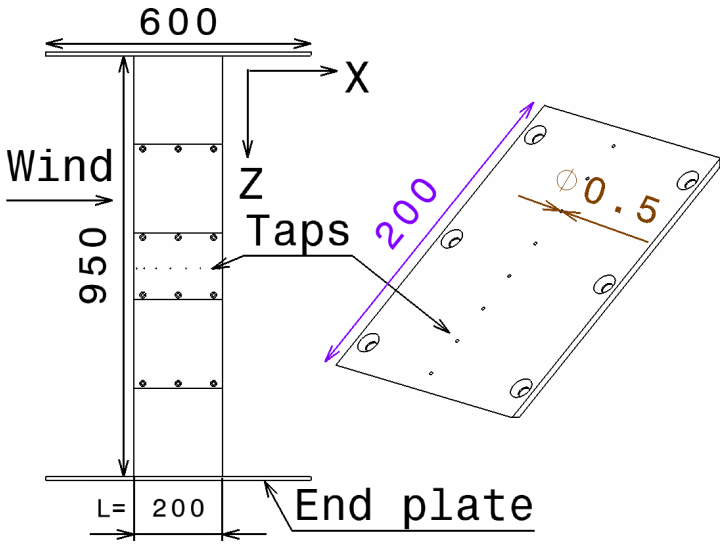


Figure 3.2: Side view drawing including main dimensions (in mm) of the square cylinder test object (left) and removable pressure tap plate (right).

A schematic side view of the model is presented in Figure 3.2 including its main dimensions and an enlarged view of the removable pressure pressure tap plates used. The model height H is 950 mm and side length L is 200 mm which results in an aspect ratio of 4.75. Since surface pressure distributions are measured using transducers (specifications discussed in Section 3.3.1), small holes are present around the perimeter of the object at its mid-height of 475 mm. Each of the holes is fitted with a 0.5 mm inner diameter steel pressure tap with a length of 12 mm. For the pressure tap pattern

around the model perimeter to be changed easily, the cylinder faces containing the taps are removable and referred to as pressure (tap) plates. An enlarged view of an isolated pressure tap plate is shown on the right of Figure 3.2. The pressure plates and square sectioned parts of the model are FDM 3D printed out of PLA plastic (Polylactic Acid). The inside of the model is kept hollow except for a steel rod spanning the entire 950 mm height which provides structural integrity and a mounting point for the pressure modules which house the pressure transducers. The aspect ratio of the model is relatively low which can cause unwanted three-dimensional effects near the top and bottom of the model to spoil nominal two-dimensional flow as achieved by the experiments referenced in Section 2.9. To counter this, circular end plates with 600 mm diameter are mounted on the top and bottom of the object which will work in combination with the external end plates to be discussed in Section 3.4. Figure 3.2 also shows a global reference frame which is defined such that the freestream velocity is aligned with the x-axis.

Four sets of pressure tap plates are used where each set consists of four plates. As a benchmark, plates containing 8 evenly spaced pressure taps are used to obtain pressure distributions with a spatial resolution of 32 in total. From the CFD results discussed in Section 4.7.2, it is found that decreasing the spacing of pressure taps towards the corners enhances the determination of pressure drag through surface pressure when using 8 taps per face. Therefore, a second set of plates using a cosine spacing on each cylinder face is manufactured and used as well. The coordinates of the pressure taps mapped onto the extent of a cylinder face from $-\frac{L}{2}$ to $\frac{L}{2}$ using this cosine spacing follows from:

$$cos\,(\chi) \cdot \frac{L}{2}, \quad \chi = i \cdot \Delta\chi, \quad i = 1, 2 \ldots n_{\mathsf{f}}, \quad \Delta\chi = \frac{180^{\circ}}{n_{\mathsf{f}} + 1} \tag{3.1}$$

With n_{f} the number of taps to be placed on a face of the square cylinder. Equation 3.1 is illustrated visually for $n_{\mathsf{f}} = 8$ below.

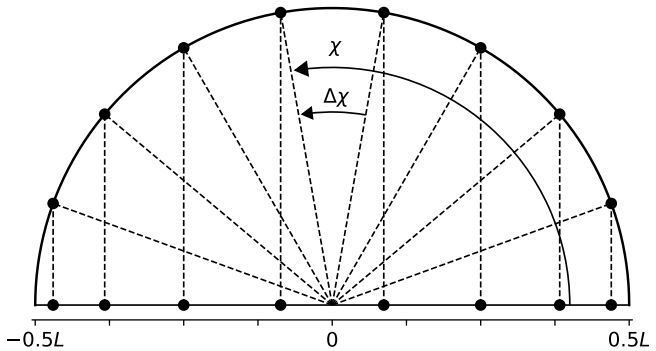


Figure 3.3: Cosine sampling for tap locations skewed towards square cylinder edges

The CFD results to be discussed in Section 4 have been used in combination with the QR-CP sensor placement algorithm to identify CFD based sensor locations for experimental use. More information regarding the determination of these sensor locations is provided in Section 4.7.2 however two sets of pressure tap plates referred to as Opt V1 and Opt V2 are manufactured to accommodate them. Each of the aforementioned

pressure tap distributions is shown schematically in Figure 3.4 as a section view at the model mid-height. Pressure taps are indicated as dots with their position determined either by local x' and y' coordinates or sensor angle ψ as shown. The face definitions and angle of attack are the same as used in Section 2.9 but are included regardless for completeness.

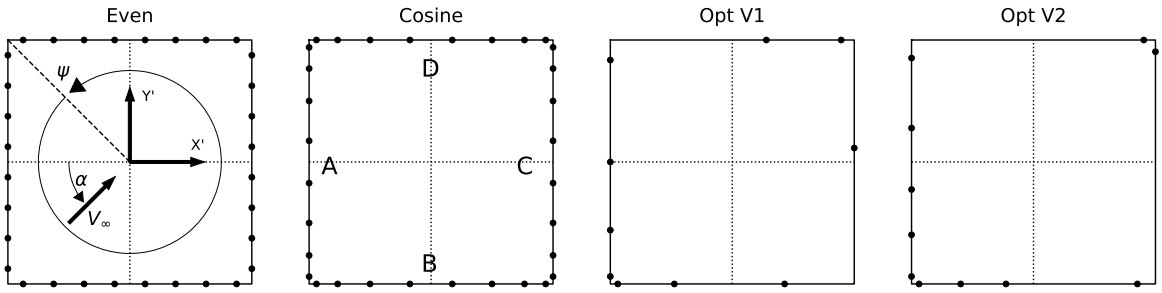


Figure 3.4: Tap locations for different plate sets, including local coordinate system and angle/face definitions.

3.3. Measurement equipment

With the test object itself designed around removable plates containing pressure taps, the main part of the data acquisition is performed by measuring surface pressure at the tap locations by means of pressure modules. Mainly serving as a validation tool of the measured pressure distributions, the complete test object is mounted on an external balance. The external balance is used to obtain force measurements independent of the number and locations of the pressure taps. Specifications of the pressure measurement system and external force balance are provided below.

3.3.1. Pressure modules

The pressure measurement system used is a Nub systems pressure scanning system consisting of two components; an acquisition unit and one or more pressure modules. A single pressure module houses 16 pressure transducers measuring differential pressure at the same operating range. The transducers in these modules are digital Honeywell-HSC series. To simultaneously measure the 32 pressure taps, two modules are used with a range of 600 Pa and 2500 Pa respectively. At these operating ranges, the transducers are stated to both have a total error band on the measurements of ± 1 % of the full scale span (Honeywell, TruStability board mount pressure sensors, HSC-series, 2022). The pressure modules inside the test object are connected to the data acquisition unit outside of the model by a cable running trough the bottom end plate. The data acquisition unit can support up to five pressure modules at once and is capable of sampling all individual transducers at a frequency of 2 kHz. The data acquisition unit connects to an external computer on which the modules can be controlled using LabVIEW software and data can be stored.

3.3.2. Force balance

The balance used is the OJF external balance delivered by the Dutch National Aerospace Laboratory NLR. This force balance is capable of measuring both forces and moments in three directions at a frequency of 2 kHz. The load ranges for the six measurable

components are tabulated in Table. 3.1.

Table 3.1: OJF external balance load- and moment- ranges

	Simultaneous loading	**Single loading**
Axial Force (F_x)	± 250 N	± 250 N
Side Force (F_y)	± 500 N	± 600 N
Vertical Force (F_z)	± 500 N	± 3500 N
Rolling Moment (M_x)	± 500 N m	± 550 N m
Pitching Moment (M_y)	± 250 N m	± 500 N m
Yawing Moment (M_z)	± 50 N m	± 125 N m

The balance is operated while simultaneously loaded, therefore the single loading ranges are not of interest. The coordinate system for the balance is similar to the global coordinate system shown in Figures. 3.2 and 3.5 however is offset with respect to the x- and y-axes by $7.25°$, indicated by dashed axes with subscript 'b' in Figure 3.5, and rotates together with the test object. The balance x- and y-axes therefore are offset to the global x- and y-axes by $\beta = 7.25° + \alpha$. The accuracy of the balance has been determined by a first order dead-weight calibration which yielded the maximum errors and standard deviations in percentage of the applied maximum load shown in Table. 3.2.

Table 3.2: OJF external balance accuracy in percentage of applied maximum calibration load

	ΔF_x	ΔF_y	ΔF_z	ΔM_x	ΔM_y	ΔM_z
Maximum error [%]	0.06	0.23	0.16	0.05	0.05	0.25
Standard deviation [%]	0.02	0.05	0.05	0.01	0.01	0.07

3.4. Experimental setup

The test object, measurement equipment and several external support objects together are used to realize an experimental setup at the OJF. A schematic view of the setup is shown in Figure 3.5 of which the main components will be discussed.

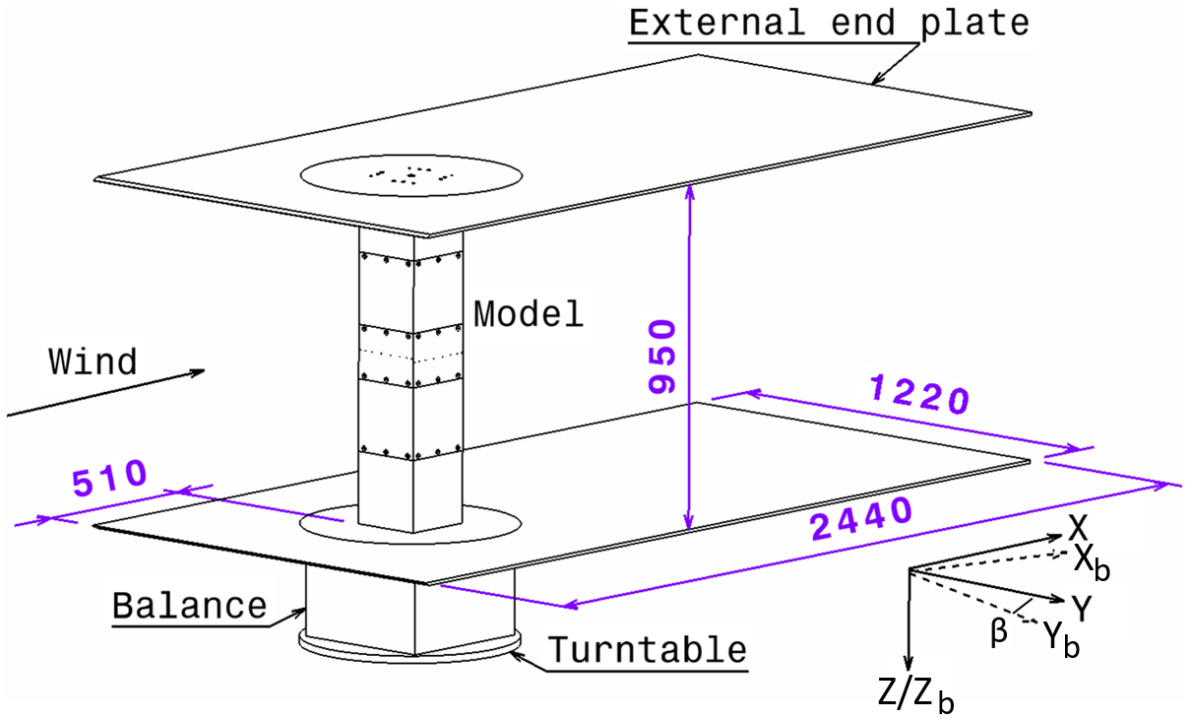


Figure 3.5: Schematic of the experimental setup including main components and dimensions (in mm)

Inside the test section of the OJF, a height-adjustable hydraulic table is used to assemble the complete experimental setup onto. The hydraulic table contains threaded holes onto which a turntable is fastened. The turntable, sketched at the bottom of Figure 3.5, is capable of rotating $360°$ while the OJF is being operated and effectively enables the angle of attack of anything mounted onto it to be varied. The turntable contains attachment points which are used to mount the OJF external balance. This therefore rotates in combination with the turntable resulting in the aforementioned balance reference frame being offset with respect to the global reference frame which is indicated by the dashed axes and angle β in Figure 3.5.

The bottom end plate of the test object is fastened to the balance using threaded holes present on a balance adapter plate. The test object is assembled on the bottom end plate part by part from the bottom up to keep access to the inside of the test object during assembly. The steel rod in the center of the model is used as an attachment point for the pressure modules and allows for the removal of the pressure plates without the model collapsing. The cables attached to the two pressure modules run through a hole in the bottom end plate inside the model and are attached to the data acquisition unit positioned next to the balance. Two pressure modules are used which allow for simultaneous surface pressure measurements across at most 32 taps (16 taps per module). Connections between the modules and the taps are made using pieces of urethane tubing with inner diameter $1\,$mm, outer diameter $2\,$mm and length $150\,$mm. The taps, tubing and pressure module ports are numbered from 1 to N based on their sensor angle ψ with 1 being closest to $\psi = 0°$, N being closest to $\psi = 360°$ and N being the number of pressure taps on the four plates combined.

With an aspect ratio of 4.75, the circular end plates are not expected to provide enough of a boundary to achieve nominal two-dimensional flow (Zhao et al., 2021). Therefore larger external end plates are used around the circular end plates as well. These external end plates are of size $2440\,\text{mm} \times 1220\,\text{mm}$ and contain a circular cutout roughly $10\,\text{mm}$ larger in radius compared to the model end plates to account for oscillations and bending of the test object. Using aluminum extrusions as support (not shown in Figure 3.5 for clarity), these plates are positioned around the model end plates such that the model can rotate freely without coming into contact with any external object. The leading edges of the external end plates are chamfered at a $60°$ angle to reduce their downstream impact on the model. The complete setup is positioned such that the lower external end plate is located $20\,\text{mm}$ above the bottom of the OJF outlet and the leading edge protrudes into the outlet by $100\,\text{mm}$.

3.5. Data acquisition procedures

The wind tunnel measurements are performed following a test matrix made in advance. In this section, a condensed version of this test matrix is briefly discussed. After this, some practical aspects of the force balance and surface pressure measurements are discussed including force/angle zeroing and a blockage correction.

3.5.1. Test matrix

Two 'categories' of measurements are performed during the wind tunnel campaign: First, balance measurements to be used in the determination of the circular end plates drag; Second, simultaneous force balance and surface pressure measurements at various angles of attack. These measurements form the main part of the wind tunnel campaign which yield the data to be used in combination with the POD framework. These wind tunnel measurements are summarized in a condensed version of the test matrix in Table. 3.3.

Table 3.3: Condensed test matrix

	V_∞ [m/s]	α [°]	Tap locations
Wind off end plates Wind on end plates	0 10	N/A	N/A
Wind off runs Wind on runs	0 10	$[0°, 2.5° \dots 45°]$	Even, Cosine, Opt V1, Opt V2

The 'wind off-' and 'on end plates' measurements are used to find an approximation for the isolated drag of the circular end plates. Despite their purpose of counteracting three-dimensional effects, the plates experience drag measured by the balance which is not representative of a nominal two-dimensional flow. For comparable results between balance and pressure taps, the end plate drag is therefore removed from the balance measurements. To isolate the force experienced by the end plates, sequential measurements of the bottom end plate, the bottom end plate + steel rod and the bottom end plate + steel rod + top end plate are taken. These measurements are performed once at a freestream velocity of $10\,\text{m}\,\text{s}^{-1}$ (wind on) and once at wind off conditions to act as zeroing measurements.

The 'wind off' and 'wind on runs' in Table. 3.3 include balance and surface pressure measurements for angles of attack from $\alpha = 0°$ to $45°$ in $2.5°$ intervals. Such a sequence of measurements will be referred to as an α-sweep and are performed for each of the four sets of pressure plates. This means four α-sweeps are measured at wind off conditions and four at wind on conditions however all measurements within this 'category' are performed twice for redundancy meaning a total of 16 α-sweeps are performed. Unlike what is shown in Table. 3.3, the angle of attack order for measurements is randomized to decouple the effect of time dependent variables.

3.5.2. Angle of attack zeroing

The measured forces and surface pressure around a square cylinder are sensitive to the true angle of attack experienced. Due to geometric constraints, the local coordinate system of the square cylinder could not be kept aligned with the turntable through which control over the angle of attack is possible. The resulting offset in the angle of attack is known in theory from the CAD model of the experimental setup however differs in practice as a result of many uncertainties ranging from manufacturing (in)accuracy to hydraulic table alignment. For the offset between the controllable turntable angle and the true square cylinder angle of attack to be known, the model was equipped with the evenly spaced pressure taps which are symmetric w.r.t. the freestream velocity at $\alpha = 0°$. At a freestream velocity of $10\,\mathrm{m\,s^{-1}}$, the turntable angle was varied until a symmetric time averaged pressure distribution was measured around the model. This identified the offset between the turntable and square cylinder angle of attack to be $7.25°$, hence the earlier shown $\beta = 7.25° + \alpha$ with α the true angle of attack experienced by the square cylinder and β the angle of the turntable with respect to the global coordinate system as shown in Figure 3.5. This offset was validated to result in a symmetric pressure distribution at $\alpha = 45°$ as well.

3.5.3. Force balance and surface pressure measurements

In Table. 3.3 it is shown how every measurement point (combination of angle of attack and used set of pressure plates) at wind on conditions is preceded by a wind of measurement point. These wind off measurements are used for zeroing of balance results only and therefore consist of time averaged balance measurements only over a period of $10\,\mathrm{s}$ measured at $2\,\mathrm{kHz}$. Any measurement point at wind on conditions will have corresponding zeroing measurement subtracted such that aerodynamic forces are isolated from possible offsets in the balance readings resulting from a shifting model center of gravity with angle of attack for example.

Besides removing the zero measurements from balance measurements obtained at wind on conditions, the drag coefficients of the isolated circular end plates is removed as well as they are not representative of the drag experienced by a nominally two-dimensional square cylinder. The dedicated measurement points for the determination of this end plate drag are briefly mentioned in Section 3.5.1 but require three wind on measurements of the bottom end plate only, the bottom end plate with the vertical rod and the bottom end plate, vertical rod and top end plate together. After removal of corresponding zero measurements, the following drag coefficients are measured

using the cylinder model area $A = 0.2\,\text{m} \times 0.95\,\text{m}$ as a reference.

Table 3.4: Measured drag coefficients used for determination of the end plates drag contribution

	Bottom plate	Bottom plate + rod	Bottom plate + rod + top plate	**End plates**
C_D	0.063	0.160	0.210	0.113

Assuming no effect from the interaction between rod and plates, the drag of the rod is obtained by subtracting the drag coefficient of the bottom plate from the drag coefficient of the bottom plate + rod. The drag contributions of the isolated end plates is then obtained by subtracting the drag coefficient of the rod from the drag of the bottom plate + rod + top plate to obtain the C_D stated in the rightmost column of Table. 3.4. This drag coefficient will be subtracted from all wind on balance measurements.

While considering the aforementioned corrections, the actual wind on measurement points are performed at a freestream velocity of $10\,\text{m}\,\text{s}^{-1}$. This results in a Reynolds number based on the model side length L of approximately $Re = 135,000$. Initially, measurements were performed at $20\,\text{m}\,\text{s}^{-1}$ however due to excessive oscillations of the model, the freestream velocity was reduced. For each combination of angle of attack and pressure tap plates, simultaneous force balance and surface pressure measurements are performed for a duration of $30\,\text{s}$. The force balance operates at its maximum acquisition frequency of $2\,\text{kHz}$ and the surface pressure is measured at the same frequency for all individual pressure taps. Reported square cylinder Strouhal numbers vary in the range of 0.11 to 0.18 depending on the angle of attack (Lee (1975), Zhao et al. (2021), Chen and Liu (1999)). Given the current model's sidelength of $L = 0.2\,\text{m}$ and the used freestream velocity of $V_\infty = 10\,\text{m}\,\text{s}^{-1}$, this corresponds to a maximum vortex shedding frequency of $f_s = 9\,\text{Hz}$. To capture the vortex shedding in the pressure measurements while avoiding resonance in the tubing affecting measurements by more than $5\,\%$, the tube length should comply with the following equation from organ resonance theory (Irwin et al., 1979):

$$\left|\frac{p_{\text{measured}}}{p_{\text{actual}}}\right| = \left|\left(\cos\left(\frac{2\pi f l}{c}\right)\right)^{-1}\right| < 1.05 \quad \text{with} \quad f = 2f_s = 18\,\text{Hz},\ c = 343\,\text{m}\,\text{s}^{-1} \tag{3.2}$$

Where c is the speed of sound and f is the frequency of interest being taken twice the vortex shedding frequency to comply with the Nyquist criterion. Following from 3.2, the tube length l should not exceed $0.94\,\text{m}$ and thus the tube length of $150\,\text{mm}$ suffices for the unsteady surface pressure measurements. Besides surface pressure and balance measurements for each measurement point, ambient pressure, density, freestream velocity and temperature are recorded as well to be used for surface pressure and force nondimensionalization.

3.5.4. Blockage velocity correction

At an angle of attack of $0°$, the frontal area of the square cylinder model equals $0.2\,\text{m} \times 0.95\,\text{m}$. With an OJF outlet of size $2.85\,\text{m} \times 2.85\,\text{m}$, this model area results in a blockage ratio of approximately 2%. At this small blockage ratio, no significant effect of an

increased effective velocity was expected however pressure coefficients obtained on the model normalized with dynamic pressure using a freestream velocity of $10\,\mathrm{m\,s^{-1}}$ resulted in values above one. At $\alpha = 0°$ when using plates with a pressure tap at $\psi = 45°$ (Opt V1 plates, see Figure 3.4), the tap effectively measures the stagnation pressure and should therefore read a pressure coefficient of approximately one. The excess in the pressure coefficient measured at this tap was converted into a velocity correction and added to the measured freestream velocity. The value of this velocity correction was found to be $0.86\,\mathrm{m/s}$ which is 8.6% of the used freestream velocity. This velocity correction, which is much larger than one might expect based on the model blockage with respect to the OJF outlet, is likely a result of the external end plates effectively decreasing the outlet area. Using the distance between the external end plates as the outlet height instead, results in a blockage ratio of approximately 7% which falls in line with the magnitude of the required correction.

4

Square cylinder CFD

Before the experimental measurement campaign, CFD simulations of flow around a two-dimensional square cylinder were performed to test the GPOD implementation. The numerical setup is discussed first in Section 4.1. Resulting flow fields at various angles of attack are shown in Section 4.2 and discussed briefly. A more detailed discussion on the numerical results follow from the time-averaged surface pressure distributions in Section 4.3 were also comparisons to literature are made. Numerical pressure distributions could eventually be used as training snapshots for sparse experimental measurements in combination with GPOD but first it is tested how well purely numerical reconstructions based on dominant POD modes and sparse sensors perform. Section 4.4 illustrates the dominant POD modes while the assessment of Gappy POD reconstruction performance is split up into a section for surface pressure reconstruction accuracy and drag estimation accuracy in Section 4.4 and Section 4.5 respectively. The CFD results have also been used in an attempt to enhance the quality of the experimental measurements through a sensor placement strategies, these are discussed in Section 4.7.

4.1. Numerical setup

A two-dimensional mesh around a square cylinder model is generated using the Ansys ICEM CFD software. A schematic of the CFD setup is shown in Figure 4.1 together with parameters used for the simulations in Table. 4.1. The cylinder has a length and height of $L \times H = 0.2\,\mathrm{m} \times 0.2\,\mathrm{m}$ and is centered in a domain of size $30L \times 30L$. Each of the four sides of the cylinder is discretized using 136 cells resulting in $4 \times 136 = 544$ points at which variables are computed on the cylinder surface. The total number of cells in the mesh is 124,032. Simulations are performed at a freestream velocity of $20\,\mathrm{m\,s^{-1}}$ yielding a cylinder side length based Reynold's number of roughly $Re = 300,000$.

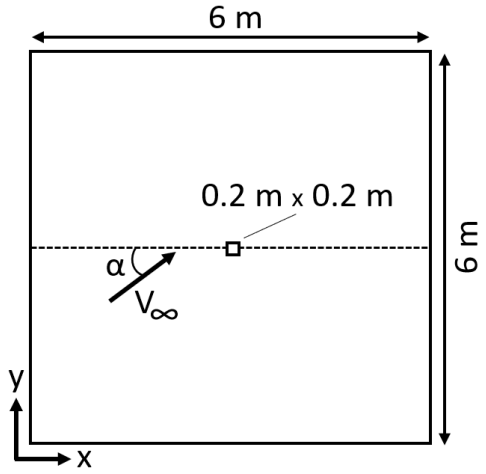


Figure 4.1: Schematic CFD setup

Table 4.1: CFD setup parameters

Domain size	$6\,\text{m} \times 6\,\text{m}$
Cylinder size	$0.2\,\text{m} \times 0.2\,\text{m}$
No. cells	124,032
No. cells cylinder	4×136
y^+	<1
Re	300,000
α	$0°$ to $45°$, $\Delta\alpha = 2.5°$
Type	Transient, incompressible
Simulation time	$6\,\text{s}$
Δt	$1\,\text{ms}$
Turb. model	$k-\omega$ SST

The Ansys CFX solver is setup to run unsteady RANS simulations for a total time of $6\,\text{s}$ using a timestep Δt of at most $1\,\text{ms}$ (timesteps to initialize being smaller). Without resorting to LES-type approaches, the $k-\omega$ SST turbulence model (Menter, 1994) is expected to be best suited for this application. A total of 19 simulations are run using an angle of attack from $0°$ up to and including $45°$ at an angle of attack 'resolution' $\Delta\alpha$ of $2.5°$. Surface pressures over the cylinder are saved for each time step in the time interval from $1.5\,\text{s}$ to $6\,\text{s}$ to exclude the startup period of the vortex shedding. The average over these saved time steps are used as time-averaged surface pressure distributions for each angle of attack.

4.2. Example flowfields

Four flowfields as obtained from the CFD simulations are visualized in Figure 4.2. The flowfields show streamlines around the cylinder to visualize the flow topology and are displayed on a background of the time averaged velocity magnitudes.

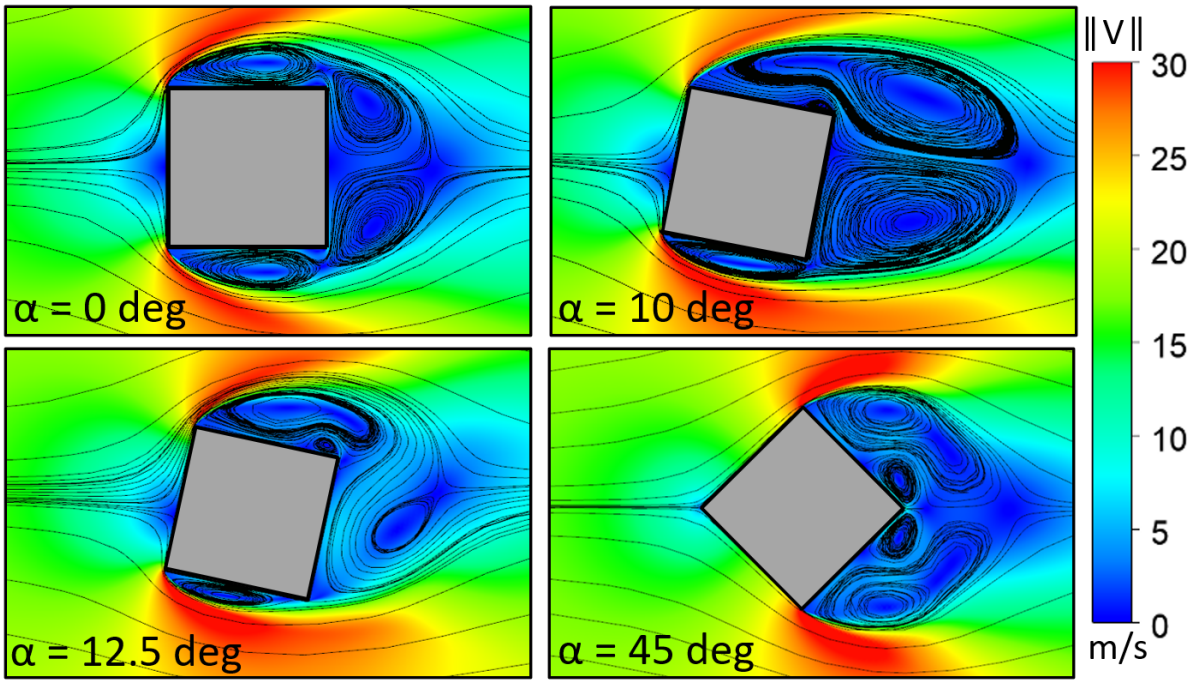


Figure 4.2: uRANS time averaged velocity fields including streamlines at angles of attack in both the sub- and supercritical flow regime.

The flowfield at an angle of attack of $0°$ shows a flow topology consisting of four clearly identifiable recirculation regions. Flow separates from the upstream corners and forms a shallow recirculation region on both face B and face D. The wake consists of two large recirculation regions which are relatively close to the the surface of the body. At this angle of attack, the flowfield is symmetric as expected and agrees very well with the earlier shown flow patterns measured experimentally by van Oudheusden et al. (2008) in Figure 2.23. This is also the case for the uRANS flowfield visualization at $\alpha = 10°$. The rotation of the cylinder has caused the recirculation region on face D to merge with the wake and the detached flow from the bottom left corner gets closer to reattachment as the shear layer approaches the bottom right corner. In accordance with Figure 2.23, the length of the wake grows considerably. At $\alpha = 12.5°$, the shear layer has steadily reattached on face B and the recirculation region is therefore pinched off. The recirculation regions in the wake seem to have broken up into a very asymmetric state which deviates from the double recirculation flow topology still visible in experimental results. Also at an angle of attack of $\alpha = 45°$ does the uRANS flowfield deviate from experimental results such as shown in Figure 2.24. The behaviour of the flow around the square cylinder surface is discussed in more detail in combination with the time averaged surface pressure distributions. These results are shown in the following section and will further point out the clear discrepancies between CFD and experimental results which seem to have appeared from the shear layer reattachment at the critical angle of attack onwards.

4.3. Surface pressure results

The obtained time averaged surface pressure distributions as a function of the angle of attack are shown in Figure 4.3. For clarity these figures include only the pressure

distributions at a $5°$ angle of attack interval. The pressure coefficients C_p are obtained by normalizing the surface pressure by the dynamic pressure $q = \frac{1}{2}\rho V_\infty^2$.

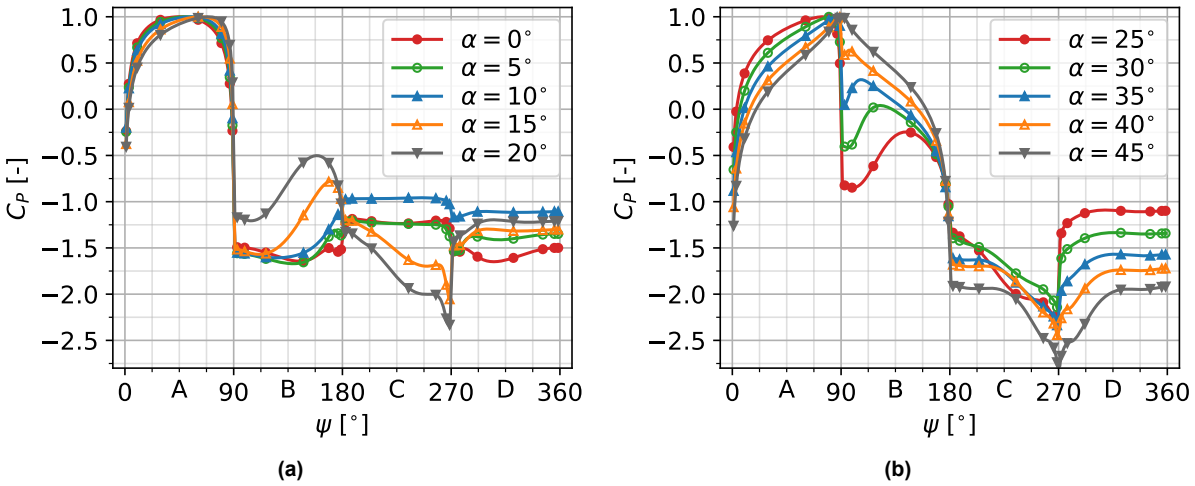


Figure 4.3: uRANS time averaged pressure distributions for $\alpha = 0°$ to $20°$ (a) and $\alpha = 25°$ to $45°$ (b)

The results indicate a stagnation point on face A of the cylinder for $\alpha < 45°$. At $\alpha = 0°$, the pressure distribution is symmetric and achieves $C_p = 1$ around $\psi = 45°$. For increasing angle of attack, the stagnation point moves in the direction of the corner at $\psi = 90°$ and skews the distribution at face A until eventually at $\alpha = 45°$, the distribution on face A is symmetric with respect to face B. Following the flow regime classifications described in Section 2.9.1, Figure 4.3a shows a transition from the subcritical to supercritical regime between $\alpha = 10°$ and $15°$. Where for $\alpha < 15°$ pressure on face B is low over the majority of the face and rises towards face C, $\alpha = 15°$ shows a clear peak in pressure on face B corresponding to a reattached shear layer near $\psi = 180°$. Despite not being shown in the above figures, the reattachment of the shear layer indicated by this peak in pressure is visible already at $\alpha = 12.5°$ meaning that $10° \leq \alpha_{crit} \leq 12.5°$. Below α_{crit}, the pressure on face C should increase for increasing angle of attack as the vortex formation region in the wake moves downstream according to Lee (1975). This does not seem to be the case for the CFD results at $\alpha = 0°$ and $5°$ however the base pressure increases considerably from $\alpha = 5°$ towards a maximum at $\alpha = 10°$. Since the reattachment of the shear layer at α_{crit} coincides with a maximum in the base pressure, the critical angle of attack using this CFD setup is thus likely closer to $\alpha = 10°$.

The CFD pressure distributions within the subcritical regime seem to follow the characteristics described in Section 2.9. In the supercritical regime, CFD results on face B show the growth of the reattached region for increasing angle of attack as expected. Faces C and D however experience a rapid drop in pressure for increasing angle of attack, especially near $\psi = 270°$ which is not described in any of the experimental results referenced in Section 2.9. The transition from sub- to supercritical regime therefore seems to mark the boundary from which the RANS results start to deviate significantly from literature. This is further illustrated by the comparisons between CFD and results from literature shown in Figure 4.4.

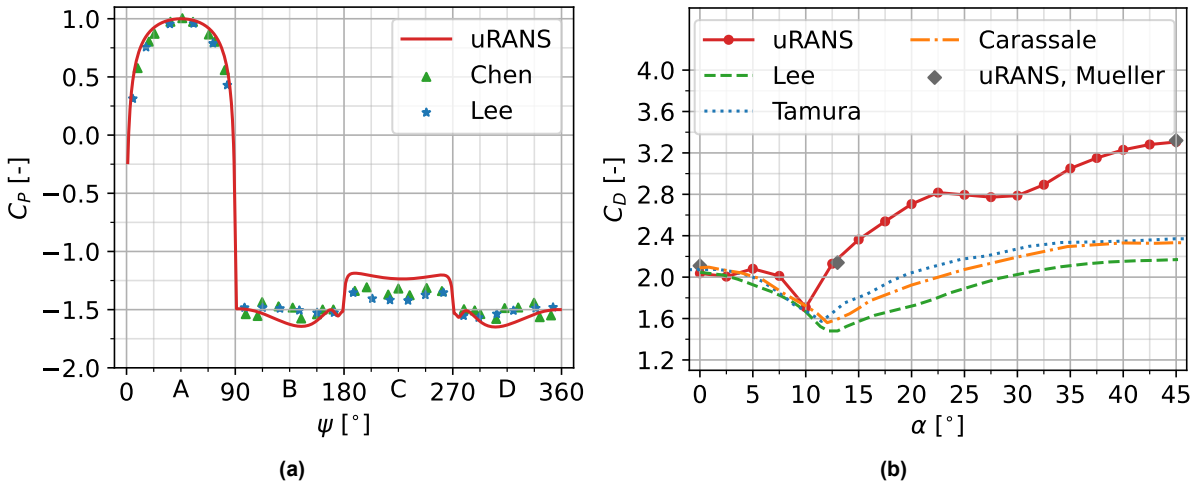


Figure 4.4: uRANS pressure distribution comparison to experimental results at $\alpha = 0°$ (a) and uRANS drag curve comparison to experimental results (b)

Figure4.4a, compares the CFD pressure distribution results at $\alpha = 0°$ to experimental results found by Chen and Liu (1999) and Lee (1975). Even though differences are visible, the general shape of the time-averaged uRANS pressure distribution agrees well with the experimental results which holds true for all angles of attack in the sub-critical regime. A decreasing pressure on faces C and D for increasing angle of attack above α_{crit} is reported throughout literature (see Lee (1975) for example) however the effect seems to be greatly overpredicted in the CFD results judging from the overestimation of the drag coefficients in Figure 4.4b from $\alpha = 12.5°$ onwards.

A similar overprediction of the drag by uRANS was experienced by Mueller (2012). Mueller compared the pressure distributions obtained through uRANS simulations to those from LES simulations and experiments. The experimental and time-averaged uRANS pressure distributions at $\alpha = 45°$ are shown in Figure 4.5a together with the uRANS results at the same angle of attack from Figure 4.3b.

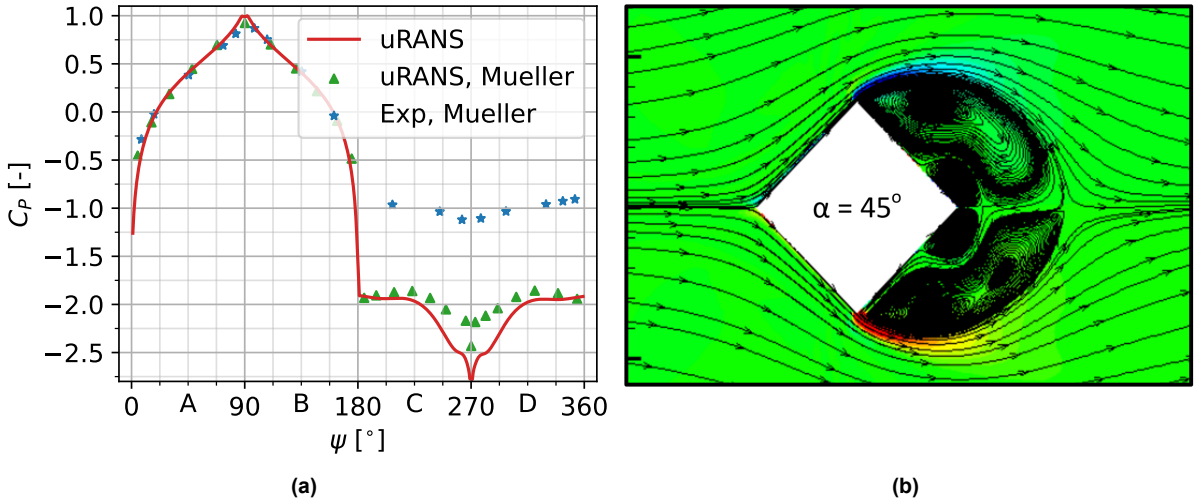


Figure 4.5: uRANS pressure distribution comparison at $\alpha = 45°$ to experimental and uRANS results from Mueller (2012) (left) and flowfield visualization at $\alpha = 45°$ of the uRANS results from Mueller (2012) (right)

The numerical pressure distributions in Figure 4.5a show much similarity and underpredict the experimental pressure distribution by more than 1 in terms of C_p on faces C and D. The drag coefficient corresponding to the uRANS pressure distribution of Mueller at this angle of attack is included in Figure 4.4b as well as those for $\alpha = 13°$ and $0°$ which all match the current report's numerical results. By comparing experimental flow visualizations to the numerical results, Mueller found that uRANS incorrectly produces a set of secondary vortices near the corner at $\psi = 270°$, not present in experiments, in addition to the large recirculation regions in the wake. These vortices decrease pressure with an overprediction of the drag as a result. A visualization of the flowfield at $\alpha = 45°$ as obtained by Mueller is shown in Figure 4.5b. The secondary vortices are clearly visible and the overall flowfield resembles that shown in Figure 4.2 very well. This leads to the conclusion that the large differences between literature and the current report's numerical results are caused by poor performance of the RANS approach in the supercritical regime and instead an LES-type approach would be necessary to obtain accurate numerical results.

Despite these disappointing results from the numerical simulations in especially a quantitative sense, the general expected trends in the behaviour of the pressure distributions can still be recognized. Since POD is a purely data-driven approach, there is no inherent reliance on the quantitative accuracy of the simulations. The numerical pressure distributions can therefore still serve as a more complex case for GPOD reconstructions and provide sensor locations suited for experiments as well based on the dominant POD modes identified.

4.4. Surface pressure POD analysis

Since the square cylinder is symmetric about the diagonal through $\psi = 90°$, the pressure distributions from $\alpha = 0°$ to $42.5°$ can be mirrored to obtain 37 pressure distributions in total covering $\alpha = 0°$ to $90°$ at a $2.5°$ interval. The 19 pressure distributions corresponding to integer multiples of $\alpha = 5°$ (including $\alpha = 0°$) have their mean across the angles of attack subtracted and are stored in a matrix for POD to be applied. This matrix thus consists of $m = 19$ rows and $n = 544$ columns and is used as training data to characterize the square cylinder flow pressure distributions from the CFD results. The remaining pressure distributions not included in this training data will be used for GPOD reconstructions in Section 4.5.

4.4.1. Eigenvalue spectrum

The POD procedure applied to the training matrix yields the orthogonal eigenmodes in columns of $\mathbf{\Phi} \in \mathbb{R}^{n \times m}$ and corresponding coefficients in $\boldsymbol{A} \in \mathbb{R}^{m \times m}$. Since the parameter varied throughout the training data is the angle of attack, the coefficients are a function of α and will therefore be referred to as α-coefficients. Eigenmodes are ordered following the eigenvalue magnitudes which indicate the relative variance captured per mode. The eigenvalue spectrum of the CFD pressure distributions is shown in Figure 4.6.

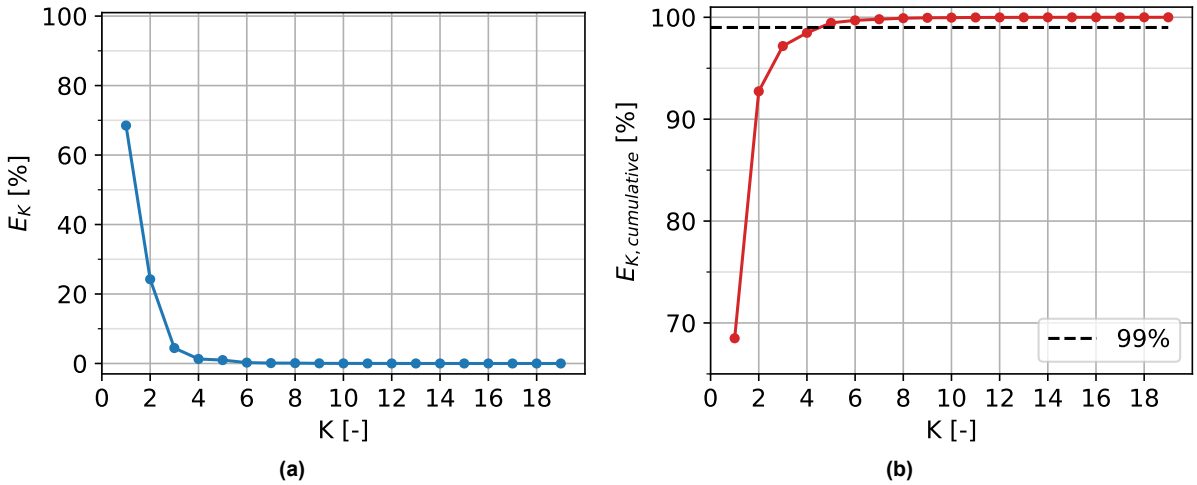


Figure 4.6: Individual (a) and cumulative (b) energy capture of the eigenvalues corresponding to CFD pressure distributions

Figure 4.6a shows the individual normalized magnitudes of the eigenvalues and Figure 4.6b the cumulative normalized magnitude. The mathematical definitions of E_K and $E_{K,cumulative}$ are stated in Equation 2.19 and are often referred to in terms of energy capture. The individual energy capture of the first eigenvalue is 68 % which is by far the largest compared to an energy capture of 24 % and 4 % of the second and third eigenvalue respectively. This dominance of the first few eigenvalues results in 5 out of the 19 eigenvalues capturing already over 99 % of the total energy. Such rapid decay in the eigenvalues is desired from a reduced order model point of view but does not necessarily result in accurate reconstructions. For this, the eigenmodes from training should be representative for the pressure distributions to be reconstructed as well and the loss due to sparse sampling should be kept to a minimum using tailored sensor locations.

4.4.2. Dominant eigenmodes

With the first few eigenvalues dominating in terms of energy capture, the corresponding eigenmodes capture prominant features present throughout the training data. In Figure 4.7 the first two eigenmodes and corresponding α-coefficients are shown which can be used to illustrate how many of the features visible in the pressure distributions from Figure 4.3 are encoded in a low-rank representation.

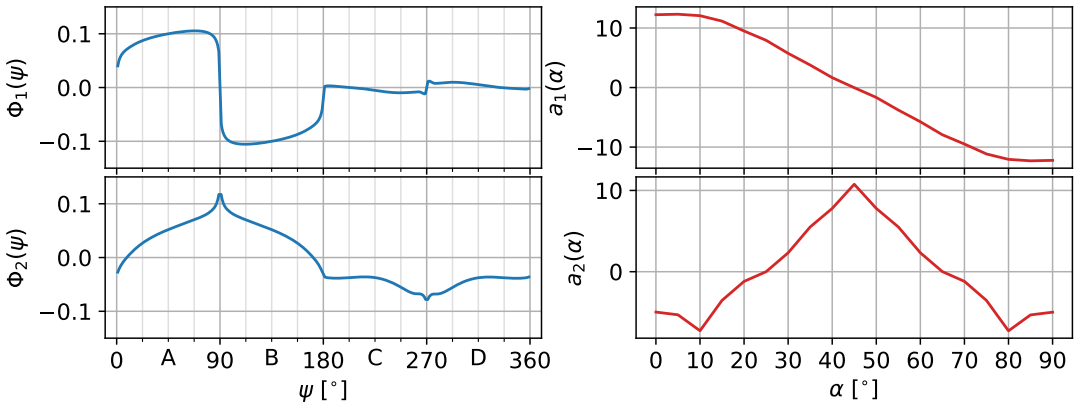


Figure 4.7: First and second spatial eigenmodes (left), first and second α-coefficients (right) for CFD square cylinder case

The eigenmodes have the same dimension as the provided training pressure distributions and can thus be visualized in the same format. The resulting figures contain pressure distribution-like shapes but contain information regarding the spatial (anti)correlation, hence: spatial eigenmodes. The first eigenmode resembles the pressure distribution at $\alpha = 0°$ or $90°$ and shows how the pressure on face A tends to be anti-correlated with that on face B. Together with the first set of $\alpha-$coefficients as $\boldsymbol{a}_1\boldsymbol{\Phi}_1^T$, this eigenmode describes the gradual shift in the high pressure region from face A at $\alpha = 0°$ to face B at $\alpha = 90°$. At $\alpha = 45°$, the pressure distribution does not contain features present in the first eigenmode and as a result, the α-coefficient is zero. The α-coefficients thus activate certain modes more or less depending on the angle of attack.

The second eigenmode instead does resemble the pressure distribution around $\alpha = 45°$ and can be seen to be most active in this angle of attack range from the α-coefficients. The two modes together as a rank two approximation therefore capture the general trend throughout the angles of attack. Close to $\alpha = 0°$, the first eigenmode is most active but the second eigenmode does not have a zero contribution. For increasing angle of attack, the presence of the first eigenmode drops as the pressure distribution deviates from resembling that of $\alpha = 0°$ (or $90°$). Around $\alpha = 45°$, the second eigenmode is most active in the rank two approximation. The pressure distributions at the remaining angles of attack are characterized by the growing/shrinking region of reattached flow on face B which both shown eigenmodes would not be able to reproduce and for which additional modes would be necessary. A small number of modes can thus be used to identify underlying trends in data however increasingly fine details require higher order modes.

4.5. Numerical surface pressure reconstruction using GPOD

The complete set of eigenmodes forms a library of square cylinder flow which can be used to infer full state approximations of the pressure distributions from sparse measurements using the GPOD approach described in Section 2.5. In this section, examples of reconstructed pressure distributions will be shown and the accuracy of

the reconstructions will be quantified in terms of the Root Mean Squared Error for different sensor budgets.

4.5.1. Reconstruction examples

The eigenmodes were obtained from 19 training pressure distributions at angles of attack being multiples of $5°$. The reconstructions will be performed on pressure distributions which have not been included in the training data as to see if the mode shapes identified by the POD represent square cylinder flow in general. Pressure distributions to be reconstructed are made artificially gappy by discarding all data except for a specific few points which therefore effectively serve as sensor locations. These sensor locations are determined through application of the QR-CP algorithm as discussed in Section 2.7. Reconstructions as a result of using 10 sensors and 10 eigenmodes at $\alpha = 12.5°$ and $42.5°$ are shown below in Figures. 4.8 and 4.9. Note that the number of sensors n_s is kept equal to the number of modes K used throughout this report for GPOD reconstructions.

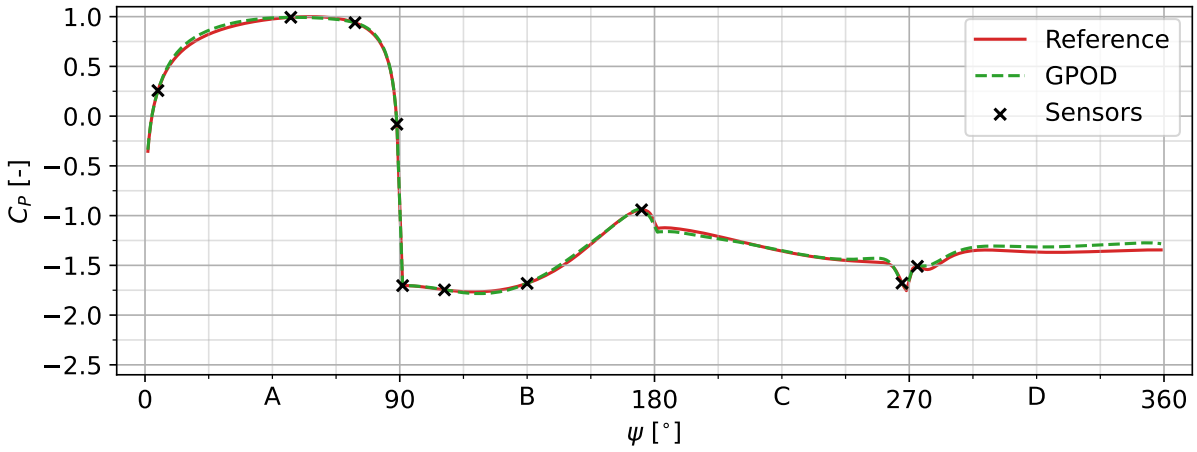


Figure 4.8: CFD GPOD reconstruction at $\alpha = 12.5°$, $n_s = K = 10$

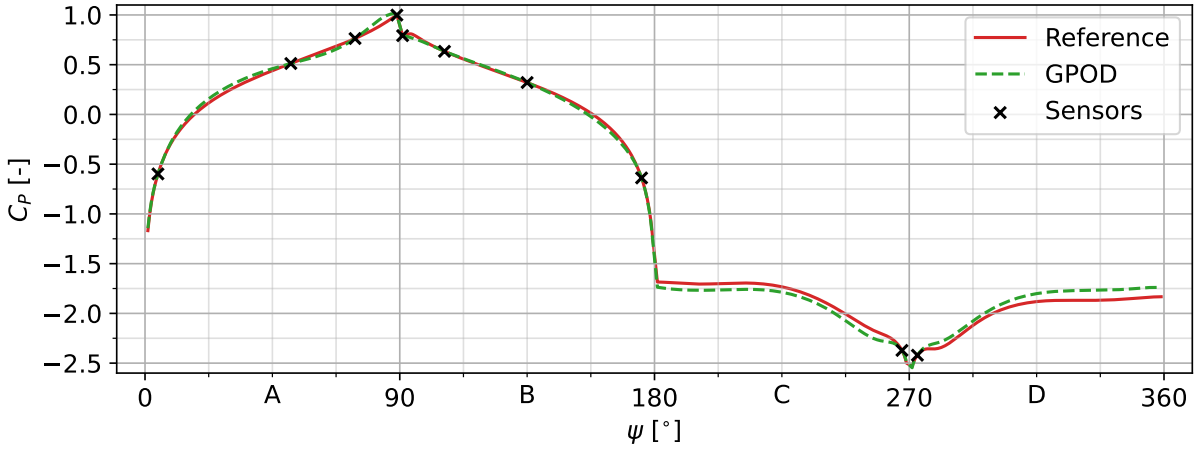


Figure 4.9: CFD GPOD reconstruction at $\alpha = 42.5°$, $n_s = K = 10$

At $\alpha = 12.5°$, the shear layer has just reattached at face B and is thus close to the critical angle of attack with relatively low drag. At $\alpha = 42.5°$, drag is almost at its

maximum instead and the two shown distributions thus provide insight in how the reconstructions perform under different conditions. Using 10 sensors, the percentage of missing pressure values from the total (gappiness) is $98.16\,\%$. The figures include the non-sparse pressure distribution as a reference for the reconstruction to be compared with. Clearly, the information contained in the 10 modes used for the reconstructions extends outside of the training pressure distributions as the reconstructions follow their respective reference closely. For both angles of attack, faces A and B seem to be reconstructed more accurate compared to faces C and D. This is better noticeable at $\alpha = 42.5°$ and is a result of the sensor placement. Since most variance in the pressure distribution occurs at faces A and B, these faces generally show high modal activity in the eigenmodes which is favored in the sensor placement through the QR algorithm. This results in faces A and B containing each 4 out of 10 sensors and faces C and D 1 each. With less sensors acting as constraints, the reconstruction accuracy suffers slightly at faces C and D. The uneven sensor distribution identified by the QR algorithm makes physical sense however; the sensor locations are symmetric about the cylinder diagonal and focus on capturing the reattachment of the shear layer through the positioning over faces A and B. The more predictable behaviour of the pressure on faces C and D would not require as much information for reconstructions but varies most extremely near $\psi = 270°$, exactly where two sensors are placed. The difference between the reconstruction and reference in terms of RMSE is $2.6 \cdot 10^{-2}$ and $3.5 \cdot 10^{-2}$ for $\alpha = 12.5°$ and $42.5°$ respectively.

4.5.2. Pressure reconstruction accuracy

The previous section showed the GPOD reconstruction at only two angles of attack however reconstructions are generated on all angles of attack in the $\alpha = 0°$ to $90°$ range and also for a sensor budget of 1 to 19. For each of these reconstructions the accuracy is measured in terms of the RMSE with respect to the corresponding reference pressure distribution. A colormap visualizing these RMSE values are shown in Figure 4.10.

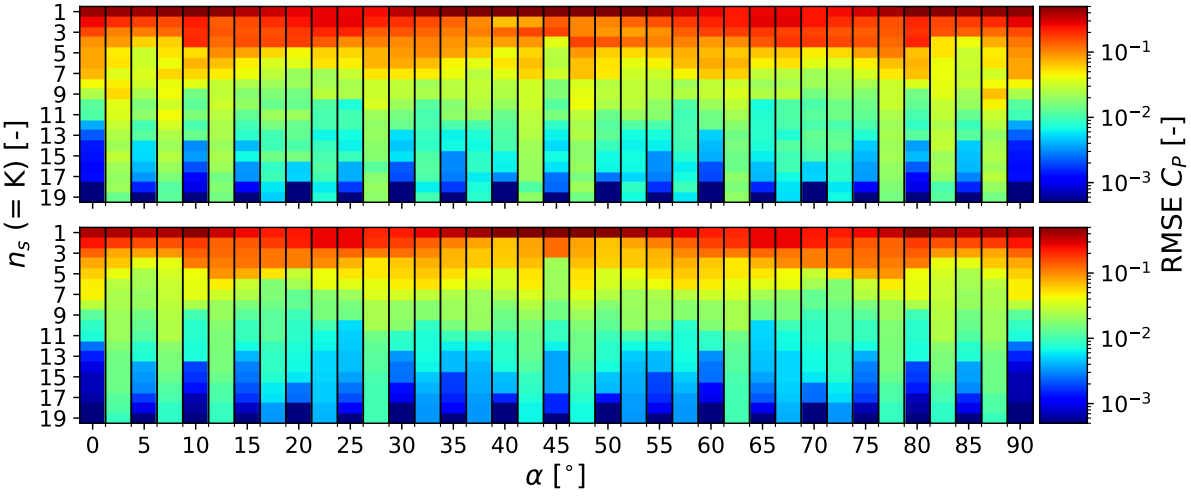


Figure 4.10: CFD GPOD pressure distribution reconstruction (top) and projection (bottom) accuracy for every combination of $n_s = K$ and α

The colormap shown on top presents the RMSE in terms of C_p for the aforementioned reconstructions. The bottom colormap shows the RMSE of the projection for any combination of angle of attack and number of bases used. These projections are slightly different from those mentioned in Section 2.3 as the coefficients in $\boldsymbol{A}$, required for such a projection, are only defined for the training data. Instead, these POD projections are achieved by performing GPOD reconstructions using $n_s = n$ (and $n_s \neq K$) as also mentioned in Section 2.5. This provides a best attainable GPOD reconstruction for the given number of modes also defined for reconstructions outside of the training data. Angles of attack for which pressure distributions were included in the training are indicated with labels on the x-axis but are not the focus of these figures. For those angles of attack, the RMSE drops to zero (or machine precision) at $K = m = 19$ for both reconstruction and projection as this reproduces the original training data exactly. Comparing the two colormaps; the projections always achieve lower RMSE's as expected however the reconstructions remain close in general. The reconstruction colormap shows individual angles of attack for which the reconstructions perform relatively poor. This is often the result of the sensor placement for a particular sensor budget benefiting certain angles of attack more than others, something often occurring with asymmetric sensor positioning for example.

The format of Figure 4.10 makes it difficult to asses how close projection and reconstruction are in a quantitative sense and specifically for the non-training angle of attack pressure distribution reconstructions. Therefore, the accuracy of reconstructions and projections across these non-training angles only are quantified using a single RMSE. This is again done for $n_s = K$ ranging from 1 to 19 which results in Figure 4.11. Note that this is not equal to averaging the RMSE across columns corresponding to non-training angles of attack due to the order of operations.

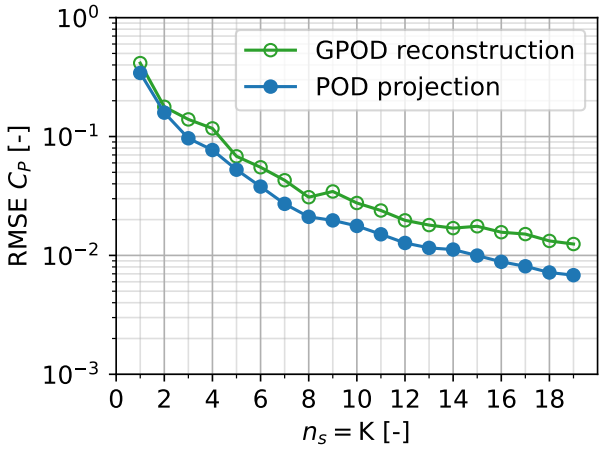


Figure 4.11: CFD GPOD surface pressure reconstruction accuracy for increasing number of modes and sensors

With each additional mode added to a projection being utilized to full extend as no 'gappiness' is considered, the projection curve decreases monotonically. The projection curve reaches RMSE values in the order of 10^{-2} already at 3 bases used and eventually drops below $1 \cdot 10^{-2}$. These low RMSE values compared to the magnitudes of the pressure coefficients of the numerical square cylinder simulations imply that

the POD bases obtained through the training data are representative for the pressure distributions used for the reconstructions. The differences between the projections and reconstructions remain small over the $n_s = K$ range with an average distance between the curves of $1.7 \cdot 10^{-2}$ (generally decreasing for increasing $n_s = K$). This demonstrates how the QR sensors are successful in sampling high modal activity across the used modes, corresponding to coherent patterns, and in the process allow the reconstructions to overcome introduced sparseness to a large extent. Despite the near-optimality of the QR algorithm for sensor placement (Brunton and Kutz, 2019), the greedy approach might lead to the 'trajectory' being followed during the sensor locations assignment for a particular sensor budget to be sub-optimal. This manifests itself in the somewhat more erratic and non-strictly decreasing behaviour of the GPOD reconstruction curve compared to the projection and is mainly recognizable at $n_s = K = 9$ for this particular case. Based on Figure 4.11, a choice on the required number of sensors for sparse sampling can be made with the aim of reaching a certain level of average accuracy in predicting the surface pressure.

4.6. Numerical drag estimation using GPOD

An approximation of the drag using sparse surface pressure measurements can be obtained by integrating the reconstructions discussed in the previous section. This allows for pressure drag approximations only which for bluff bodies like a square cylinder are expected to not differ much from the total aerodynamic drag (see Section 5.2.2 for a comparison between experimental pressure drag and balance drag results). The results of approximating drag through reconstructing pressure distributions with GPOD are discussed in the following sections. Note that all drag estimations are obtained by integrating surface pressure around the cylinder perimeter using the trapezoidal rule.

4.6.1. Drag estimation examples

An example of the approximated drag coefficients over the complete angle of attack range is provided below in Figure 4.12 for $n_s = K = 10$. The reference drag curve included is obtained by integrating the known full pressure distribution corresponding to a sparse reconstruction.

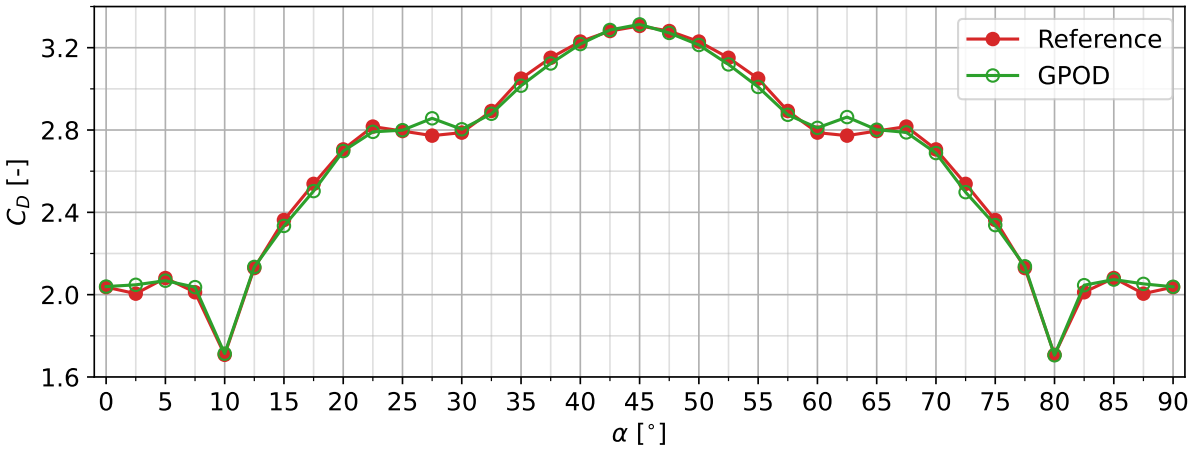


Figure 4.12: CFD GPOD drag curve reconstruction using $n_s = K = 10$

Angles of attack indicated with a ticklabel are again those originally included in the training data and will therefore not be the focus of the discussion. Using $n_s = K = 10$, the RMSE in terms of the C_p was shown earlier to be small on average and yielded reconstruction examples such as those shown in Figure 4.8 and 4.9 which clearly resemble the reference very well. It is therefore not surprising to see the approximated drag coefficients generally lie in close proximity to the reference drag across the angle of attack range. Despite visible discrepancies between reference and reconstruction in Figure 4.8 and 4.9, the corresponding approximated drag coefficients almost perfectly match those of the reference which highlights an inherent aspect of computing drag through surface pressure integration; where determination of the RMSE between surface pressure distributions does not consider the sign of the errors made, integration of the reconstructed surface pressure does depend on the reference pressure being under- or overpredicted. For $\alpha = 12.5°$ and $42.5°$, the combination of cylinder orientation and pressure reconstruction, especially on faces C and D, result in errors cancelling in the computation of the drag coefficient. This does not happen for the reconstruction at $\alpha = 27.5°$ or $63.5°$ for example where errors in the drag estimation are the result of accumulating pressure underprediction across cylinder face C or D. Such relatively large drag estimation errors do not seem to happen as clearly for any of the remaining angles of attack however illustrates how accurate surface pressure distribution reconstructions do not necessarily yield drag predictions accurate to the same extent. The reference drag curve is perfectly symmetric through $\alpha = 45°$ by construction however this does not hold in general for GPOD reconstructions. The GPOD curve being visually symmetric still is a result of the even number of sensors being positioned almost symmetrically about the cylinder diagonal by the QR algorithm.

4.6.2. Drag estimation accuracy

Drag curve reconstructions such as in Figure 4.12 are obtained for $n_s = K = 1$ to 19. By only considering the drag coefficient approximations corresponding to 'non-training' angles of attack, the GPOD drag estimation accuracy is quantified in terms of RMSE with respect to the reference drag. The resulting graphs are shown in Figure 4.13 which again includes a POD projection curve obtained by integrating the POD projection surface pressure instead of the reconstructions.

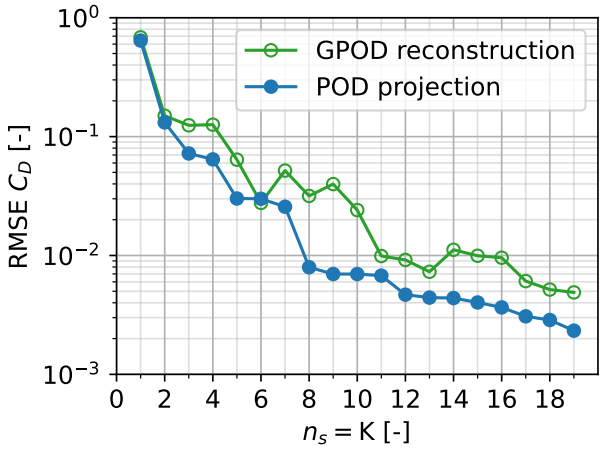


Figure 4.13: CFD GPOD drag curve reconstruction accuracy for increasing number of modes and sensors

The behaviour of both the projection and reconstruction curves in Figure 4.13 is clearly very different from that observed in Figure 4.11. The projection curve is not smoothly decreasing but instead follows decreasing steps for increasing K. These sudden drops in the projection correspond to the inclusion of eigenmodes that focus most modal activity on faces C and D of the cylinder. Such modes offer a correction to the pressure reconstructions on faces C and D which generally contain a large part of the C_D error in the form of accumulating constant under- or over prediction. The projection loses its meaning as a best attainable GPOD reconstruction when quantifying drag estimation accuracy. This is clearly visible at $n_s = K = 5$ where the RMSE of the reconstruction falls below the projection. Such phenomena can be caused by sensor locations involved in pressure distribution reconstructions constraining the eigenmodes such that 'coincidentally' accurate drag coefficient approximations result despite a larger RMSE in terms of C_p with respect to the projection. An example of this is the pressure distribution and corresponding drag coefficient reconstruction at $\alpha = 42.5°$ using $n_s = K = 10$ shown in Figure 4.9 and discussed in Section 4.6.1. With an increase in sensor budget and number of modes, one effectively increases the information available to be sampled from the QR sensors. This generally results in decreasing C_p RMSE, unless the greediness of the QR algorithm fails to exploit this. The decrease in C_p RMSE however is often realized by a significantly changed pressure distribution reconstruction around the added sensor which alters the drag estimation accuracy for each increase in $n_s = K$. This causes the erratic behaviour to appear in Figure 4.13 for reconstructions where this is not the case to the same extent for the projections. Despite accuracy gains for increasing sensor budget not being as predictable in terms of C_D RMSE compared to that of C_p, the general trend is clearly decreasing. As can be seen from the drag coefficients in Figure 4.12, which achieved an overall RMSE in terms of C_D of $2.4 \cdot 10^{-2}$, drag prediction close to the reference is possible at a significantly decreased number of sensors while providing physically valid pressure distributions as well.

4.7. CFD based sensor placement

Despite large discrepancies between the CFD results and literature in especially quantitative sense, the characteristics and patterns observed in the numerical pressure distribution results have been used for the experimental test object by forming the basis of additional pressure tap plates to be used. These were already shown in Section3.2 however the reasoning behind cosine, Opt V1 and Opt V2 pressure tap plates will be discussed briefly.

4.7.1. Tap distribution for pressure drag determination

For various angles of attack, the flow separation from the cylinder corners induces steep pressure gradients also visible in the CFD results. Such gradients might require increased spatial resolution to be resolved sufficiently and obtain representative experimental pressure distributions. For the CFD simulations this is not of concern however the spatial resolution achievable during experiments is considerably lower with at most 32 taps available around the object perimeter. The CFD pressure distribution results are sampled around the cylinder perimeter with generic even spacing and cosine spacing which has decreased distance between taps towards the corners. The resulting pressure distributions are integrated and the drag found compared to the original high resolution CFD drag coefficients. This is done using 1 to 8 taps on each cylinder face.

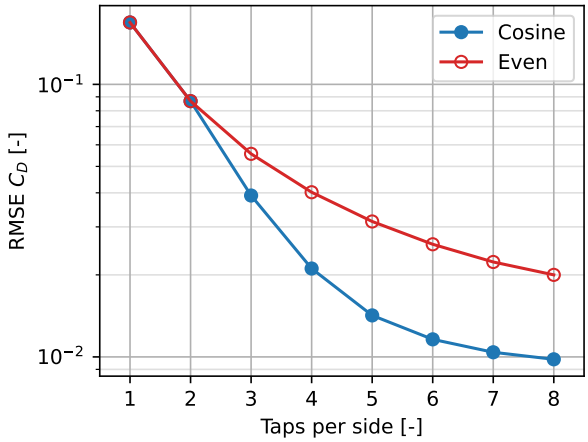


Figure 4.14: RMSE in the CFD C_D for even or cosine spaced pressure taps

Results are shown in Figure 4.14. Even using a small number of taps, clustering towards the cylinder corners with cosine spacing seems to result in slightly more accurate drag coefficients than evenly spaced taps. With at most 32 taps available for the experiments, each face can have a maximum of 8 taps installed. At 8 taps per side, the cosine tap spacing achieves a smaller RMSE C_D compared to the even spacing but the difference is small at $1 \cdot 10^{-2}$ since increasing tap density near the edges requires sacrificing it near the center of a cylinder face. Regardless, a set of pressure tap plates containing 32 taps with each plate having 8 taps following a cosine spacing is manufactured to be used during experiments as well. A schematic of the tap locations was shown earlier in Figure 3.4.

4.7.2. QR-CP sparse sensor locations

The QR-CP algorithm was found to produce sensor locations capable of keeping GPOD reconstructions close to POD projections when used in combination with the CFD results. The same sensor positioning approach would work with experimental pressure distributions obtained through 32 pressure taps as well however no sensor locations outside of the original 32 taps can be selected for sparse reconstructions. To overcome this, the high spatial resolution of the CFD results can be used to determine sensor locations for use during the experiments instead. Since QR sensor locations are determined though the POD modes, this requires a certain level of resemblance between the CFD eigenmodes and those obtained from the experimental campaign to work effectively. CFD based QR sensor locations have been made into two sets of pressure tap plates on top of the even and cosine spaced plates.

The sensor locations used in the GPOD reconstruction Figures. 4.8 and 4.9 were obtained at $n_s = K = 10$ with training data consisting of the pressure distributions corresponding to angles of attack being multiples of $5°$. These 10 sensor locations correspond to positions on the cylinder perimeter indicated by tap 1 until 11 except for tap 2 as shown in the Opt V2 row of Table. 4.2. Using $n_s = K = 5$ instead yields QR sensor locations corresponding to tap 2, 4, 7, 10 and 11 of the Opt V2 row in Table. 4.2 and are thus almost entirely nested in those of $n_s = K = 10$. Both these sensor location sequences identified for more or less sparse reconstructions are made into a set of pressure tap plates referred to as Opt V2, containing taps at the 11 locations tabulated. A visual representation of the Opt V2 tap locations was presented in Figure 3.4. A final set of pressure tap plates was made by changing the training data. By instead using the CFD pressure distributions corresponding to angles of attack $0°$ to $90°$ in $10°$ intervals including $\alpha = 45°$ as training, the eigenmodes change and therefore the QR sensor locations as well. With this training data, sensor locations for $n_s = K = 10$ are tap 1 until 10 of the Opt V1 row in Table. 4.2. Using $n_s = K = 5$, tap 1, 4, 6, 7 and 9 are used which are thus completely nested into the taps for $n_s = K = 10$. These tap locations were also included in a schematic view in Figure 3.4.

Table 4.2: Opt V1 and Opt V2 sensor locations given by sensor angle ψ.

	Tap	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Opt V1	ψ	$5°$	$45°$	$74°$	$88°$	$92°$	$110°$	$158°$	$232°$	$273°$	$299°$	
Opt V2		$5°$	$29°$	$58°$	$76°$	$88°$	$92°$	$104°$	$122°$	$175°$	$267°$	$273°$

Both these pressure tap plate sets have been used during experimental surface pressure measurements as described in Section 3.5.1. The results will therefore be shown and discussed in Section 5.1.2. Unfortunately however, asymmetry with respect to the square cylinder diagonal for both tap configurations on the Opt V1 plates and the $n_s = K = 5$ tap configuration on the Opt V2 plates makes it difficult to use the corresponding pressure measurements for GPOD reconstructions as the results can not be mirrored. The tap configuration for $n_s = K = 10$ using Opt V2 plates is symmetric and therefore its results can still be used for GPOD reconstructions which will be discussed in Section 5.8.

5

Results and discussion

In the following chapter all results based on the windtunnel campaign are presented an discussed. First, in Section 5.1, the time averaged surface pressure distributions are shown for a square cylinder at angles of attack from $\alpha = 0°$ to $45°$ and also for each of the used pressure tap layouts. The surface pressure distributions are discussed and compared to experimental results reported in literature. Force balance measurements are also used in a comparison with literature and provide a method of independent validation for the experimental surface pressure distribution measurements as well in Section 5.2. In Section 5.3, the square cylinder flow eigenvalue spectrum and dominant eigenmodes are analysed using POD. These eigenmodes provide the basis for GPOD reconstructions using only sparse surface pressure measurements of which the pressure distribution reconstruction accuracy and drag estimation accuracy is quantified in Section 5.4 and Section 5.5 respectively. Based on the insights into GPOD reconstruction behaviour obtained, some typical sources of sparse reconstruction errors are discussed in Section 5.6. Accuracy of both surface pressure reconstructions and drag predictions by GPOD are compared to linear interpolation in Section 5.7. The remaining Section 5.8 and Section 5.9 describe the attempts of fusing CFD into the GPOD procedure to an increasingly large extent.

5.1. Surface pressure measurements

Each of the tap distributions shown in Section 3.2 have been used for surface pressure distribution measurements twice following the information in Section 3.5.1. The average over the two $\alpha-$sweeps per tap distribution are computed which will be shown and discussed in the following sections.

5.1.1. Even and cosine spaced taps

For clarity of the figures, only pressure distributions corresponding to angles of attack being multiples of $5°$ are shown in Figure 5.1. Markers on the pressure distributions correspond to the actual locations of individual pressure taps on the cylinder perimeter. The coordinate system used is the same as that shown in Figure 3.4.

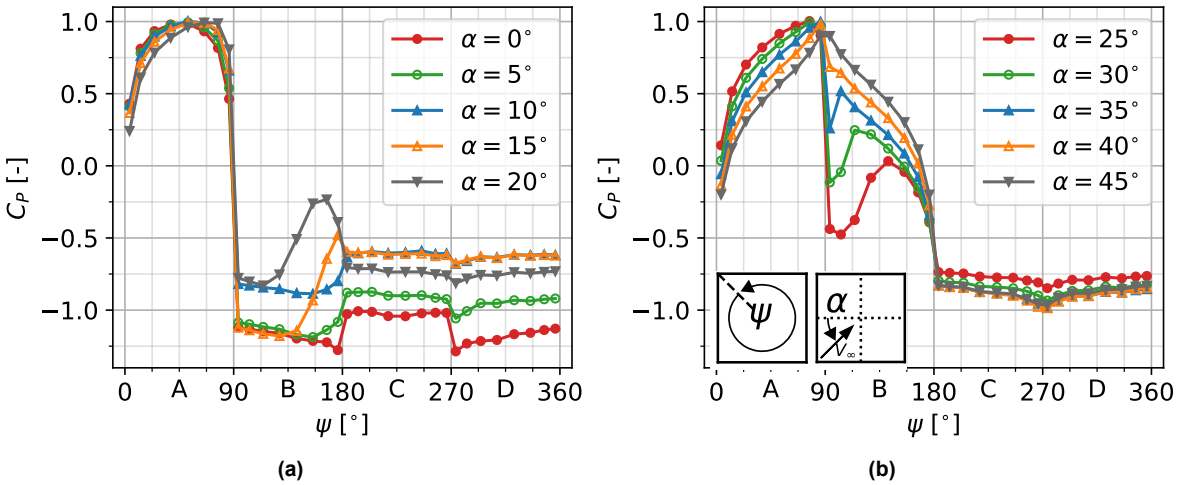


Figure 5.1: Experimental time averaged pressure distributions for $\alpha = 0°$ to $20°$ (a) and $\alpha = 25°$ to $45°$ (b), evenly spaced taps. Convention for angles ψ and α shown in bottom left of (b)

Throughout the measured angle of attack range, side A tends to face the incoming flow and experiences no flow separation. At $\alpha = 0°$, a pressure coefficient of 1 is reached around $\psi = 45°$ indicating the stagnation point. As expected, this stagnation point moves towards the corner at $\psi = 90°$ for increasing angle of attack, skewing the pressure distribution on face A in the process. Faces B, C and D experience more complex flow phenomena at different angles of attack and follow the flow regime classifications by Huang et al. (2010) described in Section 2.9.1.

Below the critical angle of attack (typically between $10°$ and $15°$), Figure 5.1a shows the pressure on face C and D increasing for increasing angle of attack. This is caused by elongation of the wake and consequent downstream shift of the vortex formation region (van Oudheusden et al. (2008), Lee (1975)). At the critical angle of attack, pressure on face C reaches a maximum (corresponding to minimum drag) and the shear layer separated from $\psi = 90°$ reattaches just upstream of $\psi = 180°$ on face B. At $\alpha = 15°$, the pressure on face C seems to reach a maximum equal to that of $\alpha = 10°$ while it shows signs of well established reattachment on face B already through the peak in pressure. A more precise identification of the critical angle of attack therefore requires the distribution at $\alpha = 12.5°$ as shown in Figure 5.2.

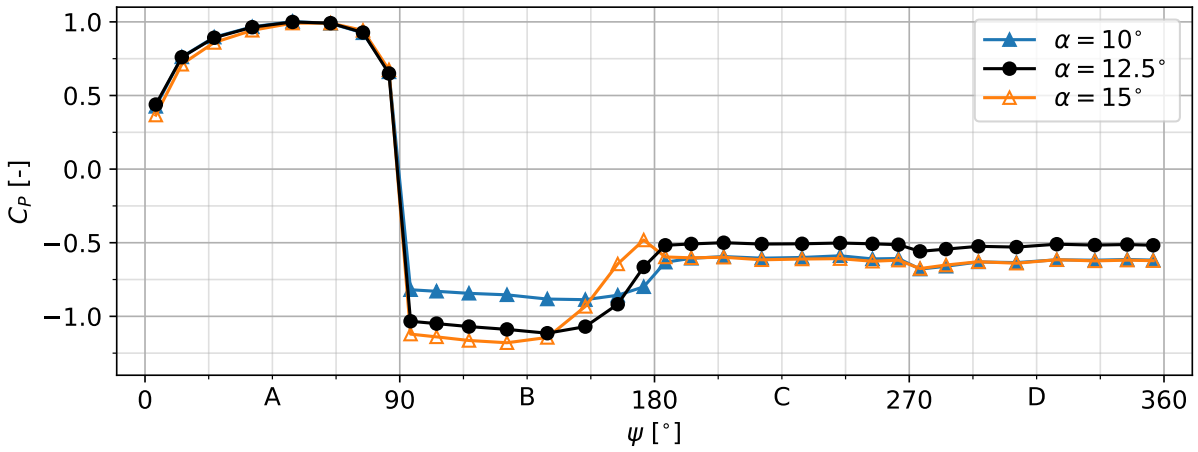


Figure 5.2: Experimental surface pressure distributions around the critical angle of attack, evenly spaced taps

The above figure clearly shows how a maximum base pressure is reached at $\alpha = 12.5°$ instead. Also, pressure on face B close to $\psi = 180°$ is increased with respect to $\alpha = 10°$ but does not contain a 'reattachment peak' as for $\alpha = 15°$ meaning that $\alpha = 12.5°$ can be identified as the critical angle of attack from the measurements. This critical angle of attack marks a transition in the behaviour of the pressure distributions, mainly on faces B, C and D. Above the critical angle of attack, the wake instead shortens bringing the vortex formation region closer to the cylinder, decreasing pressure on faces C and D. This effect is clearly visible in Figure 5.2 where the base pressure at $\alpha = 15°$ is practically equal to that at $\alpha = 10°$ and also decreases steadily for increasing angle of attack as seen in Figure 5.1b. With face B becoming increasingly exposed to the external flow above the critical angle of attack, the point of reattachment traverses upstream towards $\psi = 90°$. The exact location of reattachment is difficult to identify from the surface pressure distributions; however the peak in pressure encompassing the reattached region can clearly be seen to increase in size (spatial extent and pressure at peak) in Figure 5.1b. Above $\alpha = 5°$, the pressure distributions on faces C and D become similar in magnitude as the recirculation region on face D is expected to have merged with the recirculation region in the wake following the results and discussion by van Oudheusden et al. (2008). A slight asymmetry remains however in the form of a small pressure dip on face D near $\psi = 270°$ which gradually decreases as the angle of attack reaches $\alpha = 45°$ and the pressure distribution is roughly symmetric. This small dip seems to indicate the presence of a small secondary recirculation region on face D caused by the reversed flow not being able to follow the sharp cylinder corner at $\psi = 270°$ as it leaves face C. Note that this phenomena seems to be somewhat visible in the earlier shown results of both van Oudheusden et al. (2008) and Roosenboom (2005) but is not mentioned explicitly. Since the pressure distribution over faces C and D does not change significantly from $\alpha = 35°$ onwards, it is believed that the aforementioned secondary circulation bubble has merged with the wake and the flow topology in the wake has reached a similar state to the wedge flow reported by Huang et al. (2010) to occur at $\alpha = 45°$. This flow topology is characterized by two large recirculation regions in the wake which completely engulf faces C and D leading to little variance in the pressure distribution on those faces as seen in Figure 5.1b. This con-

firms how the formation and growth of a secondary set of recirculation regions near $\psi = 270°$ as visible in the CFD results is completely unphysical.

Pressure distributions obtained at the same angles of attack but instead using the cosine spaced pressure taps are shown below in Figure 5.3. No changes in the overall flow phenomena are expected nor observed and thus the discussion of the cosine pressure distributions will mainly focus on differences in small details compared to the even placed tap results.

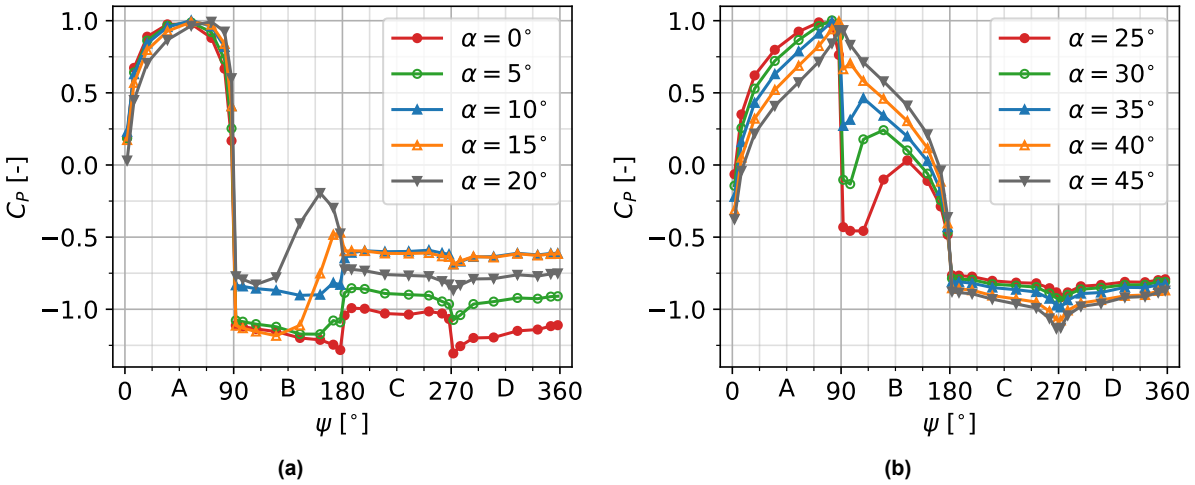


Figure 5.3: Experimental time averaged pressure distributions for $\alpha = 0°$ to $20°$ (a) and $\alpha = 25°$ to $45°$ (b), cosine spaced taps

The general shapes of the pressure distributions and relative order between the curves at different angles of attack match very well with the earlier shown results, indicating repeatability of the measurements after replacement of the pressure tap plates. Differences between the even and cosine pressure distributions as a result of the different tap spacing are mainly visible on face B and near $\psi = 270°$ on faces C and D. Within the subcritical regime, a slight dip in the pressure is visible on face B near $\psi = 180°$ which was not visible in the even distributions. This small drop is caused by a similar phenomena responsible for the dip near $\psi = 270°$ as the periodic shedding of vortices in the wake produces a time averaged recirculation region with flow along face C that is not able to follow the sharp corner at $180°$, yielding a small recirculating region on face B near this corner. This recirculation region has a smaller effect than that near $\psi = 270°$ because of the orientation of the cylinder at positive angle of attack and it disappears due to the reattachment of the shear layer from the critical angle of attack onwards. Reattachment of the shear layer produces steep pressure gradients. When the angle of attack is such that the reattachment point is near $\psi = 90°$ or $180°$, the cosine spacing can resolve the gradients near these corners better however around $\psi = 135°$ the even spacing has smaller spacing between taps. In the sub- and supercritical regime, the cosine spacing is able to better capture the small region of recirculation near $\psi = 270°$ on face D with accompanying local decrease in pressure. Once the flow topology in the wake has reached a wedge like state (roughly expected from $\alpha = 35°$ onwards), the two large recirculation regions produce flow with a downstream, time averaged, stagnation point around corner $\psi = 270°$.This locally induces relatively

large in velocity magnitude reversed flow over faces C and D leading to the decrease in pressure which is clearly better captured by the cosine spaced taps.

5.1.2. Opt V1 and Opt V2 taps

The Opt V1 and Opt V2 plates pressure taps are distributed too sparse for sufficiently resolved pressure distributions around the cylinder perimeter to be presented. Instead, the variation in pressure of individual taps are shown in Figures. 5.4 and 5.5 over the measured angles of attack. These figures provide a different perspective on the pressure variation as a function of angle of attack over the cylinder sides. Taps on both Opt V1 and V2 are numbered based on their sensor angle ψ.

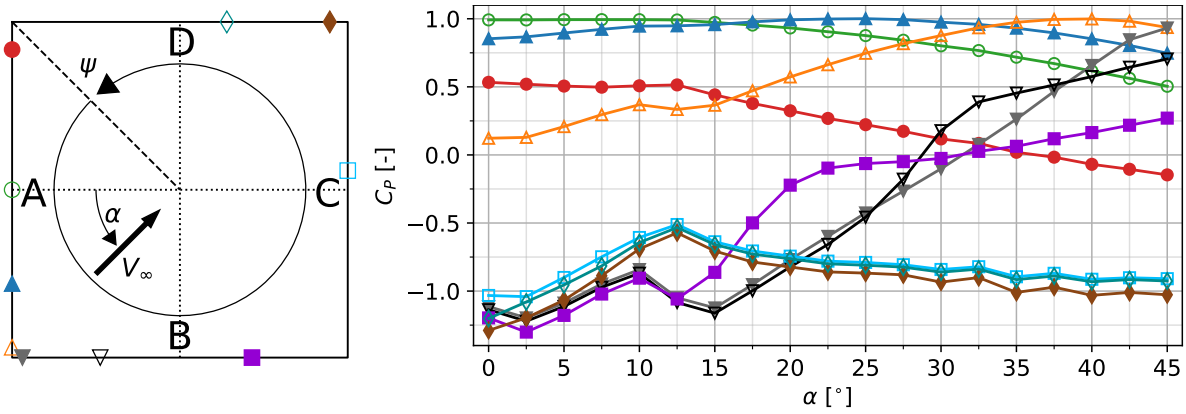


Figure 5.4: Experimental time averaged pressure tap measurements for $\alpha = 0°$ to $\alpha = 45°$, Opt V1 taps

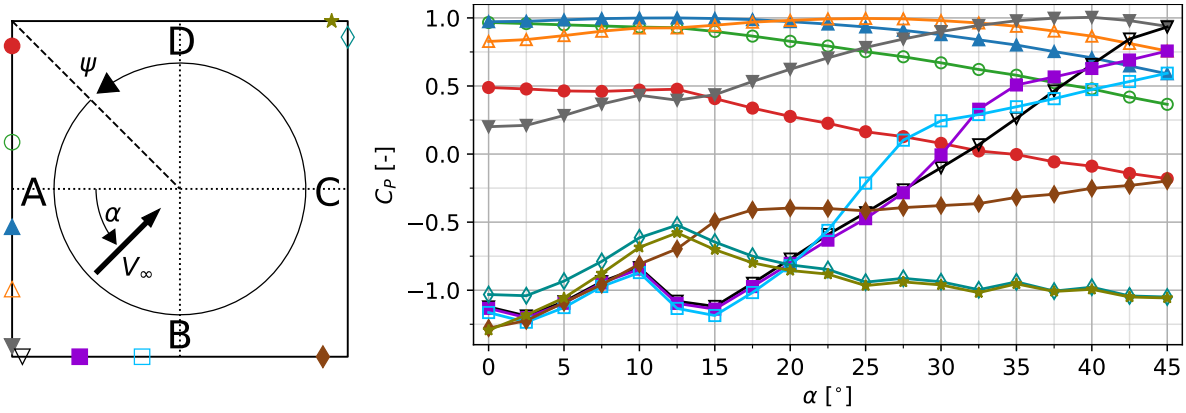


Figure 5.5: Experimental time averaged pressure tap measurements for $\alpha = 0°$ to $\alpha = 45°$, Opt V2 taps

The same square cylinder flow phenomena as discussed before are visible in these figures as well. For small angles of attack, pressures of taps located on face A for both Opt V1 and V2 are positive as they face the incoming flow. Tap 2 for Opt V1 is located at $\psi = 45°$ and therefore measures stagnation pressure at $\alpha = 0°$. Increasing the angle of attack decreases pressure measured by taps at $\psi < 45°$ as the stagnation point moves away, conversely the pressure increases for taps on face A at $\psi > 45°$. Within the subcritical regime, pressure on face B can be seen to increase (taps

5,6,7 Opt V1, taps 6,7,8,9 Opt V2) with increasing angle of attack as the shear layer intermittently reattaches. With steady reattachment of the shear layer on face B at the critical angle of attack $\alpha_{crit} = 12.5°$, pressure drops across face B except for the region in proximity to the reattachment point such as tap 9 at the Opt V2 plates. While the reattachment point traverses upstream for increasing angle of attack, the pressure measured by taps on face B increases; steeper and sooner for taps positioned closer to $\psi = 180°$.

Faces C and D contain a small number of taps for both Opt V1 and Opt V2 plates and do not experience a transition from separated to attached flow, leading to relatively small pressure changes for varying angle of attack. Clearly recognizable is the increase in base pressure in the subcritical regime towards a maximum at α_{crit} through tap 8 of Opt V1. Pressure measured by this tap decreases again after α_{crit} and is proportional to the position of the vortex formation region with respect to the cylinder; increasing base pressure in the subcritical regime as a result of an elongated wake and downstream moving vortex formation region, decreasing pressure in the supercritical regime with the recirculation regions moving closer towards the body (see Figure 2.23). Pressure measured by tap 10 on Opt V2 is equal to that of tap 11 at $\alpha = 45°$ as a result of the symmetry about the cylinder diagonal. This can be seen to hold true for all 'couples' of taps on Opt V2 as the tap distribution is symmetric with the exception of tap 2.

5.1.3. Comparison to literature

The measured surface pressure distributions adhere to the qualitative flow descriptions of Igarashi (1984) and Huang et al. (2010). To see if the results agree with those reported in literature in a quantitative sense as well, the pressure distributions at $\alpha = 0°$ and $12.5°$ are compared with earlier referenced papers. A comparison across all measured angles of attack will be discussed in Section 5.2 in terms of the drag coefficients instead.

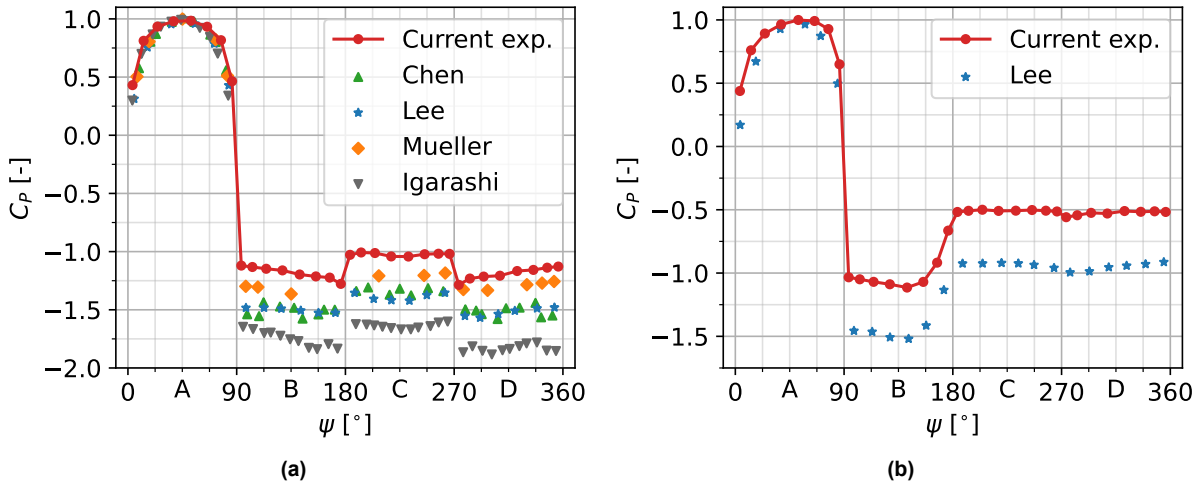


Figure 5.6: Current experiment pressure distribution comparison to available experimental result at $\alpha = 0°$ (a) and $\alpha = 12.5°$

Most of the research on square cylinder flow focuses on the $\alpha = 0°$ case. In Fig-

ure 5.6a, pressure distributions at this angle of attack from Chen and Liu (1999), Lee (1975), Mueller (2012) and Igarashi (1984) are shown together with the current experiment's results. The low pressure sides, B, C, and D at this α, are of most interest as they contain considerable variance in the by literature reported measurements. Pressure coefficients reported by Igarashi are considerably lower with respect to the other references which in turn are all offset below the current experiment's results. Igarashi himself attributes the deviation partly to wind tunnel blockage (up to 10 %) however aspect ratio and wind tunnel test section type likely also play a role. All shown pressure distributions at this angle of attack agree in shape however with pressure on face C being increased with respect to faces B and D.

Pressure distributions reported close to the critical angle of attack are more scarce and are relatively dependent on the Reynolds's number. Lee however performed experiments at a Reynold's number close to the current experiment's ($Re = 1.76 \cdot 10^5$ compared to $1.35 \cdot 10^5$) and reported the critical angle of attack pressure distribution which is shown in Figure 5.6b. Lee measured the exact same critical angle of attack being 12.5° and the shape of the distribution matches the current experiment's almost perfectly. A similar offset of nearly 0.5 in terms of C_p as for $\alpha = 0$° persists however. This seems to indicate a systematic difference between the current experiments and those of the reference papers. This will be discussed more in Section 5.2 where its effect over all angles of attack is clearly recognizable in the C_D versus α curve.

5.2. Drag measurements

Force balance measurements performed simultaneously with the surface pressure measurements provide quantitative drag forces independent of the pressure tap layout used. These balance drag measurements are compared to literature first in Section 5.2.1. After this, differences between obtaining drag through a force balance compared to surface pressure integration using different tap layouts will be discussed in Section 5.2.2.

5.2.1. Force balance C_D versus α and comparison to literature

With four sets of pressure tap plates, each used for two α-sweeps, a total of eight $\alpha-$sweeps are performed during which balance measurements have been acquired as well. Unlike the surface pressure measurements using different tap layouts, all balance measurements performed at the same angle of attack are compatible, allowing for a single force balance C_D vs α curve to be obtained by averaging over the eight times each angle of attack is used during measurements. This drag curve is shown including the standard deviation at each angle of attack in Figure 5.7. Drag curves as measured by selected literature are included as well for reference.

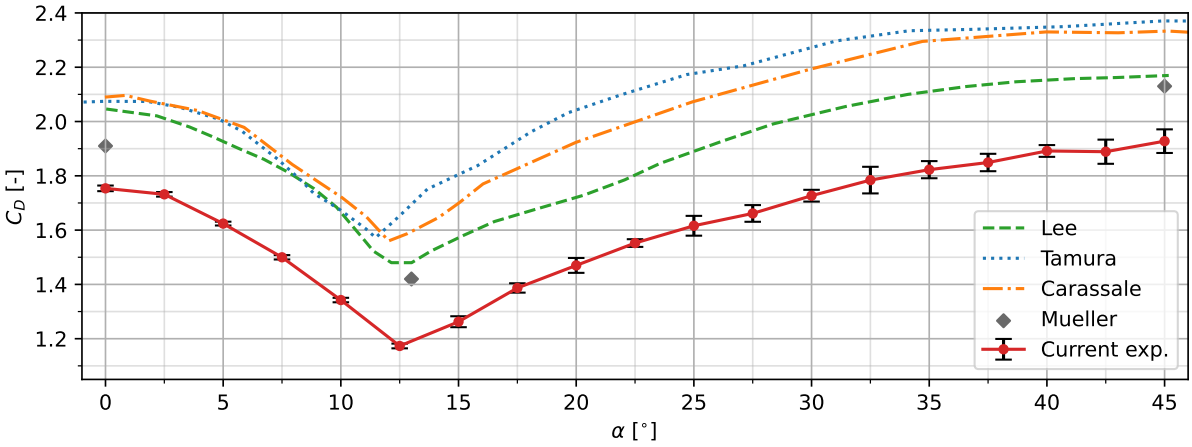


Figure 5.7: Drag coefficient versus angle of attack as measured by force balance during experimental campaign. Including reference results from literature

The phenomena causing the pressure measurements, especially on faces B, C and D, to be more negative with respect to literature seem to have been picked up by the force balance as well. This causes an almost constant offset in the drag coefficients over the entire angle of attack range. Regardless, the shape of the current experiment's drag curve adheres to the square cylinder flow regime classifications and is very similar to the provided references. A clear distinction between the sub- and supercritical flow regime is marked by the change in slope corresponding to a critical angle of attack of $12.5°$. This agrees well with those from the references despite Carassale et al. (2014) stating a strong sensitivity to test conditions. The variance in the drag coefficients of the subcritical regime are noticeably smaller than those towards $\alpha = 45°$. The magnitude of the variances seems proportional to visually observed oscillations of the test object during the measurements. Variation in stiffness of the test object as a function of angle of attack might therefore have played a role in the repeatability of measurements at different angles of attack.

The remarkable offset of the measured drag curve compared to the provided reference drag curves is possibly caused by two reasons. First, all reference drag curves, except for Mueller, have been obtained through measurements in closed test section windtunnels different from the open test section used in the current experiment. The presence of walls around the entire model eliminates the possibility of downwash over the model interfering with the vortex formation behind the cylinder. Zhao et al. (2021) investigated that for a true three-dimensional square cylinder flow, such downwash increases the downstream distance at which the vortex street forms, which increases base pressure hence decreases drag. In other words, deviation from two-dimensionality as a result of an open test section might have delayed the vortex formation despite the attempt to counter this using large end plates. The relatively low aspect ratio of the current experiment's test object could have enhanced the influence this has as well. Second, none of the included references corrects for blockage as the blockage ratio is generally low. Regardless this might lead to a slight overprediction in drag. Only the experiments performed by Mueller are performed in a somewhat similar open test sectioned windtunnel while using end plates as well which could be

the reason for those results being the closest to the current experiment's.

5.2.2. Drag through force balance versus surface pressure integration

Force balance measurements were mainly taken for validation of the surface pressure measurements as they are assumed to measure the true total drag experienced. Since gappy POD will be used for drag estimation by means of reconstructing sparse surface pressure distributions, the non-sparse surface pressure distributions themselves should yield an accurate drag estimate to the balance results. Drag from surface pressure is obtained through integration over the cylinder perimeter for the even and cosine spaced pressure tap plates. Resulting C_D versus α curves are shown in Figure 5.8 where they are compared to the balance results. Note that surface pressure results were averaged over the two α-sweeps performed for each set of plates, drag measurements shown are therefore averaged over the two sets of measurements corresponding to the surface pressure measurements.

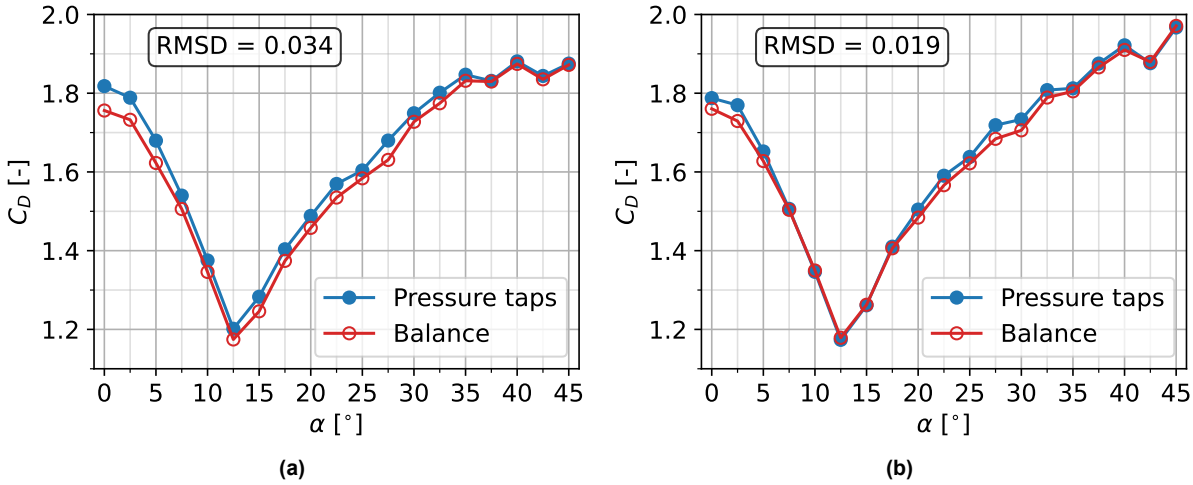


Figure 5.8: Integrated pressure drag compared to balance drag for even (a) and cosine (b) spaced taps. Root Mean Squared Difference (RMSD) between pressure tap and balance values reported in the box at the top left.

Following the expectations based on Section 4.7.1, the drag coefficients obtained by integrating the cosine spaced pressure distributions are slightly closer to the balance measurements in terms of root mean squared deviation. Regardless, with a RMSD of 0.034 in terms of C_D and a maximum difference of $3.5\,\%$ with respect to the balance, the even spaced taps results are also very representative of the experienced drag. In both figures, the pressure tap drag curve lies slightly above the balance drag curves. By obtaining drag through integrating surface pressure along a single section of the cylinder while using the model side length as a reference, an indirect assumption is made that the integrated pressure distribution spans the entire height of the model. Near the top and bottom of the model however, a boundary layer originating from the end plates will cause the pressure distribution to locally be decreased in magnitude by the decrease in flow velocity resulting in an overprediction in total drag when unaccounted for. This seems to overpower the underprediction of total drag by the

pressure taps as a result of neglecting friction, which is not surprising for bluff bodies. Overall, drag obtained through pressure taps and balance agree well and since sectional pressure distribution integration inherently assumes two-dimensional flow while the force balance does not, it seems like a nominally two-dimensional flow has thus been realized past the model.

5.3. Surface pressure POD analysis

The experimental even and cosine surface pressure measurements are in the same format as the numerical results; application of the POD procedure is therefore identical to that described in Section 4.4. Surface pressure distributions from $\alpha = 0°$ to $42.5°$ are again mirrored to obtain 37 surface pressure distributions spanning the range of $\alpha = 0°$ to $90°$ at a $2.5°$ interval. From these 37 pressure distributions, those 19 corresponding to angles of attack being an integer multiple of $5°$ (including $\alpha = 0°$) are used as training data, the remaining distributions are to be focused on for reconstruction. Since experimental pressure distributions using even and cosine tap layouts are obtained using 32 taps, the training data stored in a matrix has shape $m = 19$ and $n = 32$. Note that the average over the angles of attack has been subtracted.

5.3.1. Eigenvalue spectrum

POD applied to the aforementioned training data yields the eigenmodes in $\boldsymbol{\Phi} \in \mathbb{R}^{n \times m}$ and α-coefficients in $\boldsymbol{A} \in \mathbb{R}^{m \times m}$. The eigenvalues used for ordering the eigenmodes provide insight into the energy distribution or variance capture over the modes and are shown in Figure 5.9. Note that POD is applied to data obtained using even and cosine tap layouts separately.

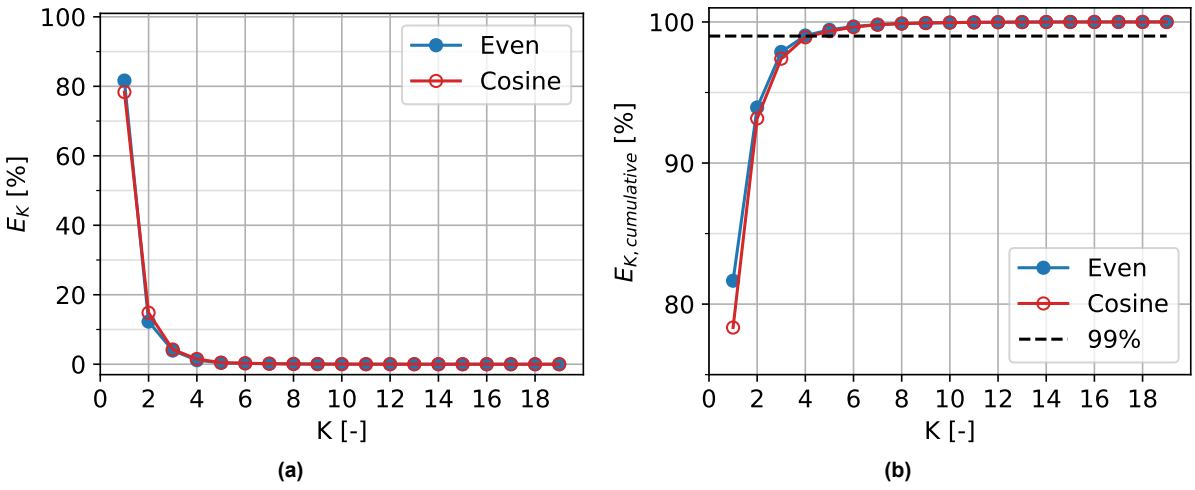


Figure 5.9: Individual (a) and cumulative (b) energy capture of the eigenvalues corresponding to experimental pressure distributions

The definitions of E_K and $E_{K,cumulative}$ are stated in Equation 2.19. The individual eigenvalues of both the even and cosine data decay rapidly in roughly the same manner. The first three eigenvalue energy levels for the even data are $82\,\%$, $13\,\%$ and $4\,\%$ respectively. The first three for cosine data are $78\,\%$, $15\,\%$ and $4\,\%$. The magnitude of the first eigenvalue for both even and cosine data is thus significantly larger

than that of the CFD results. This is caused by the much higher spatial resolution of the CFD results being able to resolve finer details and the non-physical behaviour in the numerical distributions which both dilute the energy contents of individual modes. Out of the 19 eigenvalues, 4 are enough for the even data to reach the $99\,\%$ mark in Figure 5.9b while 4 cosine eigenvalues fall just short with $98.9\,\%$. The differences between even and cosine tap spacing might seem small but the difference in energy distribution over the eigenvalues and thus modes relates back to the differences in tap layout and their effect in capturing aspects of the pressure distributions. This will become clearer in Section 5.3.2 through the spatial modes. The rapid decay in eigenmodes seems promising for accurate lower rank GPOD reconstructions, however the reconstruction accuracy will also depend on the sensor locations when only sparse measurements are available and the ability of the modes to represent pressure distributions not included in the training data.

5.3.2. Dominant eigenmodes

In Figure 5.10, the first two POD modes and corresponding α-coefficients are shown from the experimental surface pressure distributions. The eigenvalue decay was shown previously to be very rapid, the two modes shown indeed make up over $90\,\%$ of the total energy and therefore present a low rank representation of the dominant underlying features of square cylinder surface pressure distributions.

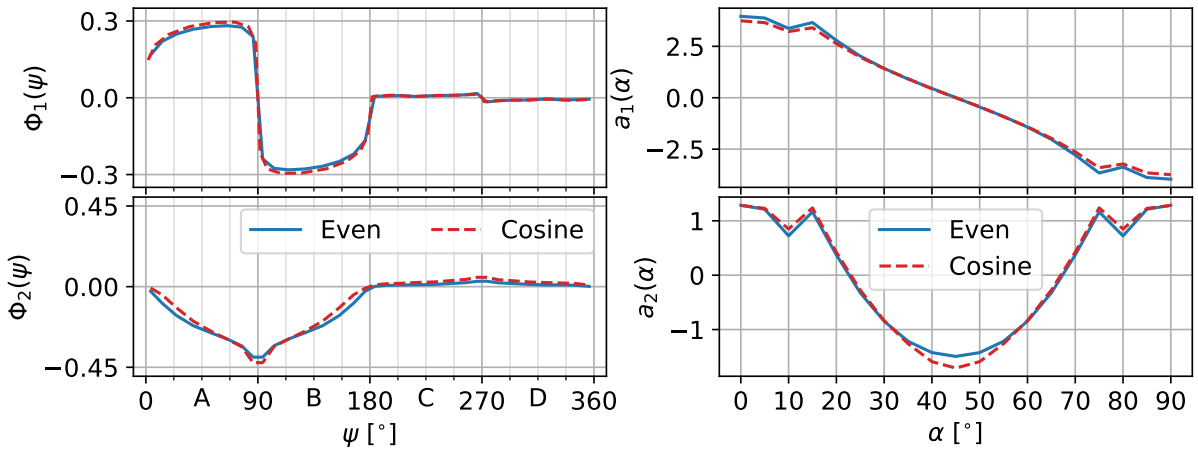


Figure 5.10: First and second spatial POD mode (left), first and second α-coefficients (right) from experimental results

Pressure distributions obtained using even or cosine spaced taps overall showed minor differences and hence especially the first few modes and coefficients are very similar in shape. As was also the case for the CFD results, the first POD mode resembles the experimental pressure distribution at $\alpha = 0°$ or $90°$.The first set of α-coefficients is almost completely linearly decreasing such that together with the first mode, the shifting high pressure from face A towards face B is captured. Little modal activity is shown on faces C and D besides the jump at $\psi = 270°$ which will be scaled according to the value of the α-coefficient. This illustrates how the pressure on faces C and D tends to be separated by a steep pressure jump near $\alpha = 0°$ and $90°$ but not close to $\alpha = 45°$.

As was also the case for the CFD results, the second experimental mode for both even and cosine tap layouts resembles the pressure distribution at $\alpha = 45°$ and is also most active in that region as visible in the α-coefficients. Unlike the first POD mode and set of coefficients, differences between even and cosine tap layouts are more pronounced in the second mode and especially in the coefficients. This links back to the differences in eigenvalues observed in the eigenspectra as shown in Figure 5.9. The effect of increased tap density near the cylinder corners for the cosine tap layout compared to the even layout is mainly visible for pressure distributions close to $\alpha = 45°$. For those pressure distributions, a sharp drop in pressure is visible in the cosine distributions near $\psi = 270°$ while it is not for the even distributions. Also, the pressure gradient near corner $\psi = 90°$ is resolved better by the cosine layout. Both these aspects are captured by the second mode which therefore is emphasized more for the cosine data as visible in the α-coefficients near $45°$. This causes the first cosine eigenvalue to be slightly smaller than for even and the second one to be larger. The rank two projection which can be formed using what is shown in Figure 5.10 describes the underlying trend in the training pressure distributions. High activity of the first mode near $\alpha = 0°$ and $90°$ as it resembles the pressure distributions near those angles of attack but the second mode takes over near $\alpha = 45°$ as the pressure distributions resemble the symmetric second mode. Relatively little pressure variation occurs on faces C and D over the angle of attack range, 'low energy' details on those faces are therefore mostly accounted for by higher order modes instead.

5.4. Experimental surface pressure reconstruction using GPOD

Two separate sets of eigenmodes are available, either obtained using the even or cosine tap layout. These have both been used for sparse reconstructions following the gappy POD procedure discussed in Section 2.5 but only results obtained using the even pressure distributions will be shown in this section. Reconstructions using the cosine pressure distributions are generally similar and are included in Appendix A. First, example reconstructions at certain angles of attack will be shown after which the reconstruction accuracy over all angles of attack and levels of gappiness are quantified in terms of RMSE.

5.4.1. Reconstruction examples

Sparse reconstructions will be performed on pressure distributions which have not been included in the training data as this will illustrate how well the training data represents square cylinder flow in general. Like the CFD GPOD reconstructions, distributions to be reconstructed are made artificially gappy through assigning sensor locations using the QR-CP algorithm of Section 2.7 and discarding the remaining surface pressure distribution. Reconstructions provided as examples below are those corresponding to $\alpha = 12.5°$ and $42.5°$. Both show a reconstruction performed using 5 sensors/bases (GPOD 1, $n_s = K = 5$) and using 10 sensors/bases (GPOD 2, $n_s = K = 10$). The reference provided in the background is the pressure distribution at the same angle of attack without any of the pressure measurements being discarded. It is therefore the non-sparse pressure distribution that should be approxi-

mated by GPOD.

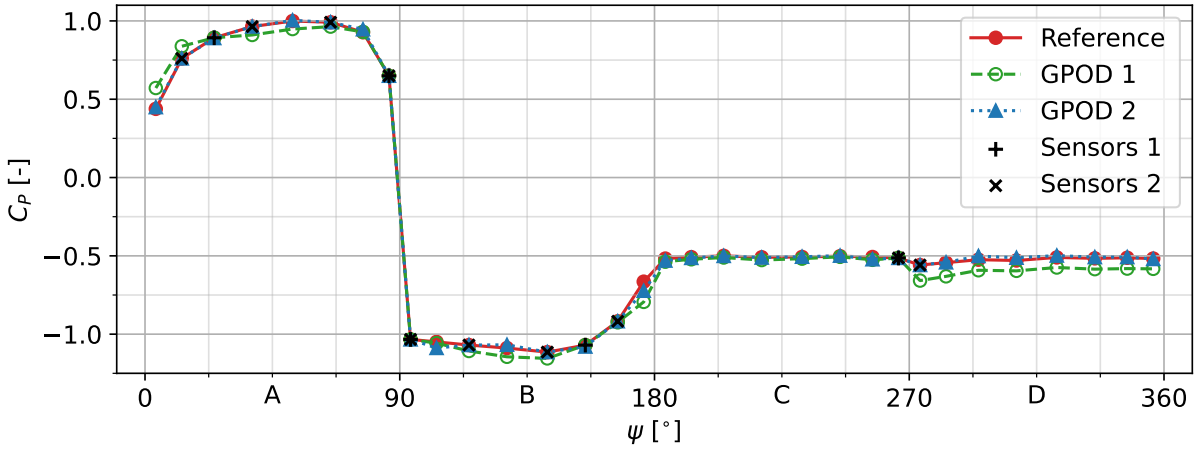


Figure 5.11: Experimental GPOD reconstruction at $\alpha = 12.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$

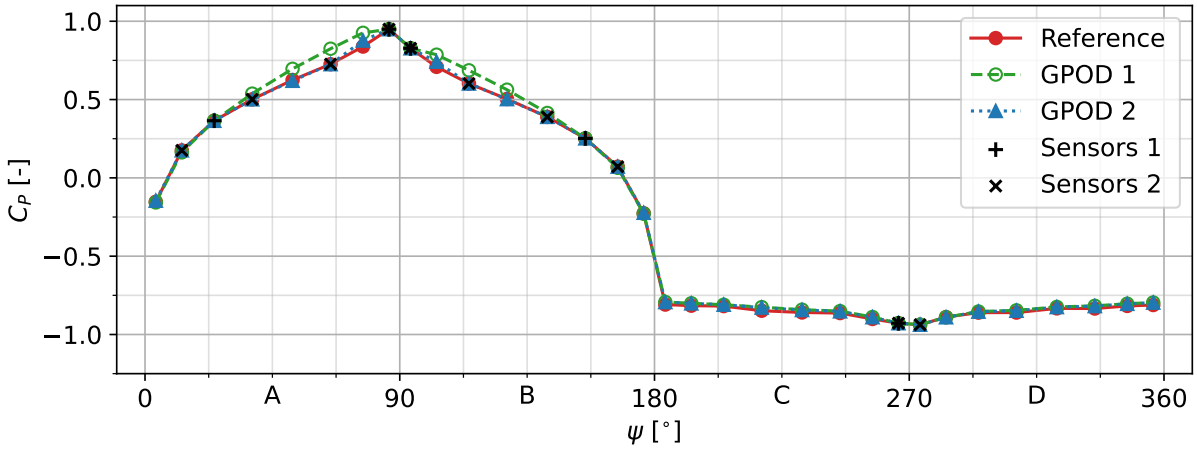


Figure 5.12: Experimental GPOD reconstruction at $\alpha = 42.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$

$\alpha = 12.5°$ and $42.5°$ correspond to pressure distributions at minimum and maximum drag respectively outside of the training data. The gappiness is roughly 84 % and 69 % using 5 and 10 sensors respectively. The GPOD 1 reconstruction in Figure 5.11 shows already great similarity with the reference. The effect of the pinched recirculation bubble on face B is predicted well even though the pressure distribution at $\alpha = 12.5°$ is quite unique. Faces A and B contain each two out of five sensors and the remaining sensor is placed at face C closest to $\psi = 270°$. This is a result of the QR-CP algorithm favoring the placement of sensors in high modal activity regions which especially for the first few modes are mainly on faces A and B. The single sensor on face C is enough to predict the surface pressure across the face however using 5 modes, apparently no modes with sufficient (anti)correlation to face D are used as to simultaneously predict pressure on it accurately, resulting in constant under prediction. GPOD 2 using $n_s = K = 10$ uses the same sensor on face C however adds a sensor close to it on face D. Together with the additional modes, this results in much closer reconstructions

to the reference on faces C and D. 8 out of 10 sensors for GPOD 2 are placed on faces A and B which also here yield an improvement upon GPOD 1. RMSE compared to the reference are $6.0 \cdot 10^{-2}$ for GPOD 1 and $1.8 \cdot 10^{-2}$ for GPOD 2. Although the increase in RMSE from $n_s = K = 5$ to $n_s = K = 10$ is quite significant, increasing $n_s = K$ beyond 10 increases the reconstruction accuracy at this specific angle of attack remarkably little (RMSE at $n_s = K = 19$ is $1.4 \cdot 10^{-2}$, see also Figure 5.14 to be discussed later). This indicates that modes above 10 are not informative towards reconstructions, possibly due to this angle of attack being underresolved by the training data.

The pressure distribution at $\alpha = 42.5°$ resembles the second POD mode in Figure 5.10 and is therefore already approximated well by the low order GPOD 1. The two sensors close to $\psi = 90°$ measure the slight asymmetry which is also incorporated into the reconstruction. Unlike the previous reconstruction example, noticeable differences between reference and GPOD 1 are mainly present on faces A and B while the single sensor close to $\psi = 270°$ seems to provide enough information to accurately reconstruct pressure over faces C and D. The reconstruction on faces A and B does not seem unphysical however does not capture smaller changes in slope which are represented in higher order modes. The additional modes and sensors used for GPOD 2 mainly help in improving upon the mistakes on faces A and B while the added sensor on face D does not seem to alter the reconstruction on C and D as it was already close to the reference. The RMSE to the reference for GPOD 1 is $4.0 \cdot 10^{-2}$ and for GPOD again $1.4 \cdot 10^{-2}$.

Especially for GPOD 1 using $n_s = K = 5$, the examples shown above are amongst the better reconstructions in terms of RMSE (see Figure 5.14, to be discussed later). This seemingly obscures aspects of the pressure distributions at some angles of attack with which GPOD struggles in reconstructing. One such pressure distribution at which the low-rank GPOD 1 reconstruction clearly suffers from only five POD modes and sensors being available is at $\alpha = 32.5°$ of which the reconstruction using both $n_s = K = 5$ and 10 is shown in Figure 5.13.

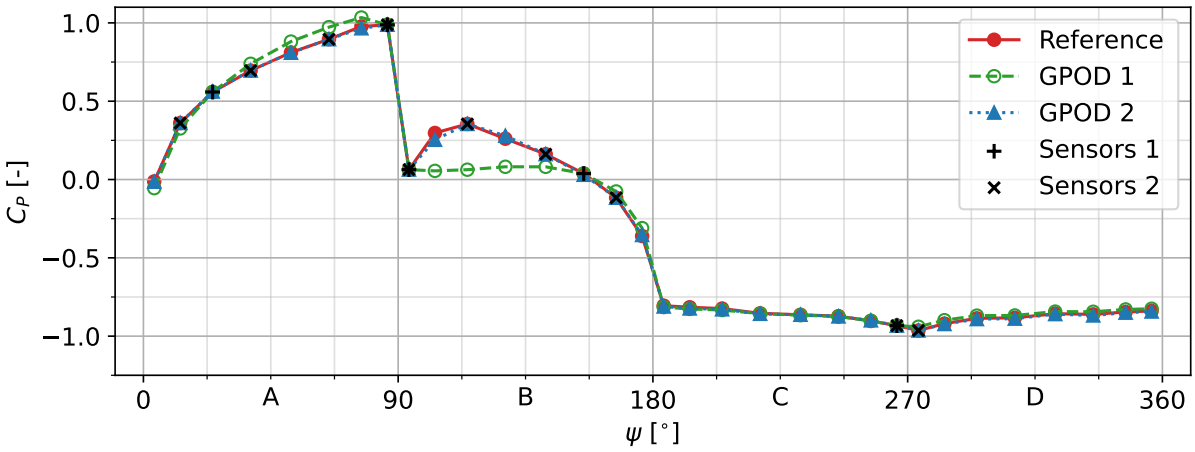


Figure 5.13: Experimental GPOD reconstruction at $\alpha = 32.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$

With an RMSE of $8.8 \cdot 10^{-2}$, the GPOD 1 reconstruction shown above is the least ac-

curate out of all reconstructions on the the non-training angles of attack. Within the supercritical angle of attack range, the pressure on faces C and D are consistently similar in shape and magnitude hence the reconstruction is able to remain close to the reference using only the single sensor near $\psi = 270°$. On face A, differences between GPOD 1 and the reference are clearly visible but overall the reconstruction follows the shape of the reference. On face B, the GPOD 1 reconstruction shows the largest errors. In this angle of attack range, the flow reattachment produces a pressure peak located in between the GPOD 1 sensor locations of $\psi = 94°$ and $155°$. For the the pressure distribution at face B to be reconstructed accurately, POD modes need to provide information regarding this reattachment spanning the gap between the sensors. This information is not contained in the first five modes (see Appendix B for spatial POD modes) however hence the peak is not reproduced. The additional modes added for GPOD 2 using $n_s = K = 10$ contain high modal activity in the form of peaks over faces A and B which are used to reproduce the growing region of reattached flow. GPOD 2 therefore receives additional sensors on these faces and is able to predict the pressure distribution in the supercritical flow regime much more accurate. The GPOD 2 reconstruction achieved an RMSE of $1.1 \cdot 10^{-2}$ which is instead the smallest out of all non-training angles of attack.

Interesting to see is how despite the large differences between the experimental and numerical surface pressure distributions, the sensor locations using $n_s = K = 10$ are similar. For both experimental and numerical results, 8 out of 10 sensors are placed symmetrically on faces A and B and 2 sensors are placed close to $\psi = 270°$. These locations make physical sense considering the transition between separated and attached flow over faces A and B while faces C and D remain fully separated throughout the angle of attack range. The sensor locations are determined purely from the POD modes. The first two modes from both experimental and numerical surface pressure distributions are similar in shape however especially the higher order modes differ significantly. This does not seem to have affected the identification of the most informative regions for sparse sensor placement.

5.4.2. Pressure reconstruction accuracy

Reconstructions like those shown in Section 5.4.1 are performed for all combinations of angle of attack and number of sensors/modes. The RMSE to corresponding reference pressure distributions for all of these reconstructions are shown in Figure 5.14.

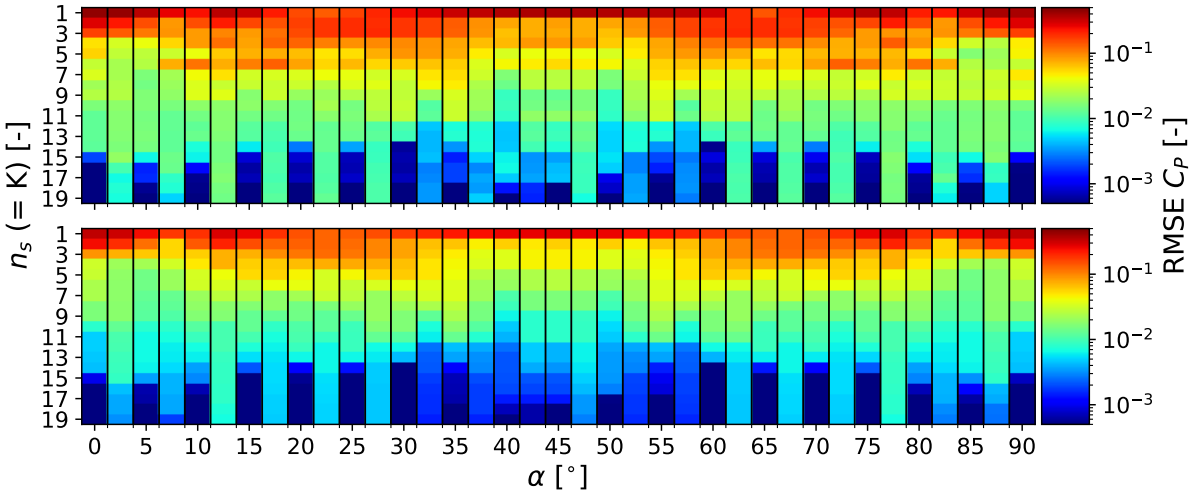


Figure 5.14: Experimental GPOD pressure distribution reconstruction (top) and projection (bottom) accuracy for every combination of $n_s = K$ and α

The upper colormap shows the reconstruction RMSE and the bottom one the projection RMSE. POD projections are again obtained as discussed in Section 4.5.2 and provide a best attainable benchhmark for GPOD reconstructions. Angles of attack included in training data are indicated with ticklabels on the x-axis but are of lesser interest for the reconstruction performance as they are per definition better represented by the eigenmodes. The reconstruction colormap shows how a certain number of modes and sensors might be better suited for reconstructions on specific angles of attack while performing worse on others. This variance in reconstruction accuracy over the angles of attack is most pronounced for smaller values of $n_s = K$ where the small number of sensors can not be used to account for all of the flow phenomena in combination with the available modes. $n_s = K = 5$ seems to significantly decrease the RMSE across the angles of attack with respect to $n_s = K = 4$ however an unexpected step back in terms of RMSE occurs when using $n_s = K = 6$ which will be discussed further below Figure 5.15. For $12.5° \leq \alpha \leq 27.5°$ and $62.5° \leq \alpha \leq 77.5°$, reconstruction RMSE's seem to decrease by only a small amount for $n_s = K > 10$. This was noticed already in Section 5.4.1 and could be the result of sensor placement being unsuited for these angles of attack. If this were the case it would however only affect reconstructions but the same trend is even more clearly visible in the projection RMSE's. Instead, the training data does not seem to provide eigenmodes capable of decreasing the RMSE significantly for these angles of attack beyond a certain $n_s = K$ which is a sign of more training data being necessary in the local angle of attack range.

Figure 5.15 provides a more quantitative way of interpreting the accuracy of GPOD reconstructions for a certain number of sensors and bases used. It is obtained by computing the RMSE to reference pressure distributions across all non-training angles of attack at once for $n_s = K = 1$ to 19. This is done for both the reconstructions and POD projections which makes this figure the experimental counterpart to the numerical results in Figure 4.11.

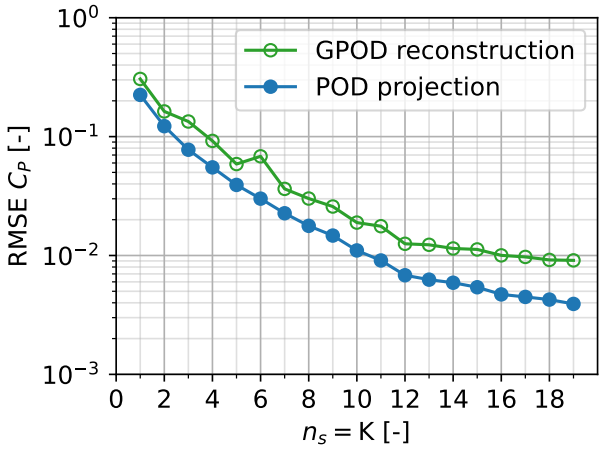


Figure 5.15: Experimental GPOD surface pressure reconstruction accuracy for increasing number of modes and sensors

The projection curve decreases for each additional mode used, the rate of which decreases with each new mode introducing less new information. The experimental projection curve achieves lower RMSE's across the $n_s = K$ range compared to the numerical results implying that the experimental modes form a better square cylinder flow library compared to CFD. The difference between GPOD reconstructions and POD projections is small and generally decreases for increasing $n_s = K$. Since POD projections are essentially GPOD reconstructions using $n_s = n \neq K$, this difference in RMSE can be attributed to accuracy loss due to gappiness. The QR algorithm successfully keeps this loss small by providing sensor locations which are informative over all modes involved in a reconstruction but independent from one another as to keep modes orthogonal. The greediness of the QR algorithm results in a sensor layout obviously far from optimal at $n_s = K = 6$ where all reconstructed pressure distributions produce relatively large RMSE's making this point clearly stick out in Figure 5.15. The sensor locations for all $n_s = K$ are shown schematically in Figure 5.16 from which it is apparent why the reconstructions at $n_s = K = 6$ fail to stick to the RMSE trend.

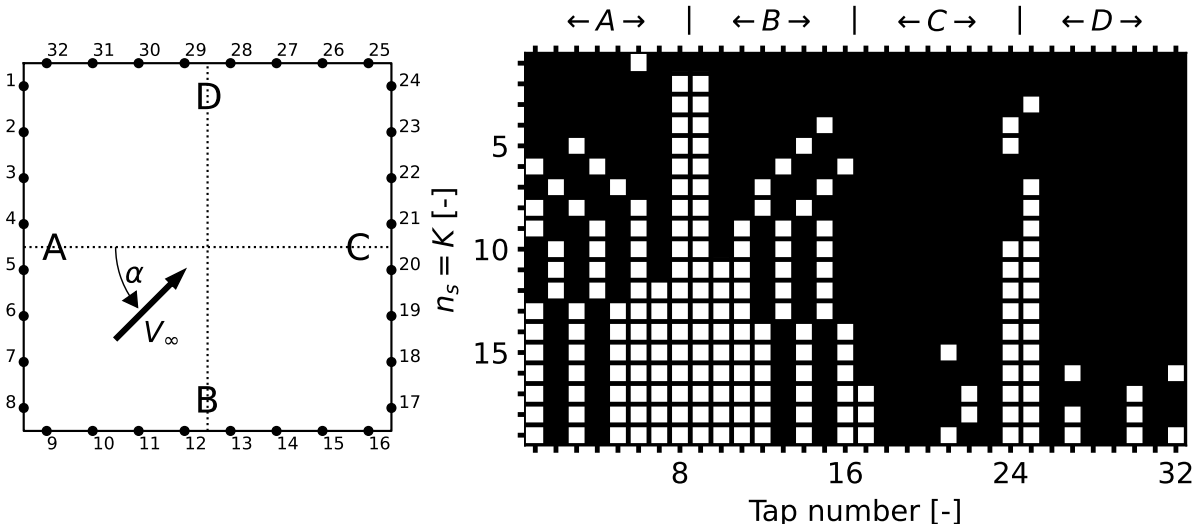


Figure 5.16: Sensor locations determined by QR algorithm for each value of $n_s = K$

The right part of the figure above shows what taps on the square cylinder surface are determined by the QR algorithm to be suited sensor locations given a certain sensor budget and number of POD modes. The left part provides a more easily interpretable view of where the taps on the cylinder surface are positioned. Despite relatively little variance in the pressure distribution on faces C and D, inclusion of the third mode at $n_s = K = 3$ already introduces a sensor on the downstream side of the cylinder near the corner of faces C and D. This sensor 'guides' the reconstructions over all angles of attack on faces C and D for $n_s = K = 3$, 4 and 5 even though it switches from face D to C. At $n_s = K = 6$, no sensor is placed on either face C or D from which the reconstructed pressure distribution on these faces suffers and yields an increase in RMSE. At $n_s = K = 7$, again a sensor is placed on face D and the RMSE can be seen to follow the general trend again. Manually repositioning a sensor at $n_s = K = 6$ to tap 24 or 25 was found to result in a lower overall RMSE confirming the QR sensor placement was not optimal for this situation.

5.5. Experimental drag estimation using GPOD

Approximations for the pressure drag are obtained by integrating the previously discussed surface pressure distribution reconstructions. These sparse drag approximations will be discussed in the following sections. First, two examples of reconstructed drag curves using different numbers of sensors are shown. After this, the overall drag estimation accuracy is quantified in terms of RMSE for different levels of sparseness used during the surface pressure reconstructions.

5.5.1. Drag estimation examples

Figure 5.17 shows a reference drag curve and two GPOD approximations. The reference curve is obtained by integrating the complete experimental surface pressure distributions obtained using the even tap layout. Drag curves for GPOD 1 and GPOD 2 are obtained by integration of the sparse reconstructions of the same experimental data using $n_s = K = 5$ and 10 respectively and should therefore approximate the reference. Angles of attack marked with a ticklabel on the x-axis are those for which the corresponding surface pressure distributions were included in the training data.

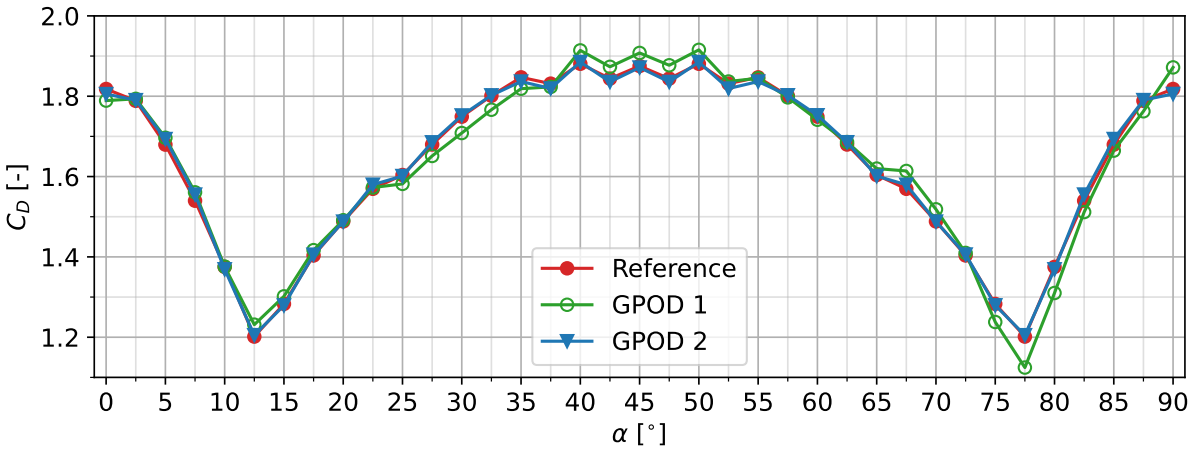


Figure 5.17: Experimental GPOD drag curve reconstruction. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$

Already using $n_s = K = 5$ for GPOD 1, the reference drag curve is approximated quite well. Features such as the angles at which minimum/maximum drag is achieved and also the slopes of the drag curve are all reproduced as the reconstructions using 5 sensors were seen earlier to capture the essence of pressure distributions well. The sensor layout for GPOD 1 is not symmetric about the cylinder diagonal as an uneven number of sensors are used. This affects the reconstructions differently depending on the angle of attack and therefore the GPOD 1 drag curve is not symmetric about $\alpha = 45°$. For small angles of attack, most of the drag force is a result of differences in pressure between faces A and C. For angles of attack closer to $\alpha = 90°$ it is instead mostly the pressure differences between faces B and D. Since the only sensor above $\psi = 180°$ is placed on face C while faces A and B both contain two sensors, the important regions of the pressure distributions for drag determination are reconstructed better for smaller angles of attack (close to $\alpha = 0°$). Therefore the GPOD 1 drag coefficients for smaller angles of attack are more accurate than for large angles of attack (close to $\alpha = 90°$). The growing region of increased pressure on face B resulting from flow reattachment does not seem to be encoded well into the first five POD modes as it is underpredicted for $25° \leq \alpha \leq 35°$ for which drag is also underpredicted as a result (see also Figure **??**. A similar partly reattached type flow but instead on face A occurs for $55° \leq \alpha \leq 65°$ which is also similarly underpredicted. This does not influence the drag approximations however as it tends to cancel with the errors made on face D due to the lack of a sensor. The overpredicted drag around $\alpha = 45°$ is mainly a result of the reconstruction pressure being slightly too large on faces A and B and not so much from errors on faces C and D.

Using $n_s = K = 10$ for GPOD 2 improves upon the drag estimation at all angles of attack compared to GPOD 1. The GPOD 2 drag curve is almost indistinguishable from the reference and also symmetric about $\alpha = 45°$ due to the symmetric sensor layout. No noticeable angle of attack ranges are present with abnormalities in the drag estimation meaning that all dominant surface pressure features are incorporated into the 10 used modes and are sampled effectively by the 10 sensors. The increase in accuracy is for a large part the result of the inclusion of the 10th mode which has high modal activity on faces C and D. This mode can thus be used to correct reconstructions on faces C and D on which reconstruction offsets to the reference are often responsible for drag prediction errors, but it also results in a second sensor being placed above $\psi = 180°$ on face C (see Figure 5.16) to make it possible to constrain pressure on faces C and D separately.

5.5.2. Drag estimation accuracy

The accuracy of the GPOD drag predictions is quantified by computing the RMSE between the drag curve reconstructions and the reference over all the non-training angles of attack only. This is done for $n_s = K = 1$ to 19 and also for the drag curves obtained using the POD projections. The results are shown in Figure 5.18.

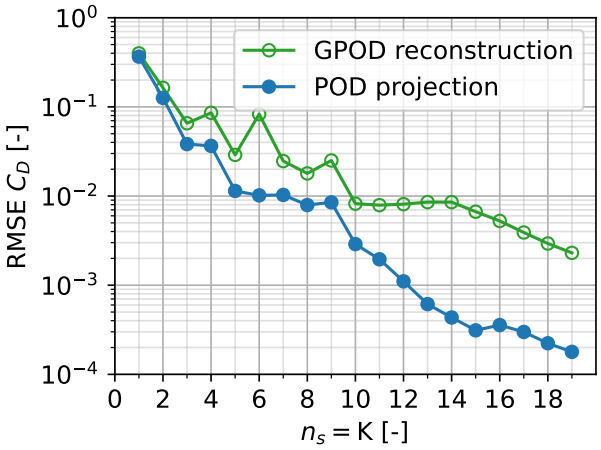


Figure 5.18: Experimental GPOD drag curve reconstruction accuracy for increasing number of modes and sensors

Similar to the CFD results, the behaviour of both the projection and reconstruction RMSE curves is much more irregular than for the surface pressure reconstructions. This is again caused by the surface pressure integration required for the computation of the drag coefficient making it possible for small errors in surface pressure reconstructions to accumulate. Whereas for the surface pressure reconstructions increasing $n_s = K$ generally decreases RMSE C_p, this does not necessarily translate into better approximations of the drag coefficients. This also means that the POD projections do not necessarily produce a lower RMSE in terms of C_D. Unlike the surface pressure reconstruction accuracy, the drag estimation depends greatly on where around the cylinder errors in the reconstructions are made and therefore especially relies on the sensor placement shown in Figure 5.16. The lack of sensors on faces C and D at $n_s = K = 6$ yields poor reconstructions on those faces of which the errors accumulate and result in exceptionally bad drag prediction accuracy. Even using a sensor less at $n_s = K = 5$, the single sensor placed on face C is able to allow for better drag estimates. At $n_s = K = 7$, 8 and 9, no additional sensors are placed on faces C and D. The RMSE in terms of C_p decreases but as most of the additional eigenmodes included also focus mainly on faces A and B, no considerably decrease in C_D RMSE occurs. As mentioned earlier, the 10th POD mode has high modal activity on faces C and D. This enables this mode to be used for better reconstructions on those faces but also results in a new sensor layout such that both faces C and D contain a separate sensor. This decreases the C_p RMSE conform the general trend of Figure 5.15 but has a larger impact on the drag prediction accuracy as the C_D RMSE drops by over a factor of 3 from $n_s = K = 9$ to 10. POD modes 11 to 14 again focus on faces A and B and thus the QR algorithm places the additional sensors there as well. Still, these modes seem to contain the information to enhance drag prediction accuracy as the projection curve remains decreasing. This illustrates how the GPOD- in combination with the QR algorithm focus on decreasing the C_p RMSE while better drag predictions could possibly be achieved otherwise. This seems to have been realized by the earlier referenced work by Mifsud et al. (2019) as an adapted least squares formulation in the GPOD procedure was used to produce reconstructions in better accordance with integrated aerodynamic coefficients. Purely to increase the accuracy with which drag

is predicted, increasing the number of sensors above 10 results in diminishing returns. Either using 5 or 10 sensors seems to be optimal for providing physically meaningful surface pressure reconstructions which yield drag predictions in accordance with the true measured pressure drag.

5.6. GPOD error sources

During the discussion of the results, several aspects which affect the accuracy of GPOD reconstructions have been identified. In general, the effect of these aspects on the GPOD accuracy are most easily identified through the surface pressure reconstruction accuracy in terms of RMSE in the C_p as in the drag estimation accuracy they are obscured by the accumulation of errors. Observed errors in the RMSE C_p plots seem to originate from three 'sources of error' which are schematically depicted in Figure 5.19. Note that even though the sources of error are discussed separately, almost always a combination of all is responsible for the observed RMSE C_p in surface pressure reconstructions.

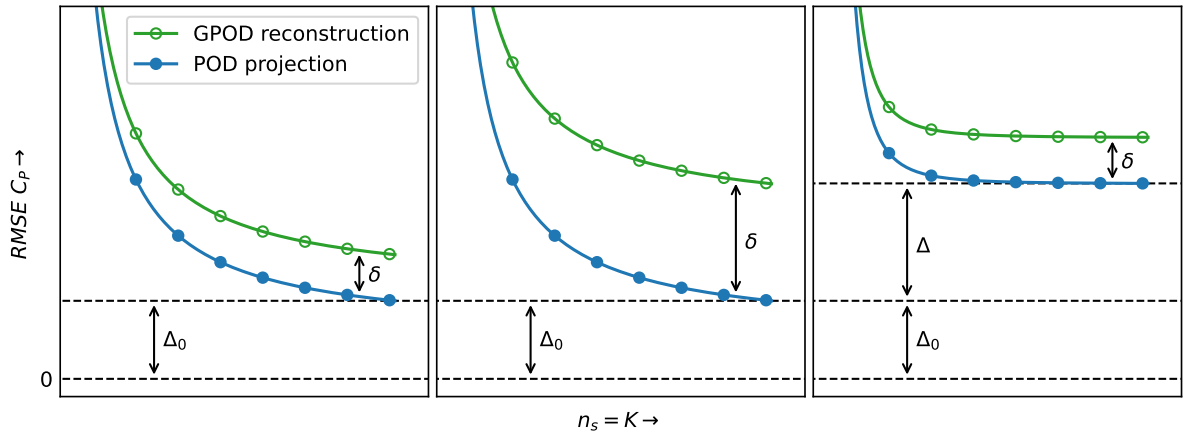


Figure 5.19: Schematic RMSE C_p plots representing sources of error: overall training inadequacy (left), sensor placement (middle) and training data skewness (right)

The left figure depicts a schematic projection and reconstruction RMSE C_p curve similar to those in Figures. 4.11 and 5.15 which represented the RMSE across all non-training angles of attack for certain $n_s = K$. This figure illustrates the presence of an inherent offset in RMSE C_p from 0, denoted as Δ_0. This offset is defined with respect to the last point of the POD projection curve and illustrates to what extent the provided training data is adequate for reconstructions performed. Hence, the error source Δ_0 is referred to as the training inadequacy. At the last point of the projection curve, all available modes are used ($K = m$) and as discussed in Section 4.5.2 the projections are performed using the GPOD procedure without any missing elements in the modes ($n_s = n$). This yields optimal least squares approximations using a linear combination of all modes. Since the modes were obtained from training data, rank m projections on training data thus reproduces the data exactly. Rank m projections of data not included in training instead results in non-zero RMSE's as features not present in the training data are bound to occur. Over all reconstructed non-training angles of attack, this RMSE difference from 0 (Δ_0) thus quantifies how effective the training has been

in encoding the features of a certain use case in the space spanned by the training data. Since GPOD reconstructions approach the projections as n_s approaches n, it marks a region unattainable using sparse reconstructions given certain training data. The rate at which Δ_0 is reached by the projections depends on the relative importance of the eigenmodes and therefore on the 'low-rankness' of the use case.

The middle figure exaggerates the difference between GPOD reconstructions and the POD projections which is denoted by δ. Both projection and reconstruction use the same $K \leq m$ POD modes but reconstructions need to overcome sparseness as well. δ can thus be seen as an 'accuracy loss due to gappiness' which the chosen sensor placement algorithm should minimize given a value for $n_s = K$. Overall, the size of δ is mainly affected by the sensor placement algorithm. This is apparent from Figure 2.20 when comparing the random sensor placement performance with that of QR but also in the results of Jayaraman et al. (2019) for example where multiple algorithms are compared (although primarily in the oversampled case $n_s > K$). δ is not constant for each value of $n_s = K$ however as most sensor placement algorithms, like QR-CP, provide a greedy sensor layout solution to an otherwise intractable combinatorial search over all sensor layout possibilities. An absolute optimal sensor configuration is therefore not guaranteed and sub-optimal configurations can lead to outliers in terms of reconstruction accuracy as discussed in Section 5.4.2 regarding Figure 5.15. Still, optimal sensor placement alone would not be able to bring δ to zero as the least squares fit of K modes through $n_s < n$ sensor measurements will always have larger residual to the reference than a rank K projection using $n_s = n$ would have.

The right figure schematically illustrates the trend in RMSE C_p for increasing $n_s = K$ at individual angles of attack only. This is different from the left figure which assumed a single RMSE C_p value for reconstructions over all angles of attack at a certain $n_s = K$. Plots like the right of Figure 5.19 would be obtained when focusing on single columns of Figure 5.14. Even though an 'overall reconstruction performance' plot as shown on the left is not the average over the reconstruction performance at individual angles of attack (order in which mean and square are calculated in RMSE matters), the distance Δ, defined from Δ_0 to the projection at $n_s = K = m$, gives insight into how well pressure distributions at single angles of attack are resolved by the training data and therefore how 'skewed' the training data might be. As illustrated, Δ can stack with Δ_0 to indicate that at the specific angle of attack; training data is not very representative and additional eigenmodes from a certain point onwards do not enhance projection nor reconstruction significantly. An example of this was discussed in Section 5.4.2 where such behaviour was noticed for specific angles of attack in the supercritical flow regime. Conversely, for reconstructions at other angles of attack, Δ might be 'negative' in the sense that it decreases RMSE with respect to Δ_0. The eigenmodes from training data seem to be skewed towards pressure distributions at those angles of attack instead. Examples of this would be the columns in Figure 5.14 corresponding to angles of attack close to $\alpha = 45°$ which can be seen to remain steadily decreasing in RMSE C_p for increasing $n_s = K$. Large skewness of the training data can lead to increased differences between reconstruction performance at different angles of attack. To counter this, one could locally refine training data with additional training snapshots

as to resolve the flow phenomena better. Additionally, one could redistribute the angles of attack at which training snapshots are provided to focus on difficult to resolve flow phenomena (e.g. expected onset of stall for an airfoil at angle of attack case).

5.7. GPOD versus linear interpolation

Accuracy of GPOD surface pressure distribution- or drag curve reconstructions have been quantified in terms of RMSE with respect to a known reference distribution. This gives insight into the absolute performance of the method but not how it compares to alternatives. One such alternative is linear interpolation (LERP) which despite its limited accuracy is very robust, computationally inexpensive and simple to implement (Raben et al., 2012). Linear interpolation requires local spatial information only to fill in missing measurements and is therefore non-data-driven. GPOD on the other hand identifies spatial and 'temporal' (depends on the changing variable(s) throughout training snapshots, angle of attack in this case) structures from training data and uses this for what effectively is a data-driven interpolation technique. Linear interpolation has been implemented for reconstructions of the same surface pressure distributions and the achieved accuracy will be compared to the earlier shown GPOD results in the following sections.

5.7.1. Linear interpolation reconstruction examples

LERP reconstructions are performed using a number of sensors n_s from 1 to 19 at all available angles of attack. The sensor locations which are used as the interpolation points are kept equal to those used for GPOD reconstructions which are shown in Figure 5.16. To avoid completely unphysical LERP reconstructions, each face of the cylinder is interpolated separately using the available sensors on that face. In case no sensors on a face are available, the pressure measured by the closest sensor on a different face is used. In the examples below, the GPOD reconstructions using $n_s = K = 5$ and $n_s = K = 10$ are compared to LERP reconstructions using the same number of sensors at $\alpha = 17.5°$.

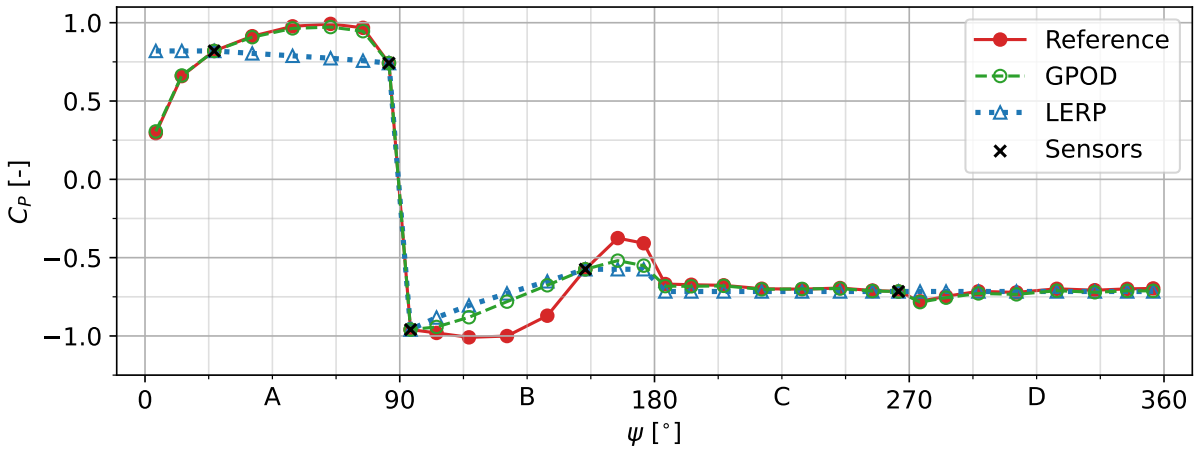


Figure 5.20: Experimental GPOD and LERP reconstruction at $\alpha = 17.5°$, $n_s = K = 5$.

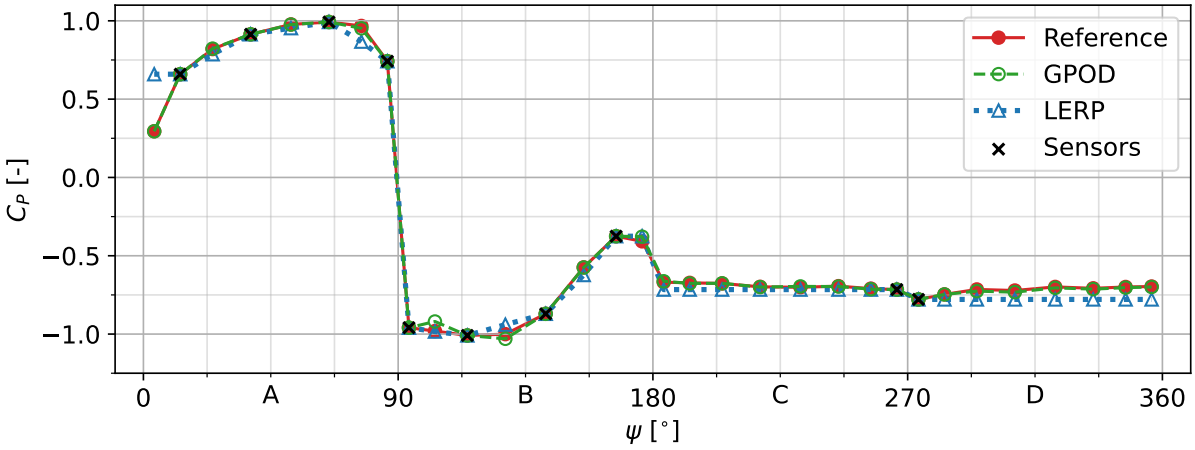


Figure 5.21: Experimental GPOD and LERP reconstruction at $\alpha = 17.5°$, $n_s = K = 10$.

Figure 5.20 shows the LERP and GPOD reconstructions using $n_s = 5$. The QR algorithm places two sensors each on faces A and B which experience the most variation in pressure at different angles of attack. This is clearly not sufficient for linear interpolation to be able to follow the reference as significant under- and overprediction by the LERP reconstruction is visible. On face A, the GPOD reconstruction instead is able to follow the reference almost perfectly as the fully attached flow is easily characterized in the first few POD modes. On face B, the GPOD reconstruction struggles with the growing region of reattached flow and because of this, GPOD reconstructions at $\alpha = 17.5°$ tend to have relatively large RMSE C_p. Regardless, GPOD is able to predict the surface pressure at each pressure tap on face B closer to the reference compared to LERP as it captures some of the curvature of the pressure distribution. Also on faces C and D the GPOD reconstruction is closer to the reference but the generally simple behaviour of the pressure distribution here lends itself well for an approximation by a constant pressure as obtained using the single sensor for LERP. The LERP reconstruction using $n_s = 5$ achieves a RMSE C_p to the reference of $1.6 \cdot 10^{-1}$ while for GPOD this is $7.4 \cdot 10^{-2}$.

Using 10 sensors as shown in Figure 5.21, faces A and B both receive four sensors each. The GPOD reconstruction on face A was already accurate using five sensors total but becomes slightly better (in the non sensor locations) through the additional modes and constraints of the sensors. The LERP reconstruction improves greatly but is still unable to capture curvature of the concave pressure distribution yielding small errors and is very far from the reference in the first tap as an artefact of the linear interpolation implementation. The reattachment on face B is able to be reproduced by both GPOD and LERP as 4 out of the 8 sensors on the side are available. Faces C and D were also reconstructed close to the reference by GPOD using 5 sensors total already and the improvement using 10 sensors instead is therefore almost unnoticeable by eye. The effect on GPOD reconstructions of placing a sensor on face D is mainly visible on angles of attack closer to $\alpha = 0°$ and $45°$. The additional sensor on face D actually hurts the LERP reconstruction accuracy as the face D sensor is placed in a local dip in pressure which is assumed to be the pressure across the whole face. For LERP reconstructions, the additional sensor on face D is mainly effective in the sub-

critical regime where pressure between faces C and D differ but is counterproductive for other angles of attack instead. The LERP reconstruction using $n_s = 10$ achieves a RMSE of $9.3 \cdot 10^{-2}$ to the reference and for GPOD this is $1.7 \cdot 10^{-2}$, demonstrating how the additional sensors w.r.t. $n_s = 5$ benefit GPOD more.

5.7.2. Comparison of surface pressure reconstruction accuracy

To gain insight into the overall performance of LERP versus GPOD, reconstructions such as shown in the previous section are constructed for all angles of attack for which the number of sensors is varied from $n_s = 1$ to 19 each. For each value of n_s, a single overall RMSE to the reference for the LERP reconstructions is calculated. Only reconstructed surface pressure values different from the supplied sensor measurements are considered as well as only considering the angles of attack which were used to quantify the GPOD reconstruction accuracy (non-training angles of attack, see Section 5.3). The LERP results are shown together with the GPOD results in Figure 5.22 below.

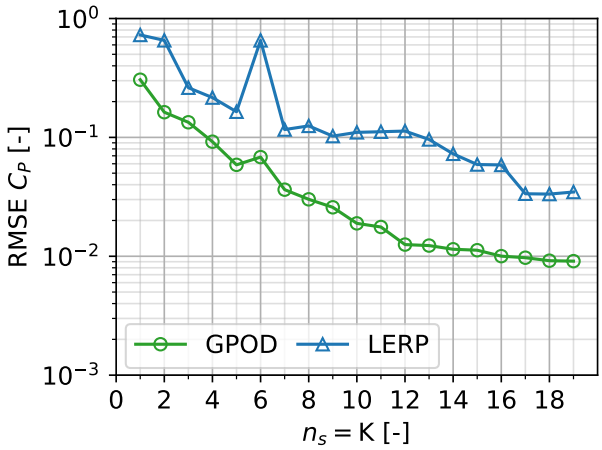


Figure 5.22: Experimental GPOD and linear interpolation reconstruction accuracy

For all values of n_s, the LERP RMSE C_p values are larger than GPOD. Overall, the LERP RMSE decreases for increasing n_s but not as consistently as for the GPOD results. For a small number of sensors, each additional sensor rapidly decreases the LERP RMSE as expected until $n_s = 6$. At $n_s = 6$ a jump in the RMSE is visible as was also the case for the GPOD results. In Section 5.4.2 this was explained to be caused by the QR algorithm not assigning any sensors on both faces C and D even though at $n_s = 5$ a sensor was placed on face C. While this certainly negatively affects the GPOD reconstructions, the POD modes allow for correlation between pressure on faces A, B and C, D to be utilized in predicting a relatively accurate surface pressure regardless. Linear interpolation, relying on local surface pressure measurements only instead, suffers greatly in terms of reconstruction accuracy compared to GPOD with two out of the four faces providing no information through sensors at all. With a sensor returning on face D at $n_s = 7$, the LERP RMSE falls in line with the previous trend again but remains stagnant onwards as additional sensors are placed on faces A and B while the errors mainly originate from the single sensor on face C and D combined. GPOD remains steadily decreasing for increasing the number of sensors beyond 7 and achieves RMSE values around $n_s = 10$ which are almost a full order of magnitude

smaller than for LERP. For very sparse reconstructions (e.g. $n_s = 5$), GPOD thus greatly improves upon LERP in an absolute sense while the decreasing RMSE trend for both is similar but since GPOD remains steadily decreasing in RMSE for increasing n_s while LERP does not, the relative accuracy gain using GPOD increases.

5.7.3. Comparison of drag estimation accuracy

Similar to GPOD, LERP reconstructions fill in the missing surface pressure measurements such that a reconstruction defined at the $n = 32$ tap locations is produced which can be integrated for a drag coefficient approximation. Such LERP drag approximations are performed for all angles of attack and at $n_s = 1$ to 19. For each value of n_s, the drag estimation accuracy is quantified as the RMSE between the LERP and reference C_D vs α curve while only considering the non-training angles of attack as was also done for the GPOD results. The LERP and GPOD drag estimation accuracy is shown in Figure 5.23.

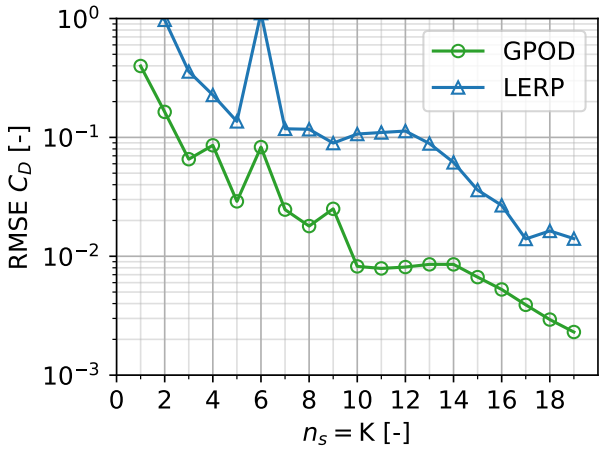


Figure 5.23: Experimental GPOD and linear interpolation drag curve reconstruction accuracy

Unlike the RMSE in terms of C_p, the RMSE C_D is also influenced by the location around the cylinder at which errors in a surface pressure reconstruction are made. Especially at small n_s, each additional POD mode can significantly change the GPOD reconstruction leading to reconstruction errors at different locations which either accumulate or cancel in the integration of the pressure distribution. This leads to the erratic behaviour seen in the GPOD results below $n_s = K = 10$ but the LERP reconstructions, not relying on a set of modes, are less sensitive to this and instead decrease in RMSE C_D until $n_s = 9$ (ignoring the outlier at $n_s = 6$ which is caused by reasons already discussed in the context of Figure 5.22). From $n_s = 9$ to 10, the addition of a second sensor on faces C and D combined actually results in a slightly worse reconstruction on faces C and D for the majority of angles of attack using LERP, see Figures. 5.20 and 5.21. This affects the drag estimation as well as the RMSE C_D can be seen to have increased. At $n_s = 11$ and 12 the additional sensors being placed at faces A and B do not solve this problem and the drag curve approximations remain at a similar RMSE C_D which are more than a factor 10 larger than for GPOD. From $n_s = 13$ onwards, a sensor is first placed on tap 1 which eliminates the linear interpolation artefact shown in Figure 5.21 and additional sensors are placed on faces C

and D which decrease the RMSE. At this point however, the number of supplied surface pressure measurements for both LERP and GPOD being sparse is questionable.

5.8. CFD based sensors GPOD reconstructions

In the previous sections, both the training for POD modes and the sparse sensor placement are obtained from the experimental surface pressure distributions. In the following section, POD modes will still be obtained from experimental surface pressure distributions but the sensor placement will instead be based on the QR algorithm applied to the CFD results discussed in Section 4.3. The sensor layouts obtained using the CFD results are referred to as Opt V1 and Opt V2 and were mentioned earlier in Section 4.7.2 and shown in Figure 3.4. Experimental surface pressure measurements using these tap layouts were shown in Figure 5.4 and 5.5. The small number of taps on the Opt V1 and V2 pressure tap plates does not allow for accurate determination of the drag through spatial integration of the surface pressure measurements (see LERP results Section 5.7.3). A GPOD reconstruction however can be used to fill in the gaps in the pressure distribution and increase the spatial resolution if a suitable set of POD bases is available. Unfortunately, due to an oversight also discussed in Section 4.7.2, only the taps on the Opt V2 plates for the 10 sensors and bases layout can be used for GPOD reconstructions.

A set of representative POD bases is readily available from either the even- or cosine taps surface pressure distributions but both do not contain taps at the spatial locations of taps present on the Opt V2 plates, hence the POD modes do not either. The POD modes can be interpolated to support modal values at the positions of the Opt V2 taps but this destroys orthogonality. Instead, the surface pressure distributions themselves are interpolated to contain surface pressure values at the additional locations present on the Opt V2 plates. These interpolated measurements serve as placeholders to carry through to the POD modes while keeping orthogonality but will be replaced by the actual experimental measurements for the reconstructions. This approach introduces unwanted effects however which will be discussed later.

5.8.1. Opt V2 reconstruction examples

Since the cosine spaced pressure plates contain taps closer to the cylinder edges than any of the taps on the Opt V2 plates, it is used to interpolate the additional datapoints in the pressure distributions and obtain the POD modes. These POD modes are not used for the QR algorithm to determine sensor locations as the sensor locations will be fixed on the Opt V2 tap locations where experimental measurements are available. GPOD reconstructions can be performed as usual by using the Opt V2 experimental measurements on the tap locations instead of the interpolated placeholders. Note that with the merging of cosine and Opt V2 experimental measurements, the total number of 'taps' has become $n = 42$ instead of $n = 32$. Pressure distributions at the same angles of attack as before, being integer multiples of $5°$, are used as training data for the modes. Since the training data and sparse measurements on the sensor locations have been obtained independently in this case, differences between reconstructions

on training versus non-training angles of attack are smaller and thus no distinction is made in showing the results.

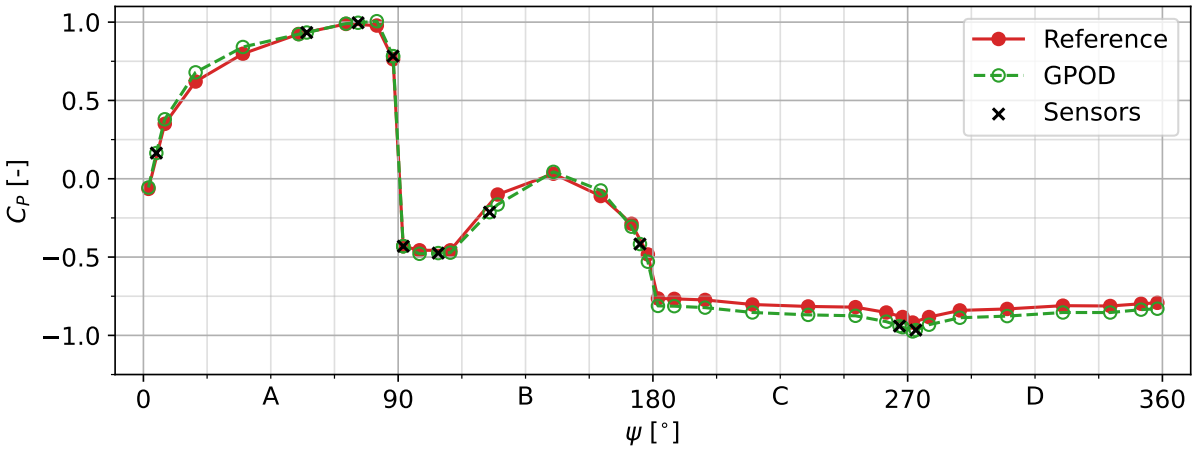


Figure 5.24: Opt V2 GPOD reconstruction at α = 25°, $n_s = K = 10$

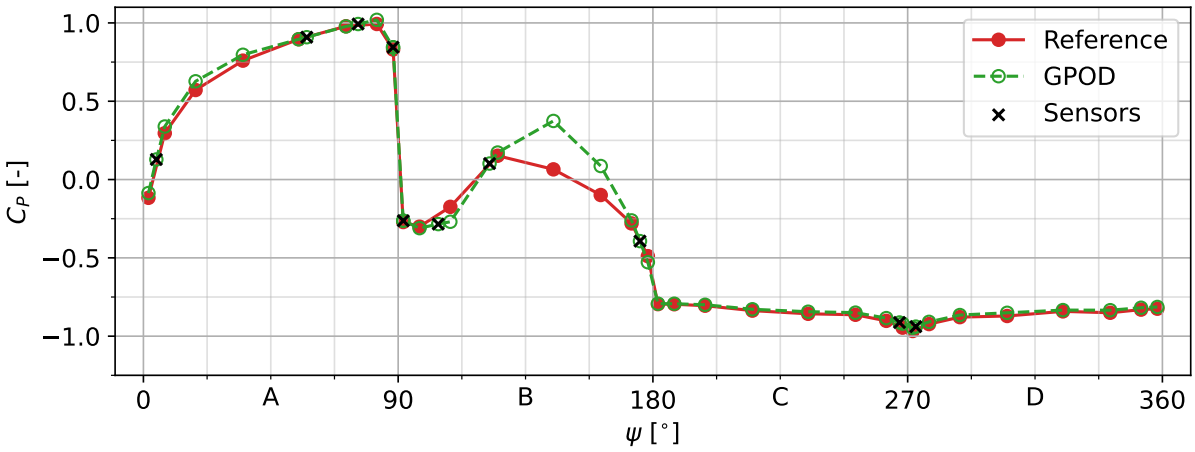


Figure 5.25: Opt V2 GPOD reconstruction at α = 27.5°, $n_s = K = 10$

Figures. 5.24 and 5.25 show the reconstructed pressure distributions at $\alpha = 25°$ and $27.5°$ respectively. The Opt V2 pressure taps provide experimental pressure measurements only on the sensor locations. Therefore, the references are taken as the cosine pressure distributions and do not necessarily coincide with the C_p values at the sensor locations but still provide an idea of what the reconstruction should look like. The difference between the two shown reconstructions is only $2.5°$ in angle of attack but they are illustrative of the reconstruction behaviour using this GPOD implementation with experimental training and sparse measurements but CFD based sensor locations.

Similar to earlier shown reconstructions using experimental data, the reconstruction on the 'pressure face(s)', in this case face A, is rather trivial and therefore unsurprisingly is in good agreement with the reference for both reconstructions shown. For $\alpha = 25°$, the reconstruction on face B agrees well with the reference as it follows the increase in pressure due to the reattached flow. On faces C and D, experimental measurements using the Opt V2 taps near $\psi = 270°$ are often offset to the reference (cosine)

pressure distributions which is also visible in Figure 5.24. Since the training data included the interpolated placeholder points at the locations where Opt V2 results are different from the reference, these discrepancies are not incorporated into the POD modes. The reconstructions however are constrained by the provided sensor data and thus pass through the Opt V2 measurements yielding a reconstruction similar in shape to the reference but positioned slightly below it. When assuming the lower pressures measured using the Opt V2 taps are indicative of the true pressure distribution (the distribution experienced during the Opt V2 measurements) being more negative than the reference on faces C and D, then the reconstruction might have predicted the pressure closer to the truly experienced distribution. Since the Opt V2 sensor locations determined using the CFD results are similar in layout to those identified from experimental data only (see Fig. 5.12 for example), the reconstruction at $\alpha = 25°$ is similar to 'regular' surface pressure reconstructions as discussed in Section 5.4.1.

The reconstruction at $\alpha = 27.5°$ shown in Figure 5.25 is close to the reference at faces A, C and D but less so at face B. This results in a larger RMSE C_p to the reference as shown in Figure 5.26 which will be discussed in Section 5.8.2. This overprediction of the pressure at face B seems unexpected judging from the similarity in the overall pressure distribution to Figure 5.24 and is possibly caused by one of or a combination of three reasons. First, despite the sparse measurements for GPOD being obtained separate from the cosine data used for building the POD basis, $\alpha = 27.5°$ not being included in the training data while $\alpha = 25°$ was makes it such that the pressure distribution at $\alpha = 25°$ is likely better represented through the POD modes. The difference between reconstructing pressure distributions which were or were not included in the training data however is still expected to be much smaller than it was for reconstructions in Section 5.4.1, making this only part of the cause for differences to the reference. Second, the GPOD reconstructions are forced to use the locations from the QR-CP algorithm applied to the CFD results as sensor locations. Using the sensor locations as shown in the Opt V2 reconstruction examples yields a condition number of matrix M (see Section 2.5) of 11 instead of 10 which would be obtained when using the QR-CP algorithm for sensor determination based on the experimental POD modes (relation between condition number and GPOD accuracy mentioned in Section 2.6). The difference is relatively small indicating that the CFD based sensor locations on the Opt V2 plates should not be significantly 'less optimal'. Yet, allowing the sensor layout to be determined by the QR algorithm applied on the cosine experimental modes instead yields a reconstruction in line with the results shown earlier in Section 5.4.2. Despite this seemingly confirming sub-optimality of the used sensor layout, it actually demonstrates how a third problem is mainly responsible; the interpolation performed to create placeholder points in the modes. Where at $\alpha = 25°$ the pressure measured by the Opt V2 sensors on face B are in line with what would be obtained by linearly interpolating the reference, this is not the case for the the sensor at $\psi = 104°$ for $\alpha = 27.5°$. Since the modes are 'trained' to expect a linear relation between the pressure at the sensors and nearby taps of the reference, a sensor not obeying this linear expectation locally disrupts the reconstruction.

5.8.2. Opt V2 pressure reconstruction accuracy

GPOD reconstructions as shown in Figure 5.24 and 5.25 do not have the true experienced pressure distribution available for comparison and hence the RMSE C_p is defined with respect to the cosine reference distribution. The Opt V2 pressure taps are suited for reconstructions at $n_s = K = 10$ only and therefore the accuracy in terms of the RMSE's is shown for $n_s = K = 10$ only in Figure 5.26.

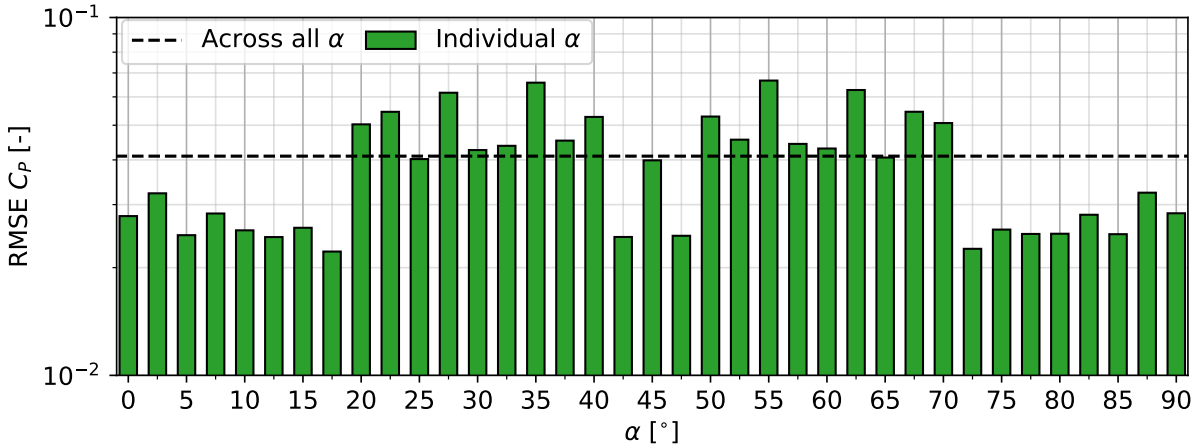


Figure 5.26: Opt V2 based reconstruction accuracy w.r.t. reference pressure distributions

The green bars illustrate the RMSE between the reconstruction and reference pressure distribution as computed for each angle of attack separately. The RMSE value at the dashed line is instead found by computing the RMSE between all reconstructions and reference pressures at once and thus represents how far off from the reference a single reconstructed pressure value is expected to be for any angle of attack within $0° \leq \alpha \leq 90°$. Note that on top of the error sources introduced in Section 5.6 and Section 5.8.1, the RMSE's in this plot additionally contain an error caused by the difference between the measurements obtained with the cosine plates and the Opt V2 plates. Considering this, the overall RMSE C_p of $4 \cdot 10^{-2}$ does not seem like a significant step back from the RMSE of roughly $2 \cdot 10^{-2}$ at $n_s = K = 10$ in Figure 5.15. This implies that despite the major difference between CFD and experimental measurements, the CFD based sensor locations are adequate for use with experimental GPOD reconstructions. RMSE values for individual angles of attack are clearly larger in the supercritical regime compared to the subcritical regime. The reason for this is expected to be twofold. First, the training data is distributed evenly over the complete angle of attack range while more focus on the supercritical regime might be necessary to capture the growing reattached region. This problem was identified earlier while discussing results in Section 5.4.2 and is one of the error sources discussed in Section 5.6. Second, especially this growing reattached region leads to spatial variation in pressure on a smaller scale than the distance between taps on the cosine plates. This invalidates the 'linear behaviour' expected by the POD modes at the locations of the Opt V2 sensor locations with the reconstruction errors discussed in Section 5.8.1 regarding Figure 5.25 as a result. An obvious solution to this artificially produced source of errors is to include the sensor locations identified by the QR algorithm applied to CFD data as pressure taps onto the plates responsible for measuring training data as

well (effectively combining Opt V2 and cosine plates that is).

Especially outside of the critical angle of attack range, where the aforementioned interpolation error has a small effect on reconstructions, the CFD based sensor locations are viable for experimental surface pressure reconstructions using GPOD. Using $n_s = K$, sparse reconstructions are extremely sensitive to the sensor layout. A single sensor being positioned in either an overall low-modal activity region or too close to other sensors such that there is a strong correlation would yield large condition numbers and an effectively underdetermined system involving matrix $\boldsymbol{M}$. Arbitrary sparse sensor placement, such as evenly spacing 10 sensors around the cylinder perimeter, will therefore almost certainly result in inaccurate reconstructions at $n_s = K$ (see example Figure 2.18). The CFD based sensor locations on the Opt V2 plates are not able to be better than experimental QR sensor positioning in terms of $\kappa(\boldsymbol{M})$ and also the Opt V2 reconstructions yielded generally larger RMSE values compared to Section 5.4.2. Regardless, both these indicators of sensor placement and reconstruction performance are unfortunately mostly limited by the linear interpolation procedure involved rather than GPOD or QR-CP in this case. This problem can be overcome relatively easily with the suggestion made at the end of the previous paragraph. Also, despite the large discrepancies between CFD and experimental results, the QR-CP algorithm was still able to identify a suitable sensor layout for experimental reconstructions, indicated by the similarity between CFD and experimental sensor layouts. Using CFD for prior determination of tap locations in an experimental setting is therefore expected to be viable when CFD results are expected to be trustworthy in at least a qualitative sense. Even when CFD results are inaccurate, as is the current case, QR-CP applied to the numerical POD modes yields a good initial guess for experimental sensor placement and can aid in determining optimal experimental sensor locations by indicating regions that might benefit from decreased tap spacing.

5.8.3. Opt V2 drag estimation accuracy

Drag coefficient estimations based on the sparse Opt V2 tap measurements are obtained by integration of the reconstructed surface pressure distributions. Since balance measurements were taken simultaneous with the surface pressure measurements, the drag coefficient as measured by the balance will be used as a reference instead of integrating the earlier used cosine pressure distributions. Both reconstructed and balance drag curves are included in Figure 5.27 below for comparison.

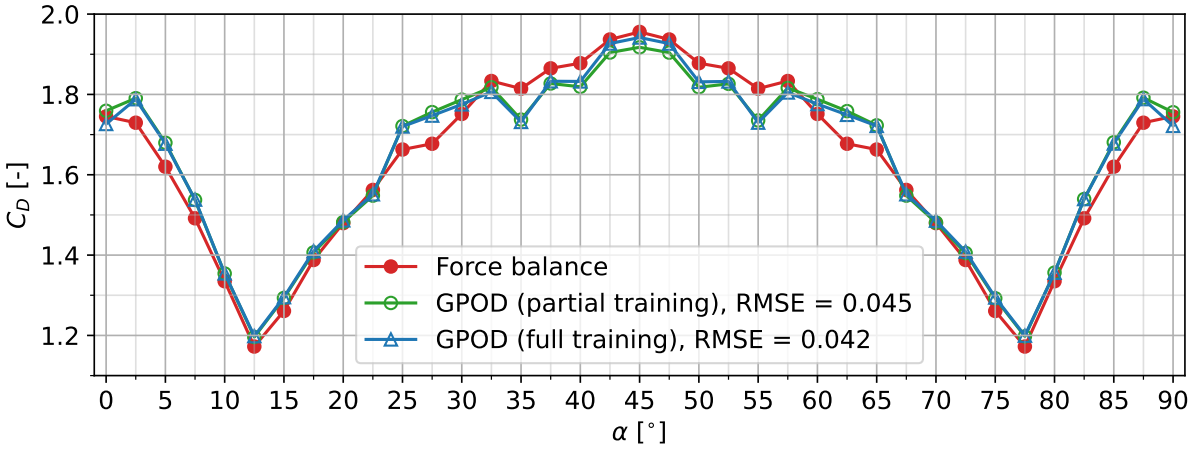


Figure 5.27: Opt V2 GPOD drag curve reconstruction using $n_s = K = 10$

The GPOD results consist of two drag curves; partial and full training. The 'partial training' curve is obtained by using the pressure distributions at angles of attack which are multiples of $5°$ as training data. Taking such a subset of the measured pressure distributions as training data has been the GPOD procedure throughout this report thus far and produce the figures as shown in Section 5.8.1. The 'full training' drag curve uses all available pressure distributions at the $2.5°$ interval in angle of attack for training data instead. Since the sparse measurements are obtained independently from the training measurements, this 'full training' approach would be closer to a practical application.

The overall RMSE of the partial training GPOD drag curve is 0.045 in the C_D. This is considerably worse compared to the results shown in Section 5.5.2 but relatively small compared to the magnitudes of the drag coefficients. Especially when taking into account the inherent difference between force balance and surface pressure integrated drag coefficients which were found to differ within the same order of magnitude in Section 5.2.2. In the region $25° \leq \alpha \leq 35°$, the partial training GPOD curve seems to correlate less with the force balance curve compared to outside of this angle of attack range (focusing on $\alpha \leq 45°$ as the curves are all symmetric). This is again mainly caused by the sparse measurements on the Opt V2 plates not following the linear behaviour with respect to nearby tap measurements while being forced to do so in the reconstruction, leading to erroneous reconstructions especially on face B in the supercritical regime. Surprisingly, the reconstruction at $\alpha = 25°$ significantly overpredicts drag while not suffering from the aforementioned problem at all as visible from Figure 5.24. Instead, the downward shift in the pressure reconstruction caused by the Opt V2 measurements on faces C and D is the cause. Either the difference between balance and integrated pressure drag coefficients is larger than expected and the reconstruction is accurate or the truly experienced dip in pressure near $\psi = 270°$ is steeper when using the Opt V2 plates compared to the cosine plates used for the training data. Despite using roughly double the number of surface pressure distributions as training data, the full training GPOD curve achieves only a marginally lower RMSE of 0.042. The reason for this disappointing increase in accuracy is the sensor locations on the Opt V2 plates being based on the partial training pressure distribu-

tions which are therefore not necessarily suited for reconstructions using a different set of training snapshots. CFD based sensor locations on the Opt V2 plates should have been determined using all available numerical pressure distributions instead, as would be the case for a practical scenario.

5.9. CFD trained experimental surface pressure reconstruction using GPOD

In Section 5.8, sensor locations based on CFD snapshots of square cylinder flow were used in combination with experimentally obtained POD modes and sparse surface pressure measurements to apply GPOD. In the following section this will be taken a step further as the POD modes used will additionally be taken from the numerical results such that no experimental training data is involved and reconstructions are generated directly from sparse experimental measurements. The CFD results have been discussed separately in Section 4.

5.9.1. Reconstruction examples

The tap locations on the Opt V2 plates were determined using the CFD results as discussed in Section 4.4 and the experimental surface pressure measurements from these taps will be used for the GPOD reconstructions. Unfortunately, the POD modes used to obtain these tap locations were determined using numerical training snapshots of a subset of the total angle of attack ensemble (partial training, see Section 5.8.3) instead of surface pressure distributions at all available angles of attack. Reconstructions will therefore not benefit significantly from using numerical POD modes from training at all angles of attack and the 'partial training' numerical POD modes are used instead. These numerical POD modes are thus the same as used in Section 4.4 of which the two dominant modes are shown in Figure 4.7. The ten surface pressure measurements at each angle of attack obtained using the Opt V2 taps are used with the first ten numerical POD modes to achieve $n_s = K = 10$. The numerical POD modes have the same spatial resolution as the mesh around the square cylinder with $n = 544$. Reconstructions will therefore use the ten Opt V2 pressure measurements to obtain a pressure distribution also at a spatial resolution of $n = 544$.

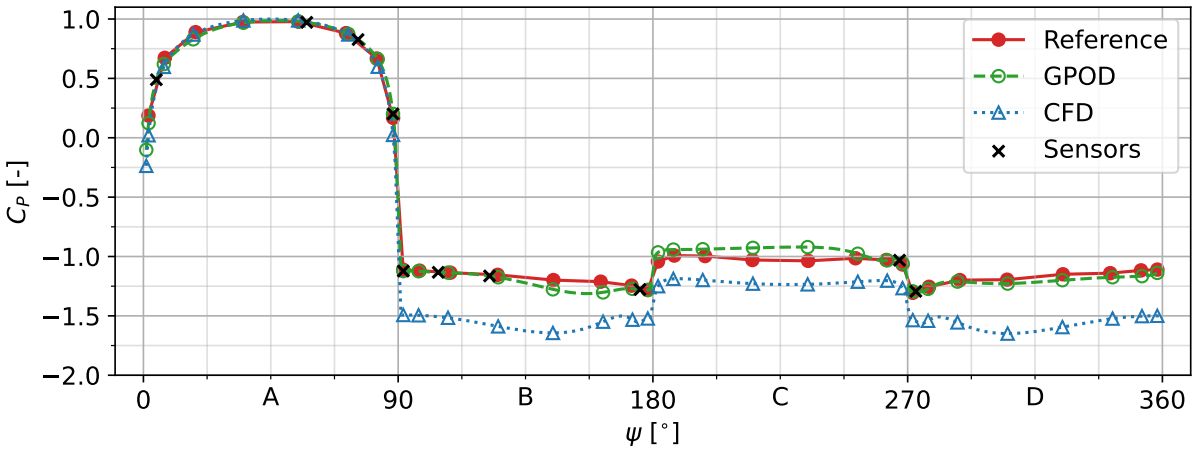


Figure 5.28: CFD trained GPOD reconstruction using experimental Opt V2 sensor measurements compared to cosine reference- and CFD pressure distribution at $\alpha = 0°$, $n_s = K = 10$.

Figure 5.28 above shows the GPOD reconstruction using the CFD POD modes and the experimental surface pressure measurements indicated by a cross at $\alpha = 0°$. The complete experimental pressure distribution is unknown and therefore the cosine pressure distribution at the same angle of attack is included as a reference. Also the CFD pressure distribution which, depending on the angle of attack, has been included in the training is shown. At this angle of attack, the CFD results agree well with the reference in a qualitative sense but still the GPOD reconstruction improves upon the CFD pressure distribution at almost every reference tap location (indicated with markers). The GPOD reconstruction shows asymmetry on faces B and C which is caused by the POD modes being 'active' at this angle of attack also representing features of pressure distribution from different angles of attack. Even though the CFD results overall resemble experimental results very poorly, the difference between CFD and experimental results in both qualitative and quantitative sense is smallest in the subcritical regime. This yields reasonably accurate GPOD reconstructions in this regime, as demonstrated in Figure 5.28, but even for reconstructions below α_{crit}, the POD modes are 'diluted' by the inaccurate numerical pressure distributions at $\alpha > \alpha_{crit}$ being used as training. This is also part of the reason for the differences between reference and GPOD in Figure 5.28. Starting close to the critical angle of attack, where CFD results start to become increasingly bad w.r.t. experimental, the GPOD reconstructions become increasingly inaccurate as well. This is demonstrated in Figure 5.29 which shows the reconstruction at $\alpha = 12.5°$.

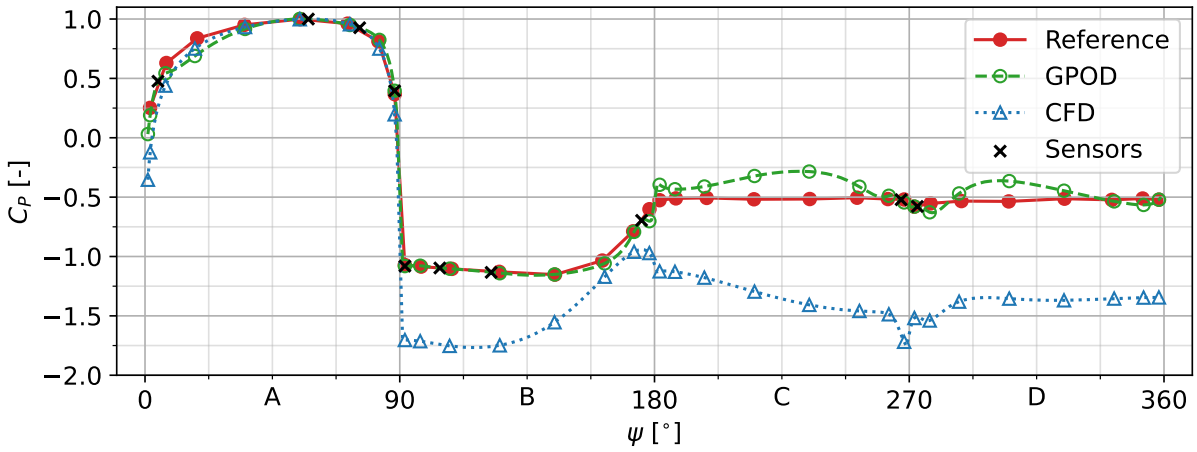


Figure 5.29: CFD trained GPOD reconstruction using experimental Opt V2 sensor measurements compared to cosine reference- and CFD pressure distribution at $\alpha = 12.5°$, $n_s = K = 10$.

The CFD results match the experimental results well on the 'pressure sides' of the square cylinder for every angle of attack and hence the GPOD reconstruction at those faces is generally quite accurate, as is the case on face A at this angle of attack. The CFD results predict well established reattachment of the shear layer on face B at $\alpha = 12.5°$ visible through a peak in pressure while experimental results as discussed in Section 5.1 identify this angle of attack to be α_{crit} with no pressure peak. The GPOD reconstruction seems to be able to correct for this misidentification of the critical angle of attack by the CFD results as it matches the reference well on face B. $\alpha = 12.5°$ marks the start of CFD results greatly overpredicting drag through the presence of unphysical secondary recirculation regions in the wake, underpredicting pressure on face C and D while changing the shape of the distribution. This is already visible in Figure 5.29 with the negative slope in the CFD pressure towards $\psi = 270°$ which through the training snapshots is encoded in the POD modes. This forces the the GPOD reconstruction on faces C and D to include this same curvature while complying with the sparse experimental measurements, leading to inaccurate results that instead overpredict pressure and behave unphysical. As the effect of the unphysical secondary recirculation regions in the CFD results gets stronger for increasing angle of attack, the reconstructions become worse as exemplified by the reconstruction at $\alpha = 45°$ in Figure 5.30.

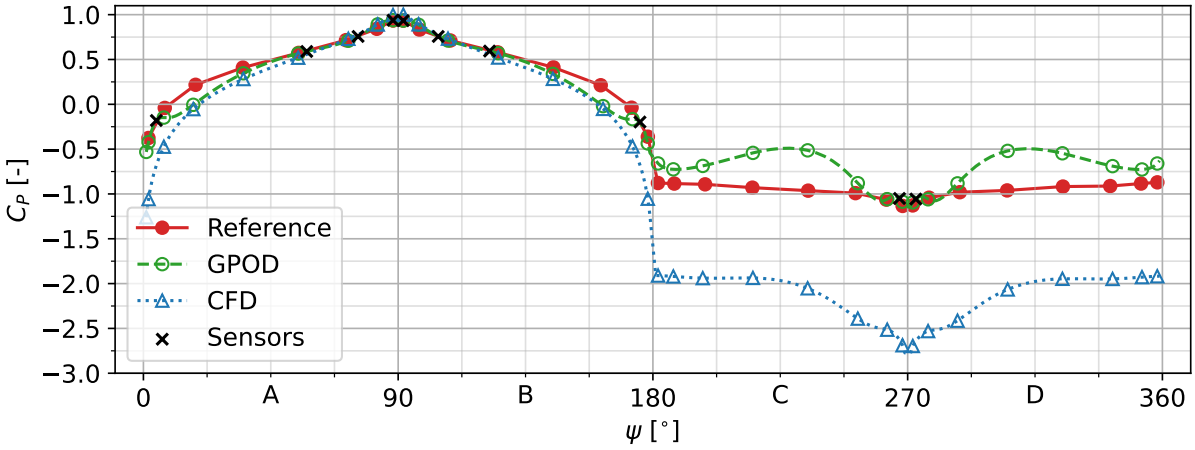


Figure 5.30: CFD trained GPOD reconstruction using experimental Opt V2 sensor measurements compared to cosine reference- and CFD pressure distribution at $\alpha = 45°$, $n_s = K = 10$.

At $\alpha = 45°$, similar inaccurate curvature and overprediction on faces C and D are present. Unlike non-data driven interpolation methods, such as linear interpolation, that make use of local spatial information only, the POD bases involved in GPOD reconstructions enable all sparse measurements to be used for reconstructions simultaneously based on their relative spatial correlation. This was seen in Section 5.7 to yield a more accurate method of overcoming sparseness compared to linear interpolation however also enables 'locally' inaccurate results, as seen in the CFD results on faces C and D, to be strongly correlated in the modes with other regions to be reconstructed and therefore spoil the reconstruction overall. This is seen in the figure above near the edges of faces A and B where the high 'activity' of unphysical modes (such as those resembling the CFD pressure distribution at $\alpha = 45°$, see Figure 4.7) force the reconstruction into unphysical behaviour to comply with all experimental measurements on the sensor locations.

5.9.2. Pressure reconstruction accuracy

The accuracy of reconstructions such as those shown in Section 5.9 are quantified in terms of RMSE C_p with respect to the cosine reference pressure distribution at all angles of attack from $0°$ to $90°$. This yields RMSE C_p values for individual reconstructions but also a single RMSE C_p value across all reconstructed surface pressure distributions is computed such that results in the same format as used before in Section 5.4.2 and 5.7.2 are available.

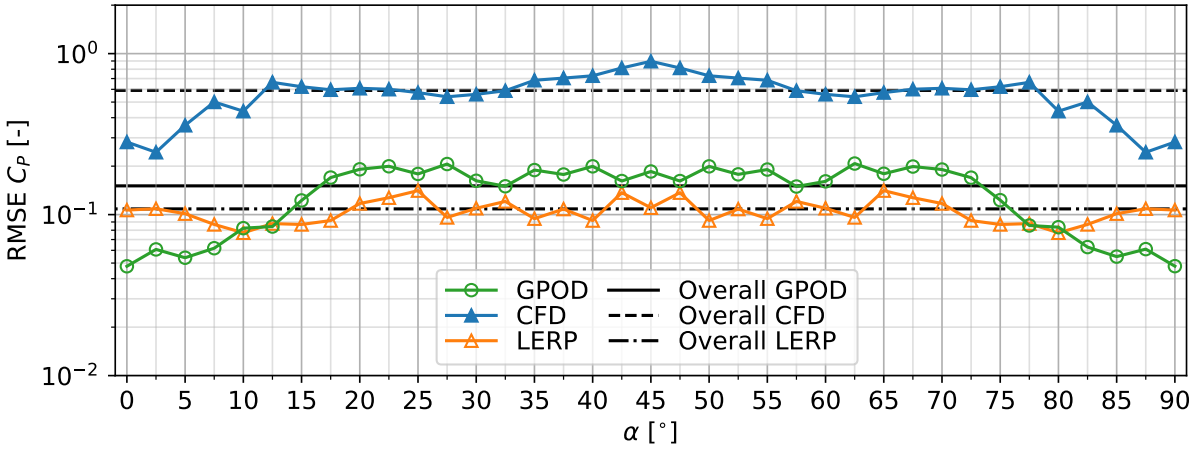


Figure 5.31: RMSE C_p to cosine reference achieved by CFD trained GPOD reconstructions using experimental Opt V2 sensor measurements. Also included: RMSE C_p of the CFD surface pressure distributions to the same cosine reference as well as the RMSE C_p achieved by linear interpolation (LERP) using the same Opt V2 sensors.

CFD results and GPOD reconstructions have a spatial resolution of $n = 544$ compared to the $n = 32$ of the cosine reference. In the computation of the RMSE's therefore only those entries in the CFD and GPOD results at spatial locations at which cosine taps are positioned as well are considered (corresponding to the markers in the examples of Section 5.9). Unsurprisingly, the overall RMSE value for the CFD results is extremely large at approximately $6 \cdot 10^{-1}$. At $\alpha < 12.5°$ and $\alpha > 77.5°$, individual RMSE values are significantly smaller as the numerical results qualitatively match the experimental results however remain large in an absolute sense. The overall GPOD reconstruction results have an RMSE C_p of $1.5 \cdot 10^{-1}$ and thus improve upon the CFD results by a factor 4 but still remains significantly larger than the results at $n_s = K = 10$ of Figure 5.15. RMSE values for individual angles of attack however show that there is much variation in the GPOD results. Below the critical angle of attack, GPOD RMSE values are below 10^{-1} and show a decrease with respect to the CFD results in the RMSE of up to a factor 8 at $\alpha = 7.5°$. At $\alpha = 0°$ (and $90°$), the GPOD reconstruction reaches a minimum RMSE of $4.7 \cdot 10^{-2}$ which for the same number of sensors is comparable to the RMSE achieved using experimental training data, without actually requiring experimental training. The figure also includes the results one would obtain when using linear interpolation as done in Section 5.7 on the experimental measurements otherwise used for GPOD reconstructions. The overall RMSE C_p is considerably smaller than the overall GPOD results at $1.1 \cdot 10^{-1}$. This overall LERP RMSE value is also close to the RMSE C_p value at the comparable $n_s = K = 10$ in Figure 5.22 (note that the LERP results in Section 5.7 were obtained using a different but similar sensor layout and used the even pressure distributions as reference instead). Within the subcritical regime, individual RMSE values for the LERP reconstructions lie above those for the GPOD reconstructions however for all remaining angles of attack (roughly $80\,\%$ out of the total of 37 angles of attack) the GPOD reconstructions are very much unphysical/inaccurate and are not able to perform better than the simple to implement linear interpolation which does not require training in any form.

5.9.3. Drag estimation accuracy

The LERP and GPOD surface pressure distributions are integrated to obtain drag coefficients for all angles of attack. Together with the CFD drag coefficients, these are shown in Figure 5.32 and compared to the force balance results also included. The force balance results are obtained simultaneously with the surface pressure measurements on the Opt V2 plates which were used for both GPOD and LERP reconstructions. Balance measurements were found earlier in Section 5.2.2 to compare well with integrated surface pressure distributions which are not available for the sparse Opt V2 measurements.

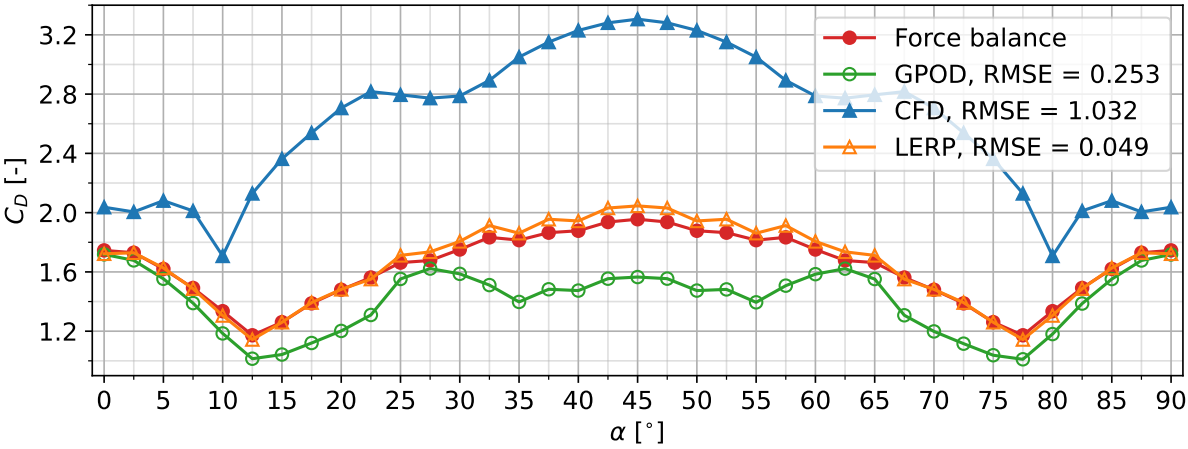


Figure 5.32: C_D versus α curve obtained by CFD trained GPOD reconstructions using experimental Opt V2 sensor measurements. Also included: CFD drag curve, drag curve obtained by linear interpolation (LERP) using the same Opt V2 sensors and the reference drag curve obtained through experimental balance measurements.

The extreme overprediction of the drag coefficients from the CFD results is not surprising considering the underprediction of pressure on faces C and D and the earlier shown results of Section 4.3. The GPOD results improve upon the CFD results by again roughly a factor 4 and the trend in accuracy with respect to the balance measurements follows the same pattern as in Figure 5.31. Within the supercritical regime, pressure underprediction in the CFD results forces the GPOD reconstructions to overpredict pressure on faces C and/or D with drag underprediction as a result. GPOD results are significantly better in the subcritical regime where the CFD results were seen to provide physically valid surface pressure distributions as training. The incorrect α_{crit} predicted around $10°$ by the CFD results has been corrected through the GPOD procedure with he help of sparse experimental measurements. The LERP results are surprisingly accurate to the force balance measurements with an overall RMSE in the C_D of 0.049. Even in the subcritical regime, where the RMSE C_p for GPOD was previously found to be smaller compared to LERP, LERP outperforms GPOD. Such a small RMSE C_D does not fall in line with expectations based on LERP RMSE C_D results shown in Figure 5.23 and indeed the specific Opt V2 tap layout seems to produce equal under- and overprediction of the surface pressure distribution at locations which cancel in the integration for the drag coefficient. Mainly in the subcritical regime therefore; LERP reconstructions might perfom better in terms of RMSE C_D but produce surface pressure distributions not as physically valid as GPOD. This

does not weigh up against the additional training steps involved required for GPOD reconstructions and limited accuracy outside the subcritical regime. For both surface pressure reconstruction and drag coefficient estimation using sparse measurements, using GPOD in combination with RANS training snapshots does not seem viable for this specific case and especially more qualitatively accurate numerical results would be required. Providing numerical training results closer to experiments in the supercritical regime is not expected to improve GPOD in the supercritical regime only but would also benefit GPOD reconstructions in the subcritical regime as the activity of POD modes involved in reconstructions often spans across a wide range of angles of attack.

6

Conclusions and recommendations

In this chapter the conclusions which can be drawn from the results presented and discussed in Chapter 5 are stated. The conclusions are split up over the different topics addressed but together provide the information to cover the research objective. The conclusions are followed up by several recommendations. These recommendations are aimed at improving the usability/accuracy of GPOD under different conditions (e.g. in combination with CFD) and would be amongst the steps undertaken by the author of the current report for follow-up research.

6.1. Conclusions

Aerodynamic drag determination through integration of the surface pressure distribution requires a large number of pressure taps/sensors to be installed on the object of interest. This increases complexity and makes it difficult for use outside of the controlled environment of a windtunnel while on-site measurements might be preferred from a practical point of view. Regardless of the setting in which measurements are to be performed, achieving a reduction in the number of required pressure taps/sensors while maintaining the ability to obtain surface pressure distributions suited for the determination of aerodynamic coefficients can be beneficial for optimization, control or applications with limited physical space/access for example.

Using adequate prior training datasets with high spatial resolution, underlying patterns in the pressure distribution in some parameter space of interest can be recognized and encoded in a reduced order model using the Proper Orthogonal Decomposition. Such a model can be used to infer full state approximations of the pressure distribution guided by only sparse measurements instead using the Gappy Proper Orthogonal Decomposition extension, if locations for such measurements are optimal for the specific use case. Training data and optimal sensor locations can be determined through windtunnel measurements or possibly CFD which would avoid the need for prior windtunnel testing all together. The goal of this thesis was to evaluate the feasibility of the outlined method for practical use, hence the following research objective:

”To assess the viability of using Gappy Proper Orthogonal Decomposition as a method to obtain surface pressure distributions and drag estimates based on experimentally obtained sparse surface pressure measurements on a 2D square cylinder”

The square cylinder model serves as a benchmark bluff body object and has been incorporated into a windtunnel setup at the Open Jet Facility at the TU Delft. Using removable plates on the square cylinder model faces, the pressure tap distribution around the circumference could be varied but contain at most 32 taps. Pressure distributions were measured at angles of attack from $\alpha = 0°$ to $45°$, at $2.5°$ intervals, which could be extended to $90°$ making use of the model symmetry. The parameter space is therefore the pressure distribution at $0° \leq \alpha \leq 90°$. Balance measurements were taken simultaneous to pressure measurements mainly for independent validation. Surface pressure distributions at the same angles of attack were also obtained using RANS simulations.

Comparison of the experimental surface pressure distributions to results obtained in literature revealed large quantitative differences however also showed that the experimental setup has been successful in realizing a nominally two-dimensional square cylinder flow which agrees well in a qualitative sense to literature as desired. Either by using evenly spaced pressure taps around the model perimeter or cosine taps, which are more densely spaced towards the edges, pressure drag coefficients obtained by integrating the surface pressure around the perimeter were in very good agreement with the balance measurements, using 8 taps per face. This further confirmed a successful nominally 2D square cylinder flow. With a Root Mean Squared Difference (RMSD) of 0.019 between balance and cosine pressure tap drag curves, using a tap layout dense towards the edges provides a slightly more accurate total drag approximation as a RMSD of 0.034 was achieved using the even tap layout. The CFD results were valid within the subcritical regime but very much unphysical and inaccurate in the supercritical flow regime.

Regardless of the inaccuracy, the POD procedure was performed on a subset of the CFD results (training). Sparse GPOD reconstructions outside the training data and based on locations provided by the QR-CP algorithm showed that overall RMSE C_p values in the order of 10^{-2} are achievable with roughly 99 % of the original pressure distribution discarded (5 to 10 sensors remain). More importantly, the QR-CP algorithm was confirmed to be a robust method of providing nearly optimal sensor locations for sparse sampling as GPOD performance remained close to that of POD projections while avoiding an expensive combinatorial search.

The identical procedure was applied to the experimentally obtained pressure distributions. The POD analysis for eigenmodes and corresponding QR-CP sensor locations were based on experimental training pressure distributions at a subset of the angle of attack ensemble. Both training and the determination of sensor locations therefore relied on prior experimental measurements. The training pressure distributions formed a representative basis for square cylinder flow at $0° \leq \alpha \leq 90°$ as sparse GPOD reconstructions outside of the provided training were capable of achieving RMSE C_p values again in the order of 10^{-2}. The number of sensors required for physically valid but relatively sparse reconstructions seemed to be between 5 and 10 which due to the 'high resolution' reference data consisting of 32 measurements translates to a gappiness percentage of 85 % and 69 % respectively. The QR-CP sensors were found to be

suited sensor locations for reconstructions in this fully experimental setting however the greedy nature of the algorithm became apparent through an instance of clearly sub-optimal sensor placement leading to large errors. Drag approximations based on the reconstructed pressure distributions were found to be capable of following the reference curve at the same number of sensors as mentioned before. The behaviour of the drag prediction accuracy for increasing the number of sensors was however very unpredictable compared to pressure distribution reconstruction accuracy. A comparison of both pressure distribution reconstruction accuracy and drag approximation accuracy with respect to linear interpolation based on the same sensor locations showed how GPOD perfroms better for both by almost an order of magnitude in RMSE depending on sparseness.

The reliance of experimental GPOD reconstructions on sensor locations determined through experimental training was tested by using the numerical pressure distributions as the basis for the QR-CP algorithm instead. Despite the large discrepancies between experimental and CFD pressure distributions, this 'numerically determined' sensor layout was very similar to that obtained experimentally. Only surface pressure reconstructions using 10 sparse measurements from these sensor locations could be performed at which the overall RMSE C_p of $4 \cdot 10^{-2}$ was not significantly worse compared to the $1.7 \cdot 10^{-2}$ obtained using a sensor layout based on experimental data. Especially when considering that the expected sources of error for these reconstructions were introduced unnecessarily through oversights in the design of the experimental campaign which can be mitigated in future experiments. Because of these errors, drag approximations based on the GPOD reconstructions suffered in accuracy as well.

Incorporation of CFD results into GPOD reconstructions using sparse experimental measurements was taken a step further by additionally relying on the numerical pressure distributions for training. Therefore both the eigenmodes and sensor locations are determined from CFD and only the sparse measurements remain experimentally obtained. Reconstructions using this approach were generally disappointing because of the unphysical CFD results being encoded into the eigenmodes. Overall, the sparse experimental measurements did provide a significant correction to the CFD results using GPOD however even linear interpolation based on the same sparse measurements provided smaller RMSE C_p values while not requiring training data in any form. Exceptions on this were seen in the subcritical regime where CFD results remained somewhat accurate to the experimental surface pressure distributions. Within the subcritical regime, GPOD reconstructions were more accurate in terms of RMSE C_p and physically valid compared to linear interpolation which implies a potential for GPOD to produce meaningful reconstructions using at least qualitatively accurate training from CFD, avoiding extensive windtunnel testing to survey the parameter space. The disappointing GPOD surface pressure reconstructions yielded disappointing drag predictions overall as well. Again, an improvement was achieved over the CFD results but far better drag coefficient estimations were obtained using linear interpolation. Within the subcritical regime GPOD drag predictions were naturally closer to the reference but still surpassed in accuracy by linear interpolation. It should be noted however that the surprisingly small errors in drag estimation through linear interpolation were

mainly a result of 'coincidental' surface pressure reconstruction errors cancelling during integration providing no physical foundation for the surprising performance. On the contrary, GPOD drag estimates within the subcritical regime were found to physically comply with the experimental references much better.

In summary, the GPOD framework was shown to provide a viable method of reducing the required number of surface pressure measurements for applications were additional sensors might be expensive, difficult to install or add unacceptable latency/weight *if* there is a possibility for high resolution experimental training to be performed. Combining experimental training with the QR-CP algorithm for the sensor layout produced both accurate surface pressure reconstructions and drag predictions hence 'fully experimental' GPOD is colored green in the summarized conclusion shown in Table 6.1. Using instead CFD for the determination of the sensor layout with the QR-CP algorithm resulted in mixed outcomes. This approach showed the potential for accurate reconstructions when the source of errors was not dominated by the procedure to ensure compatibility between different pressure tap plates. Hence the problematic implementation instead of the GPOD procedure was found to be the limiting factor. Determining the sensor layout through CFD can still be used as a prior for more effective/efficient sensor placement in an experimental setting as the QR-CP algorithm succeeded in predicting GPOD suited sensor locations despite the inaccurate numerical results. Overall, reconstructions using this approach were less accurate for both surface pressure reconstructions and drag predictions than the fully experimental GPOD but neither the sensor layout or training data are the cause of this, therefore both are colored yellow in Table 6.1. Replacing the experimental results used for training and sensor placement by numerical results has obvious practical advantages but could not be used effectively in a consistent manner in this report and thus corresponding reconstructions are colored red in Table 6.1. CFD results should agree with the physical behaviour of the system to be sampled sparsely. If this holds however, it was seen that sparse measurements can effectively be used to merge numerical- with sparse experimental results and produce practically continuous pressure distributions that are closer to 'the truth' than either numerical or sparse experimental results on their own. For the majority of the reconstructions using CFD for both sensor placement and training, the limiting factor was clearly the inaccuracy of the CFD results affecting the POD modes, hence colored red in Table 6.1.

Table 6.1: Summarized conclusion on the experimental GPOD reconstruction results depending on the source of the sensor layout and POD modes. RMSE C_p and RMSE C_D at $n_s = K = 10$ (69 % gappiness) stated in last columns for reference.

Sensor layout	**Training/POD modes**		**Experimental GPOD reconstructions**	
			C_p RMSE $\cdot 10^{-2}$	C_D RMSE $\cdot 10^{-2}$
Experimental	Experimental		1.7	0.82
CFD	Experimental	$\rightarrow$	4.0	4.5
CFD	CFD		15	25

6.2. Recommendations

In this thesis, GPOD has been applied on sparse experimental measurements in three ways; fully experimental with both training and the determination of the sensor layout relying on windtunnel measurements, partly CFD based where the sensor layout is determined through an ensemble of numerical snapshots and fully CFD based with both the training and determination of the sensor layout coming from numerical simulations. Each of the implementations achieved a different level of success and viability for practical use. Recommendations provided in this section are either aimed at increasing the viability of GPOD in general or focus on a specific implementation as used in the report.

6.2.1. Changes to the experimental pressure tap layouts

The windtunnel measurements for the generation of training data consisted of surface pressure measurements using 32 taps in total. This spatial resolution was found to be sufficient for the determination of drag but the discreteness of the measured pressure distribution limits the 'optimality' that can be achieved with the sparse sensor placement as the underlying pressure distribution is continuous in nature. Discrete pressure distributions translate into discrete POD modes which are used for the determination of the sensor positions. Since GPOD relies on effective sampling of regions with high modal activity in the modes, increasing the resolution at which these modes are captured, i.e. increasing the number of pressure taps, possibly allows for better sensor positions to be identified and hence better surface pressure reconstructions and drag approximations. An effective approach to efficiently increase the resolution of the measured surface pressure distribution near regions of expected high modal activity could be to use the CFD based sensor positions as an initial guess for regions which might benefit from an increased tap density.

A second recommendation related to the locations at which pressure taps were placed during the windtunnel campaign is aimed at resolving the problem encountered when reconstructing pressure distributions using the Opt V2 pressure tap measurements (see Section 5.8.1). To use GPOD, the eigenmodes involved in a reconstruction should contain data at the locations which are used as sensor locations. This was not the case for the Opt V2 reconstructions but was circumvented using linearly interpolated placeholder points. The inherent assumption this introduces regarding the shape of the pressure distribution at the Opt V2 tap locations leads to problems in the reconstructions at certain angles of attack. This is avoidable in general however by including the tap layout responsible for sparse sampling specifically (Opt V2) into the tap layout used for obtaining training snapshots (even or cosine). This is expected to significantly increase the accuracy of reconstructions with respect to the results shown in this report and would better illustrate the viability of using CFD for the determination of sensor locations as found RMSE values are currently spoiled by the implementation rather than the limits of GPOD.

6.2.2. Improvement of training data

Better GPOD reconstructions can obviously be achieved through training data which is more representative of the data to be reconstructed. Training data not being suited for GPOD reconstructions is the dominant factor in the disappointing reconstructions obtained when using CFD for both the training and sensor placement. For the specific application on a square cylinder, poor numerical results were caused by the inadequacy of the RANS simulations for massively separated flows. In general, RANS simulations are not expected to be viable for use on bluff bodies with large detached/reattaching flow regions (Probst et al. (2010), Ke (2019)) but using RANS to replace windtunnel testing for GPOD training was shown to be viable on aerodynamic bodies as shown in Mifsud et al. (2019). For broader usability however, numerical training for applications involving bluff body aerodynamics might require using an LES-type approach. Direct Large Eddy Simulations to survey the parameter space for training are not viable at the Reynolds numbers often encountered in practical applications. Hybrid RANS/LES approaches however try to combine the accuracy of LES and efficiency of RANS which might therefore be an alternative to RANS in applications where numerical accuracy is otherwise lacking and windtunnel measurements are not an option.

An aspect which affects GPOD reconstructions but has not been investigated in this thesis is the methodology of sampling the parameter space for training snapshots. With a few exceptions, training snapshots in this report have been provided at a constant $5°$ angle of attack interval in the parameter space. Variations in reconstruction accuracy over the angle of attack range suggest however that reconstructions at specific angles of attack might benefit from a locally decreased interval between training snapshots or in other words: uneven sampling of the parameter space. For this particular application, more closely spaced training snapshots in the supercritical regime are expected to increase reconstruction accuracy but other applications such as flow around an airfoil for example might benefit from training snapshots around the onset of stall or formation of a shockwave. Optimizing the distribution of training snapshots around such non-linearity might require prior knowledge, additional CFD simulations or a systematic approach such as described in Zhan et al. (2016). This might become increasingly important when the parameter space is spanned by more than a single variable, something not included in this report but which could have practical use as it effectively allows measuring multiple variables using the same sensors.

6.2.3. Oversampling

A final recommendation is to investigate the effect of oversampling. Throughout this report, the number of used POD modes has been kept equal to the number of available sensors. This effectively makes the least squares formulation in the GPOD algorithm redundant as instead sensor measurements are interpolated exactly. Especially in experimental settings, where sensor measurements might contain noise, this could lead to undesired effects as reconstructions are forced to obey noisy measurements. Oversampling relieves the constraint to follow measurements exactly as an overall least squares reconstruction is generated instead. This additionally opens up the possibility for weighted and regularized least squares approaches which focus on gener-

ating reconstructions specifically to conform with integrated aerodynamic coefficients such as in Mifsud et al. (2019) and Jayaraman et al. (2019).

Bibliography

B. R. Aiello, K. E. Stanchak, A. I. Weber, T. Deora, S. Sponberg, and B. W. Brunton. Spatial distribution of campaniform sensilla mechanosensors on wings: form, function, and phylogeny. *Current Opinion in Insect Science*, 48:8–17, 2021. ISSN 2214-5745. doi: https://doi.org/10.1016/j.cois.2021.06.002. URL https://www.sciencedirect.com/science/article/pii/S2214574521000602. NeuroScience . Special Section on Insects as food and feed.

G. Berkooz, P.J. Holmes, and J. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 25:539–575, 11 1993. doi: 10.1146/annurev.fl.25.010193.002543.

O. Botella and R. Peyret. Benchmark spectral results on the lid-driven cavity flow. *Computers & Fluids*, 27(4):421–433, 1998. ISSN 0045-7930. doi: https://doi.org/10.1016/S0045-7930(98)00002-4. URL https://www.sciencedirect.com/science/article/pii/S0045793098000024.

I. Bright, G. Lin, and J. N. Kutz. Compressive sensing based machine learning strategy for characterizing the flow around a cylinder with limited pressure measurements. *Physics of Fluids*, 25(12):127102, 2013. doi: 10.1063/1.4836815. URL https://doi.org/10.1063/1.4836815.

S. L. Brunton and J. N. Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019. doi: 10.1017/9781108380690.

T. Bui-Thanh, M. Damodaran, and K. Willcox. Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA Journal*, 42(8):1505–1516, 2004. doi: 10.2514/1.2159. URL https://doi.org/10.2514/1.2159.

L. A. Burelle, W. Yang, F. Kaiser, and D. E. Rival. Exploring the signature of distributed pressure measurements on non-slender delta wings during axial and vertical gusts. *Physics of Fluids*, 32(11):115110, 2020. doi: 10.1063/5.0025860. URL https://doi.org/10.1063/5.0025860.

L. Carassale, A. Freda, and M. Marrè-Brunenghi. Experimental investigation on the aerodynamic behavior of square cylinders with rounded corners. *Journal of Fluids and Structures*, 44:195–204, 2014. ISSN 0889-9746. doi: https://doi.org/10.1016/j.jfluidstructs.2013.10.010. URL https://www.sciencedirect.com/science/article/pii/S0889974613002296.

A. Chatterjee. An introduction to the proper orthogonal decomposition. *Current Science*, 78(7):808–817, 2000. ISSN 00113891. URL http://www.jstor.org/stable/24103957.

S. Chaturantabut and D. C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM Journal on Scientific Computing*, 32(5):2737–2764, 2010. doi: 10.1137/090766498. URL https://doi.org/10.1137/090766498.

J. M. Chen and C. Liu. Vortex shedding and surface pressures on a square cylinder at incidence to a uniform air stream. *International Journal of Heat and Fluid Flow*, 20(6):592–597, 1999. ISSN 0142-727X. doi: https://doi.org/10.1016/S0142-727X(99)00047-8. URL https://www.sciencedirect.com/science/article/pii/S0142727X99000478.

K. Cohen, S. Siegel, and T. McLaughlin. Sensor placement based on proper orthogonal decomposition modeling of a cylinder wake. volume 4259, 06 2003. ISBN 978-1-62410-095-6. doi: 10.2514/6.2003-4259.

CPP, 2022. URL https://cppwind.com/. Accessed: 16-11-2022.

CSIR-NAL, 2022. URL https://www.nal.res.in/index.php/en/techniques/steady-pressure-measurements. Accessed: 16-11-2022.

Delft University of Technology, 2022. URL https://www.tudelft.nl/en/ae/organisation/departments/aerodynamics-wind-energy-flight-performance-and-propulsion/facilities/low-speed-wind-tunnels/open-jet-facility. [Accessed 12-07-2022].

Z. Drmač and S. Gugercin. A new selection operator for the discrete empirical interpolation method—improved a priori error bound and extensions. *SIAM Journal on Scientific Computing*, 38(2):A631–A648, jan 2016. doi: 10.1137/15m1019271. URL https://doi.org/10.1137%2F15m1019271.

R. Everson and L. Sirovich. Karhunen–loève procedure for gappy data. *JOSA A*, 12, 08 1995. doi: 10.1364/JOSAA.12.001657.

X. He and D. R. Williams. *Aerodynamic Loads and Surface Pressure Characteristics on a Wing in Transverse Cross-Flow Gusts*. 2022. doi: 10.2514/6.2022-0044. URL https://arc.aiaa.org/doi/abs/10.2514/6.2022-0044.

J. E. Higham, W. Brevis, and C. J. Keylock. A rapid non-iterative proper orthogonal decomposition based outlier detection and correction for PIV data. *Measurement Science and Technology*, 27(12):125303, oct 2016. doi: 10.1088/0957-0233/27/12/125303. URL https://doi.org/10.1088/0957-0233/27/12/125303.

Honeywell, TruStability board mount pressure sensors, HSC-series, 2022. URL https://prod-edam.honeywell.com/content/dam/honeywell-edam/sps/siot/en-us/products/sensors/pressure-sensors/board-mount-pressure-sensors/trustability-hsc-series/documents/sps-siot-trustability-hsc-series-high-accuracy-board-mount-pressure-sensors-500
pdf?download=false. [Accessed 13-07-2022].

R.F. Huang, B.H. Lin, and S.C. Yen. Time-averaged topological flow patterns and their influence on vortex shedding of a square cylinder in crossflow at incidence. *Journal of Fluids and Structures*, 26(3):406–429, 2010. ISSN 0889-9746. doi: https://doi.org/10.1016/j.jfluidstructs.2010.01.003. URL https://www.sciencedirect.com/science/article/pii/S0889974610000137.

G. Iacobello, F. Kaiser, and D. E. Rival. Load estimation in unsteady flows from sparse pressure measurements: Application of transition networks to experimental data. *Physics of Fluids*, 34(2):025105, 2022. doi: 10.1063/5.0076731. URL https://doi.org/10.1063/5.0076731.

T. Igarashi. Characteristics of the flow around a square prism. *Bulletin of JSME*, 27 (231):1858–1865, 1984. doi: 10.1299/jsme1958.27.1858.

H.P.A.H. Irwin, K.R. Cooper, and R. Girard. Correction of distortion effects caused by tubing systems in measurements of fluctuating pressures. *Journal of Wind Engineering and Industrial Aerodynamics*, 5(1):93–107, 1979. ISSN 0167-6105. doi: https://doi.org/10.1016/0167-6105(79)90026-6. URL https://www.sciencedirect.com/science/article/pii/0167610579900266.

B. Jayaraman, S M Abdullah Al M., and Lu. Interplay of sensor quantity, placement and system dimension in pod-based sparse reconstruction of fluid flows. *Fluids*, 4: 109, 06 2019. doi: 10.3390/fluids4020109.

J. Ke. Rans and hybrid les/rans simulations of flow over a square cylinder. *Advances in Aerodynamics*, 1, 12 2019. doi: 10.1186/s42774-019-0012-9.

J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor. *Dynamic Mode Decomposition, Data-Driven Modeling of Complex Systems*. Society for Industrial and Applied Mathematics, Philadelphia, first edition, 2016.

B. E. Lee. The effect of turbulence on the surface pressure field of a square prism. *Journal of Fluid Mechanics*, 69(2):263–282, 1975. doi: 10.1017/S0022112075001437.

L.E.M. Lignarolo. On the turbulent mixing in horizontal axis wind turbine wakes. *PhD thesis, Delft university of Technology*, 2016. URL https://doi.org/10.4233/uuid:057fa33f-82a3-4139-beb8-53f184cd1d57.

K. Liu, B. Zhang, and Y. Zhou. Correlation between drag variation and rear surface pressure of an ahmed body. *Experiments in Fluids*, 62(6):1–20, 2021.

S. Lorenzi, A. Cammi, L. Luzzi, and G. Rozza. Pod-galerkin method for finite volume approximation of navier–stokes and rans equations. *Computer Methods in Applied Mechanics and Engineering*, 311:151–179, 2016. ISSN 0045-7825. doi: https://doi.org/10.1016/j.cma.2016.08.006. URL https://www.sciencedirect.com/science/article/pii/S0045782516308829.

J.L. Lumley. The structure of inhomogeneous turbulent flows. *Yaglom, A.M. and Tartarsky, V.I., Eds., Atmospheric Turbulence and Radio Wave Propagation*, 1967.

K. Manohar, B. Brunton, J. Kutz, and S. Brunton. Data-driven sparse sensor placement for reconstruction. *IEEE control systems*, 38:63–, 05 2018. doi: 10.1109/MCS.2018.2810460.

F. R. Menter. Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA journal*, 32(8):1598–1605, 1994.

M. Mifsud, A. Vendl, L. Hansen, and S. Görtz. Fusing wind-tunnel measurements and cfd data using constrained gappy proper orthogonal decomposition. *Aerospace Science and Technology*, 86:312–326, 2019. ISSN 1270-9638. doi: https://doi.org/10.1016/j.ast.2018.12.036. URL https://www.sciencedirect.com/science/article/pii/S1270963818301858.

A. Mohammed-Taifour and J. Weiss. Unsteadiness in a large turbulent separation bubble. *Journal of Fluid Mechanics*, 799:383–412, 2016. doi: 10.1017/jfm.2016.377.

T. L. Mohren, . L. Daniel, S. L. Brunton, and B. W. Brunton. Neural-inspired sensors enable sparse, efficient classification of spatiotemporal data. *Proceedings of the National Academy of Sciences*, 115(42):10564–10569, 2018. doi: 10.1073/pnas.1808909115. URL https://www.pnas.org/doi/abs/10.1073/pnas.1808909115.

A.A. Mueller. *Large Eddy Simulation of cross-flow around a square rod at incidence with application to tonal noise prediction*. PhD thesis, University of Twente, Netherlands, January 2012.

N. E. Murray and L. S. Ukeiley. An application of Gappy POD. For subsonic cavity flow PIV data. *Experiments in Fluids*, 42(1):79–91, January 2007. doi: 10.1007/s00348-006-0221-y.

A. Probst, R. Radespiel, C. Wolf, T. Knopp, and D. Schwamborn. A comparison of detached-eddy simulation and reynolds-stress modeling applied to the flow over a backward-facing step and an airfoil at stall. 01 2010. doi: 10.2514/6.2010-920.

A. Quarteroni and G. Rozza. *Reduced Order Methods for Modeling and Computational Reduction*. 01 2014. ISBN 978-3-319-02089-1. doi: 10.1007/978-3-319-02090-7.

S. Raben, J. Charonko, and P. Vlachos. Adaptive gappy proper orthogonal decomposition for particle image velocimetry data reconstruction. *Measurement Science and Technology*, 23:025303, 01 2012. doi: 10.1088/0957-0233/23/2/025303.

E.W.M. Roosenboom. Experimental analysis of the flow around a cylinder with a square cross-section. Master's thesis, Delft University of Technology, Netherlands, August 2005.

L. Sirovich. Turbulence and the dynamics of coherent structures part i : Coherent structures. 1987.

K. Taira, S. L. Brunton, S. T. M. Dawson, C. W. Rowley, T. Colonius, B. J. McKeon, O. T. Schmidt, S. Gordeyev, V. Theofilis, and L. S. Ukeiley. Modal analysis of fluid flows:

An overview. *AIAA Journal*, 55(12):4013–4041, 2017. doi: 10.2514/1.J056060. URL https://doi.org/10.2514/1.J056060.

T. Tamura and T. Miyagi. The effect of turbulence on aerodynamic forces on a square cylinder with various corner shapes. *Journal of Wind Engineering and Industrial Aerodynamics*, 83(1):135–145, 1999. ISSN 0167-6105. doi: https://doi.org/10.1016/S0167-6105(99)00067-7. URL https://www.sciencedirect.com/science/article/pii/S0167610599000677.

L.N. Trefethen and D. Bau. *Numerical Linear Algebra*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, 1997. ISBN 9780898713619. URL https://books.google.co.cr/books?id=bj-Lu6zjWbEC.

F.X. Trias, A. Gorobets, and A. Oliva. Turbulent flow around a square cylinder at reynolds number 22,000: A dns study. *Computers & Fluids*, 123:87–98, 2015. ISSN 0045-7930. doi: https://doi.org/10.1016/j.compfluid.2015.09.013. URL https://www.sciencedirect.com/science/article/pii/S0045793015003254.

B.W. van Oudheusden, F. Scarano, N.P. van Hinsberg, and E.W.M. Roosenboom. Quantitative visualization of the flow around a square-section cylinder at incidence. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(6):913–922, 2008. ISSN 0167-6105. doi: https://doi.org/10.1016/j.jweia.2007.06.030. URL https://www.sciencedirect.com/science/article/pii/S0167610507001407. 5th International Colloquium on Bluff Body Aerodynamics and Applications.

H. Wang, Q. Gao, L. Feng, W. Runjie, and J. Wang. Proper orthogonal decomposition based outlier correction for piv data. *Experiments in Fluids*, 56, 02 2015. doi: 10.1007/s00348-015-1894-x.

J. Weiss. A tutorial on the proper orthogonal decomposition. *AIAA Aviation Forum*, 17(21), 2019. doi: 10.2514/6.2019-3333. URL https://doi.org/10.2514/6.2019-3333.

K. Willcox. Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition. *Computers & Fluids*, 35(2):208–226, 2006. ISSN 0045-7930. doi: https://doi.org/10.1016/j.compfluid.2004.11.006. URL https://www.sciencedirect.com/science/article/pii/S0045793005000113.

J. Williams, O. Zahn, and J. N. Kutz. Data-driven sensor placement with shallow decoder networks, 2022. URL https://arxiv.org/abs/2202.05330.

X. Xing, M. Dao, B. Zhang, J. Lou, W. Siang Tan, Y. Cui, and B. Cheong Khoo. Fusing sensor data with cfd results using gappy pod. *Ocean Engineering*, 246:110549, 2022. ISSN 0029-8018. doi: https://doi.org/10.1016/j.oceaneng.2022.110549. URL https://www.sciencedirect.com/science/article/pii/S0029801822000269.

B. Yildirim, C. Chryssostomidis, and G.E. Karniadakis. Efficient sensor placement for ocean measurements using low-dimensional concepts. *Ocean Modelling*, 27(3-4):160–173, 2009.

Z.o Zhan, W. G. Habashi, and M. Fossati. Real-time regional jet comprehensive aeroicing analysis via reduced-order modeling. *AIAA Journal*, 54(12):3787–3802, 2016. doi: 10.2514/1.J055013. URL https://doi.org/10.2514/1.J055013.

C. Zhao, H. Wang, L. Zeng, Md. M. Alam, and X. Zhao. Effects of oncoming flow turbulence on the near wake and forces of a 3d square cylinder. *Journal of Wind Engineering and Industrial Aerodynamics*, 214:104674, 2021. ISSN 0167-6105. doi: https://doi.org/10.1016/j.jweia.2021.104674. URL https://www.sciencedirect.com/science/article/pii/S0167610521001586.

A

Cosine results

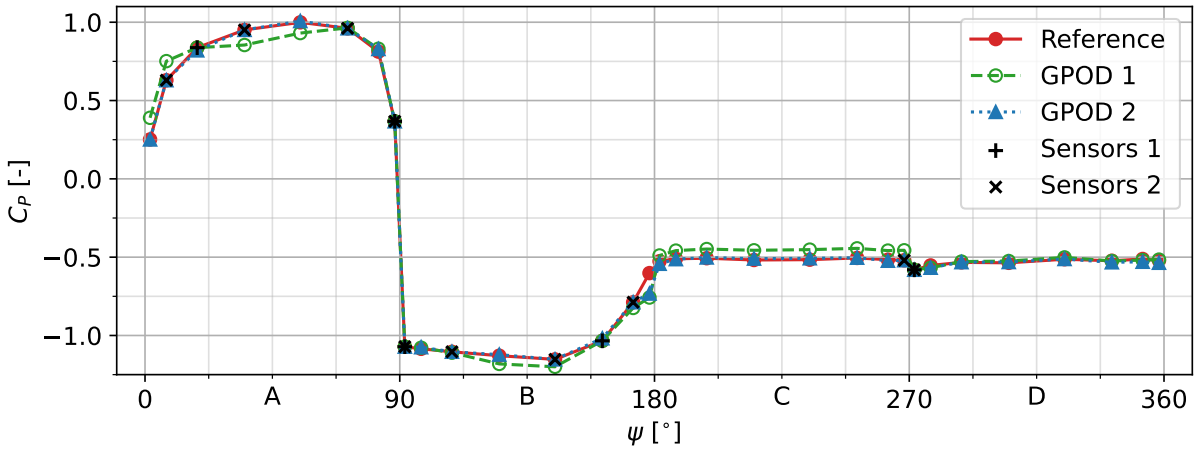


Figure A.1: Experimental GPOD reconstruction at $\alpha = 12.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$, using cosine spacing pressure distributions.

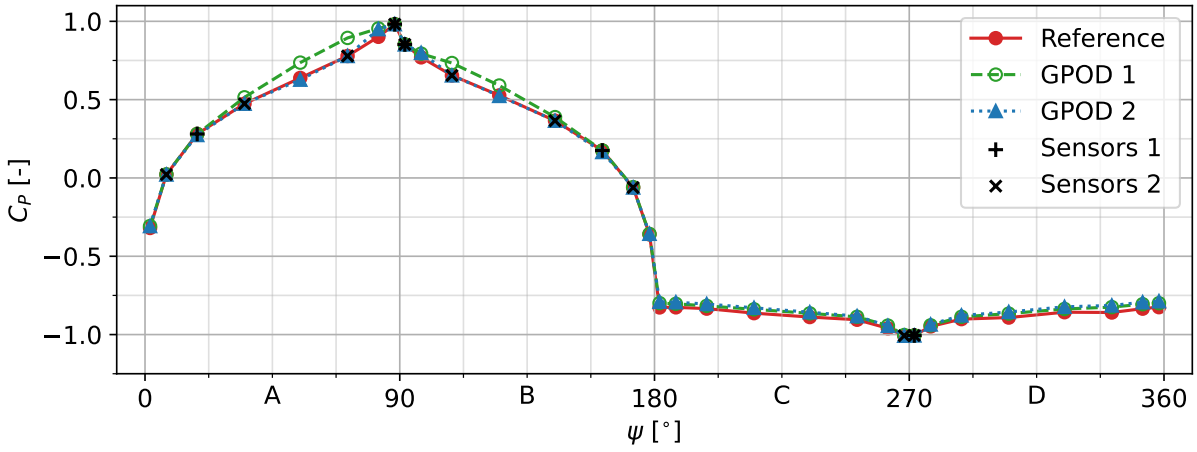


Figure A.2: Experimental GPOD reconstruction at $\alpha = 42.5°$. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$, using cosine spacing pressure distributions.

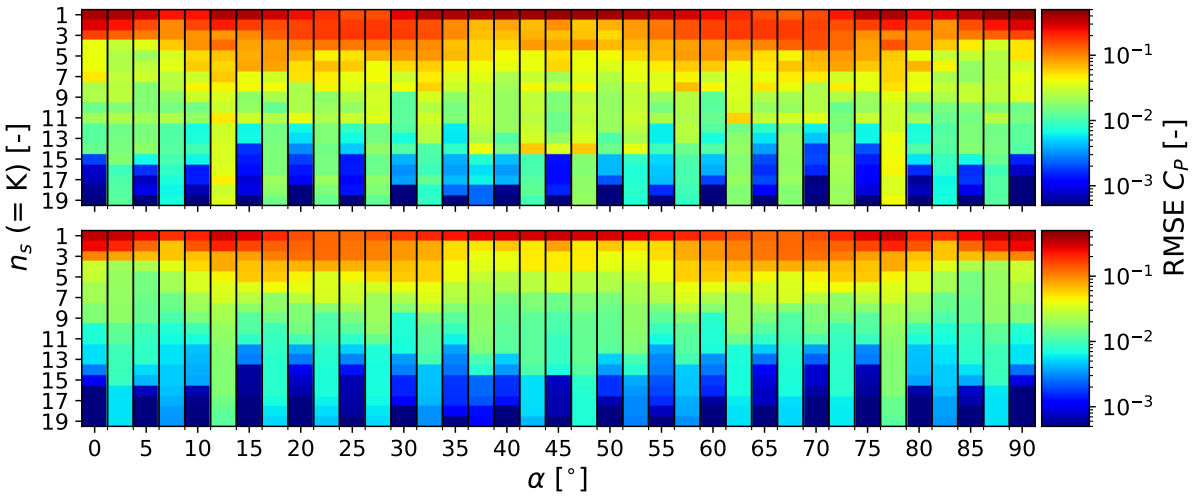


Figure A.3: Experimental GPOD pressure distribution reconstruction (top) and projection (bottom) accuracy for every combination of $n_s = K$ and α, using cosine spacing pressure distributions.

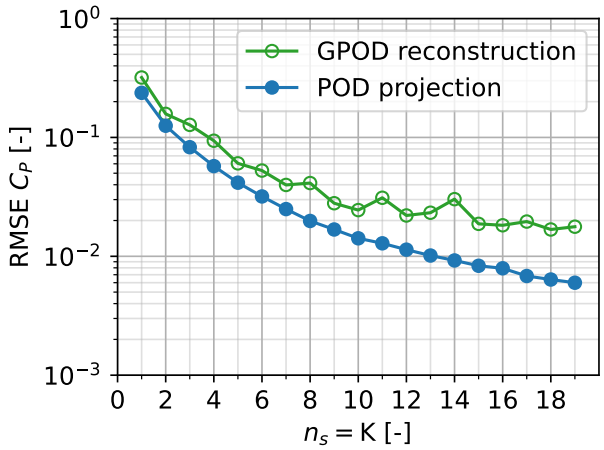


Figure A.4: Experimental GPOD surface pressure reconstruction accuracy for increasing sensor budget, using cosine spacing pressure distributions.

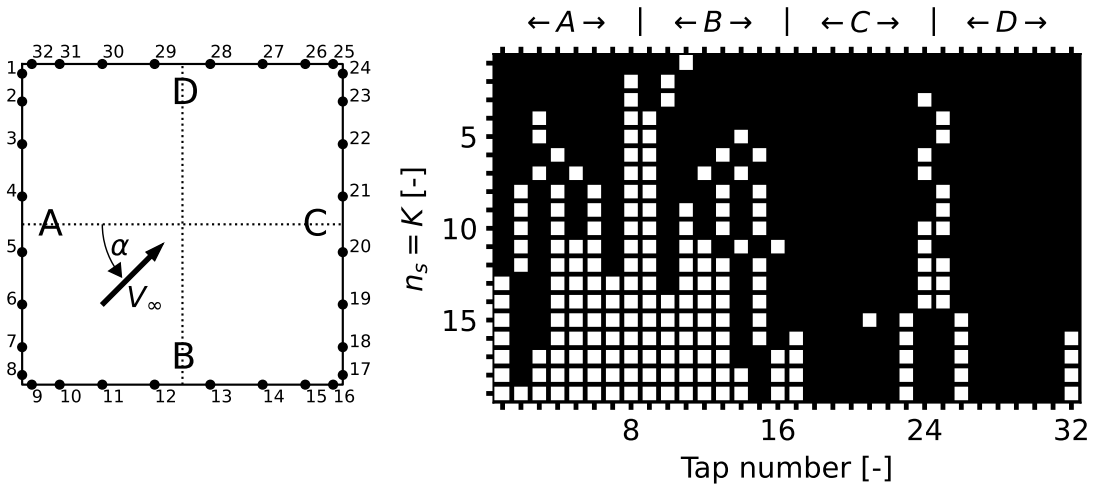


Figure A.5: Sensor locations determined by QR algorithm for each value of $n_s = K$, using cosine spacing pressure distributions.

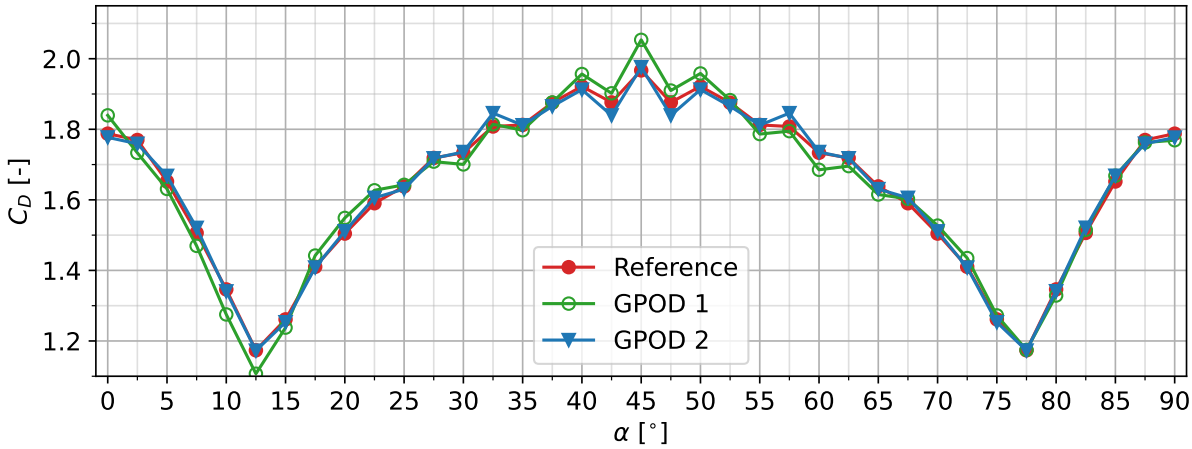


Figure A.6: Experimental GPOD drag curve reconstruction. GPOD 1: $n_s = K = 5$, GPOD 2: $n_s = K = 10$, using cosine spacing pressure distributions.

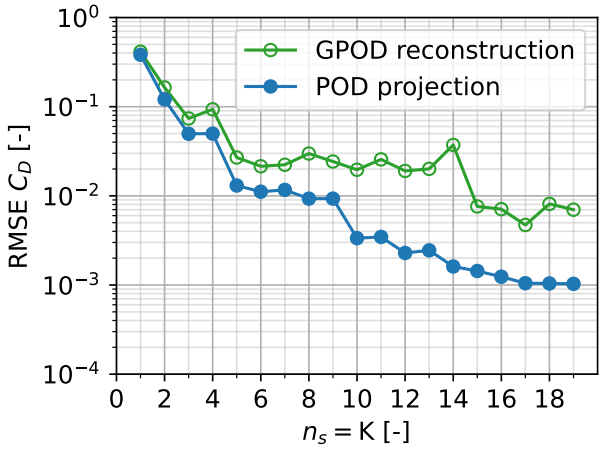


Figure A.7: Experimental GPOD drag curve reconstruction accuracy for increasing sensor budget, using cosine spacing pressure distributions.

B

Experimental even spacing POD modes and coefficients

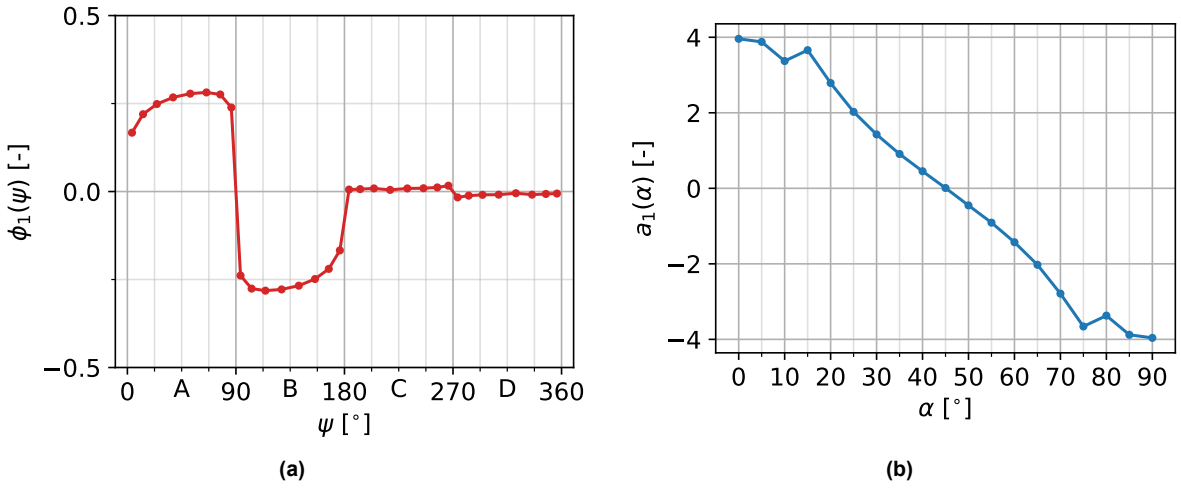


Figure B.1: Spatial POD mode and α-coefficients 1 from pressure distributions using even tap spacing

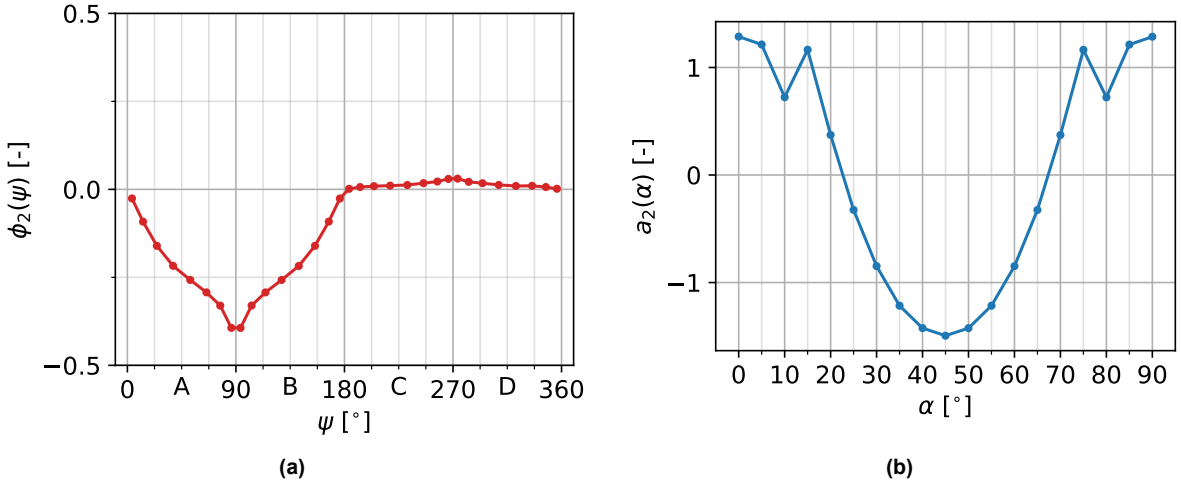


Figure B.2: Spatial POD mode and α-coefficients 2 from pressure distributions using even tap spacing

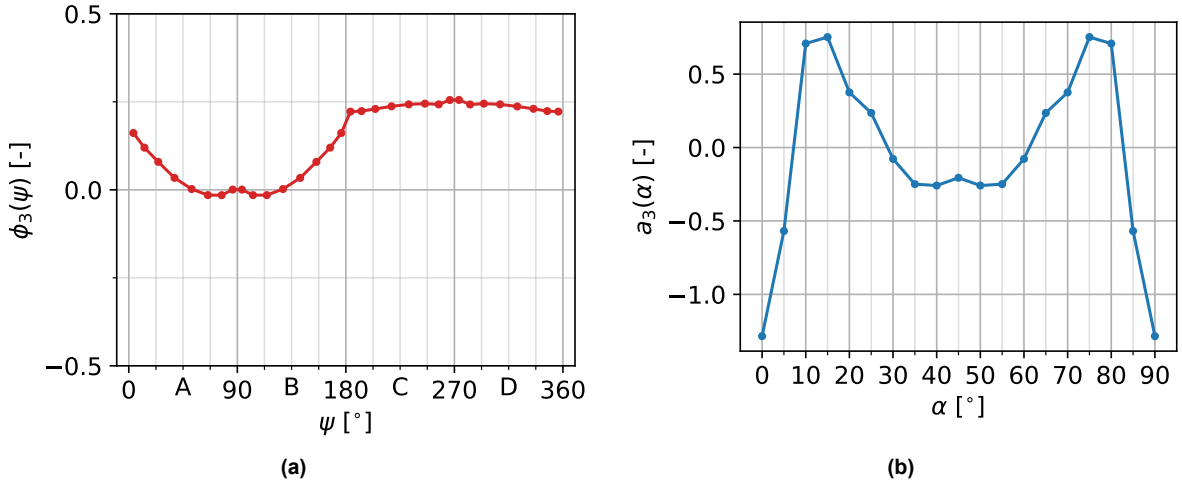


Figure B.3: Spatial POD mode and α-coefficients 3 from pressure distributions using even tap spacing

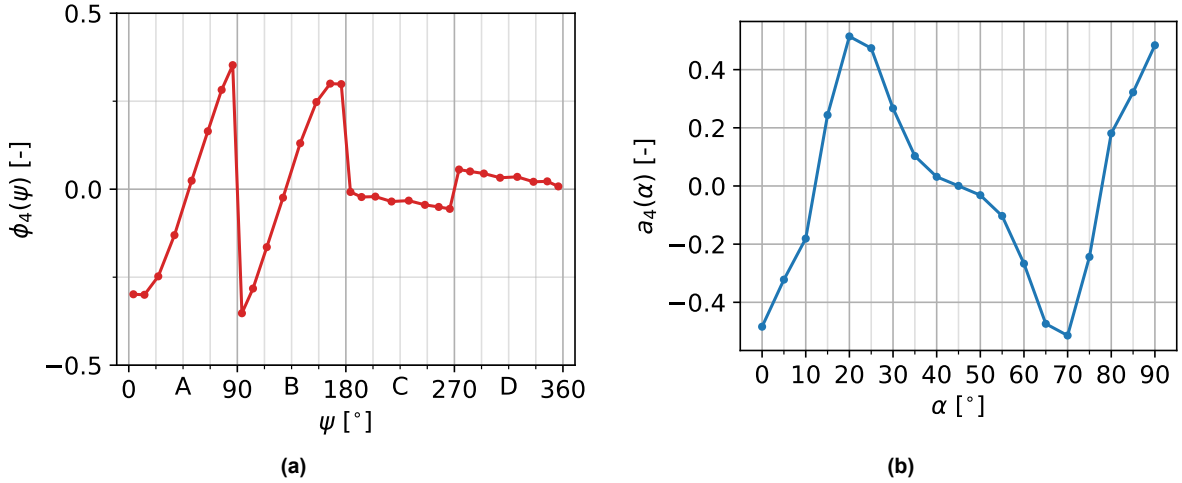


Figure B.4: Spatial POD mode and α-coefficients 4 from pressure distributions using even tap spacing

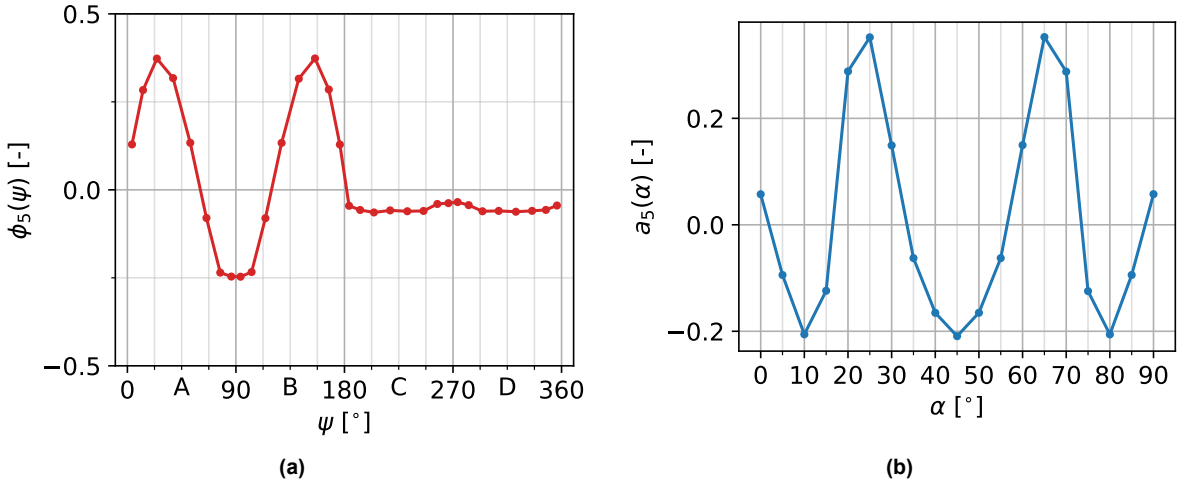


Figure B.5: Spatial POD mode and α-coefficients 5 from pressure distributions using even tap spacing

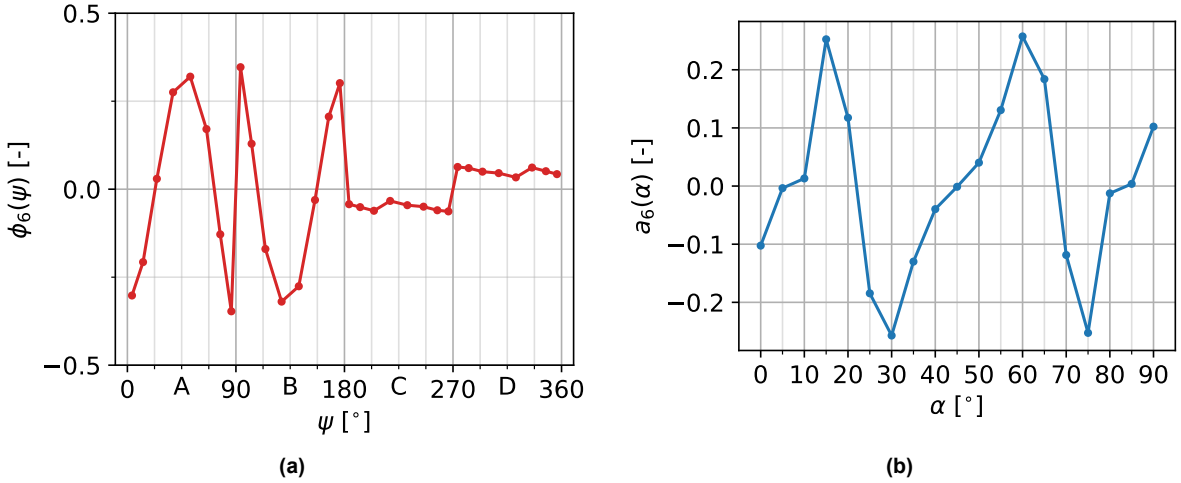
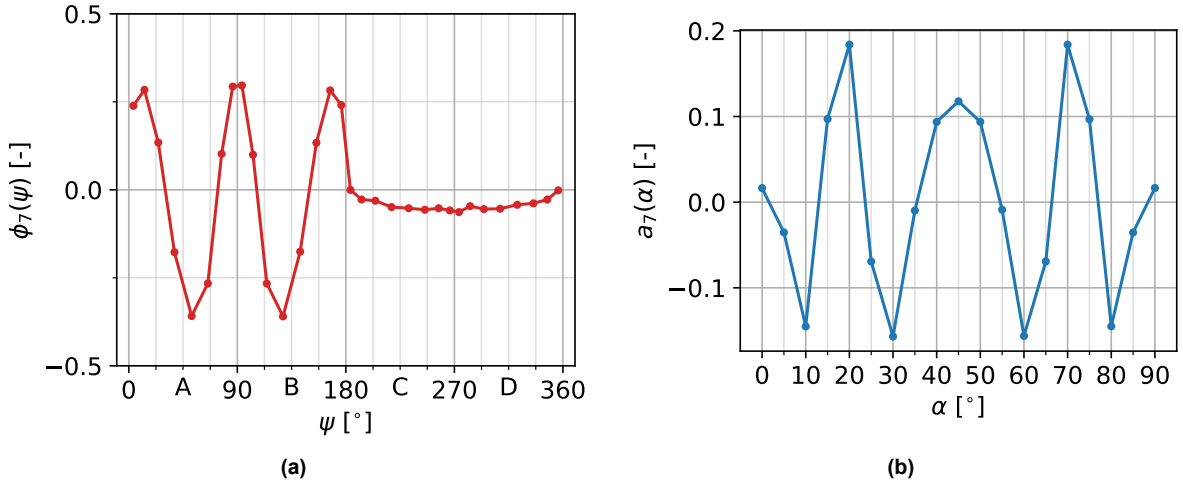
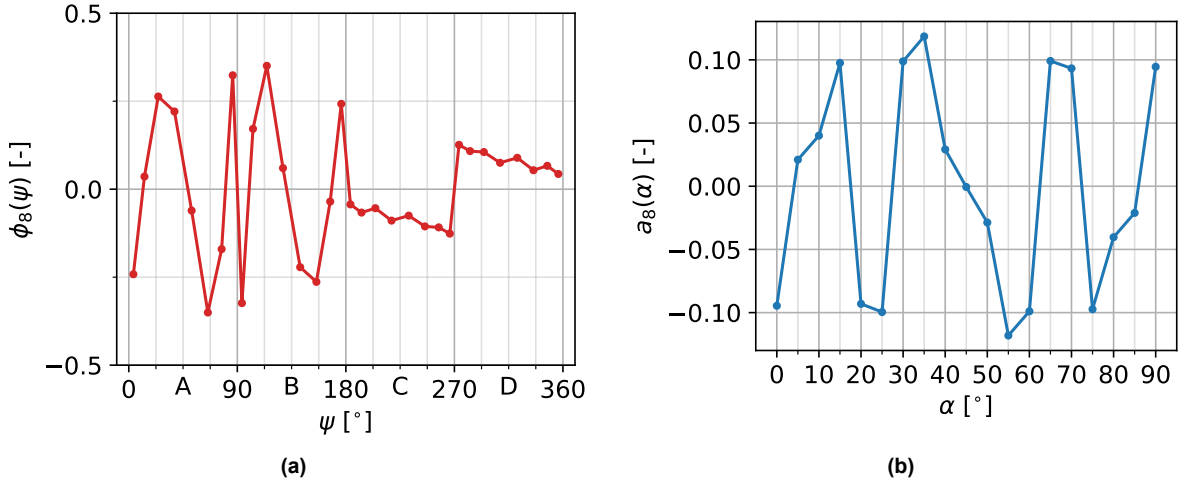


Figure B.6: Spatial POD mode and α-coefficients 6 from pressure distributions using even tap spacing



(a) (b)

Figure B.7: Spatial POD mode and α-coefficients 7 from pressure distributions using even tap spacing



(a) (b)

Figure B.8: Spatial POD mode and α-coefficients 8 from pressure distributions using even tap spacing

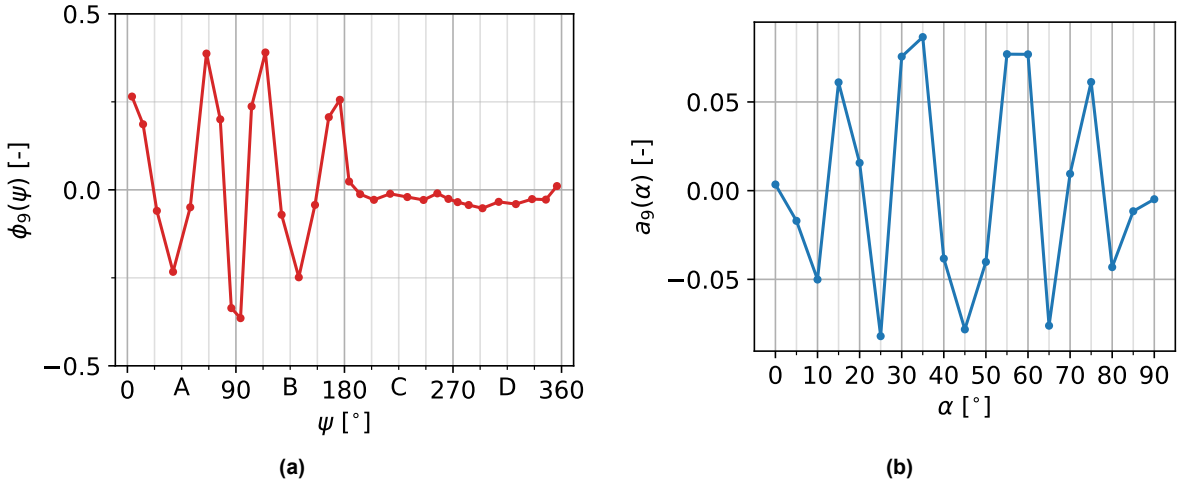


Figure B.9: Spatial POD mode and α-coefficients 9 from pressure distributions using even tap spacing

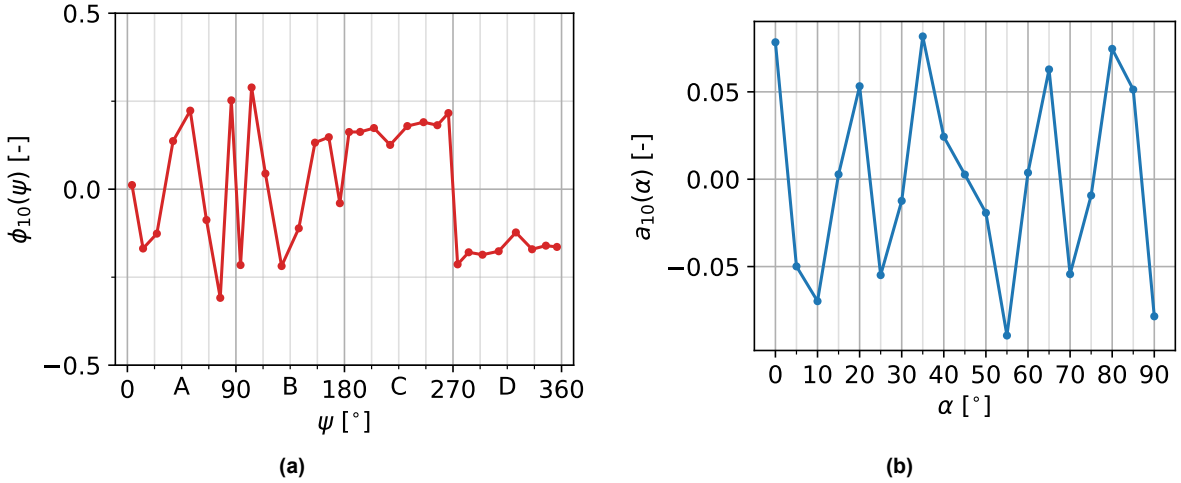
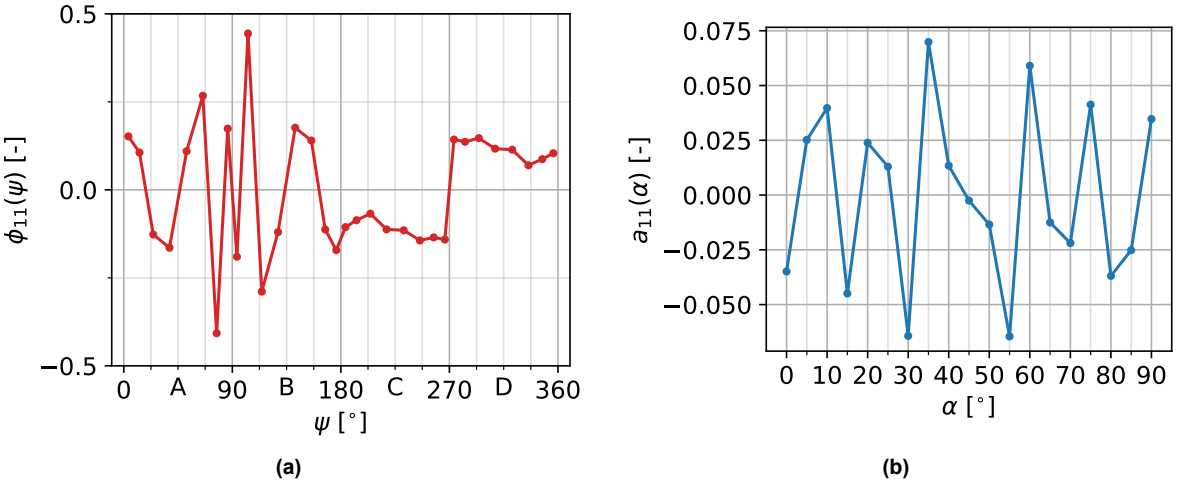
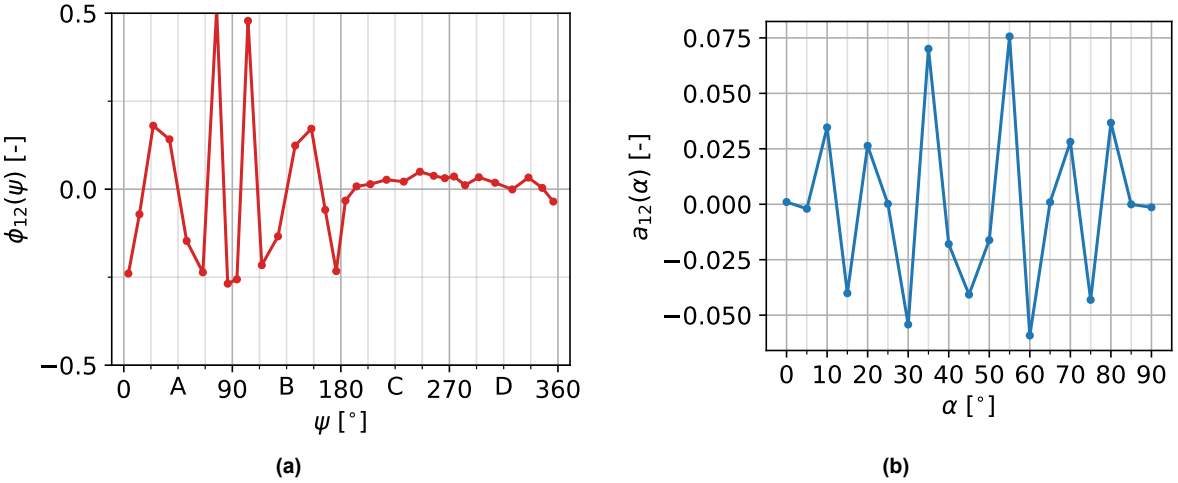


Figure B.10: Spatial POD mode and α-coefficients 10 from pressure distributions using even tap spacing



(a) (b)

Figure B.11: Spatial POD mode and α-coefficients 11 from pressure distributions using even tap spacing



(a) (b)

Figure B.12: Spatial POD mode and α-coefficients 12 from pressure distributions using even tap spacing

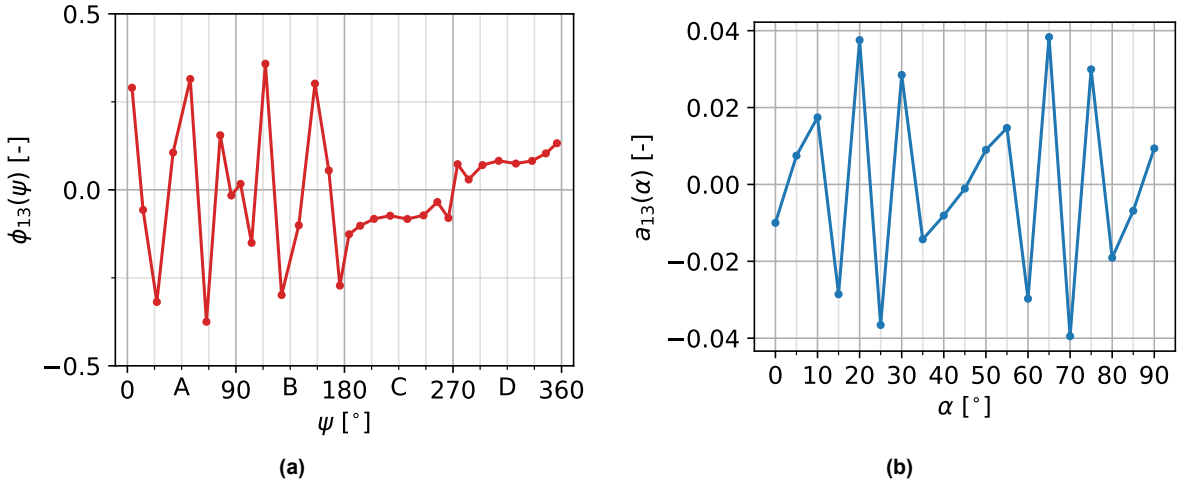


Figure B.13: Spatial POD mode and α-coefficients 13 from pressure distributions using even tap spacing

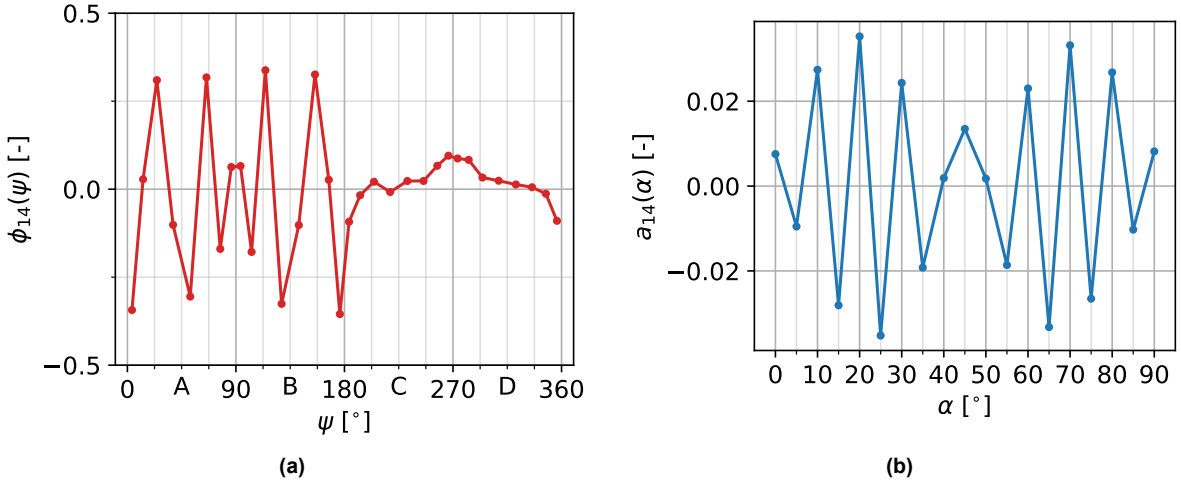
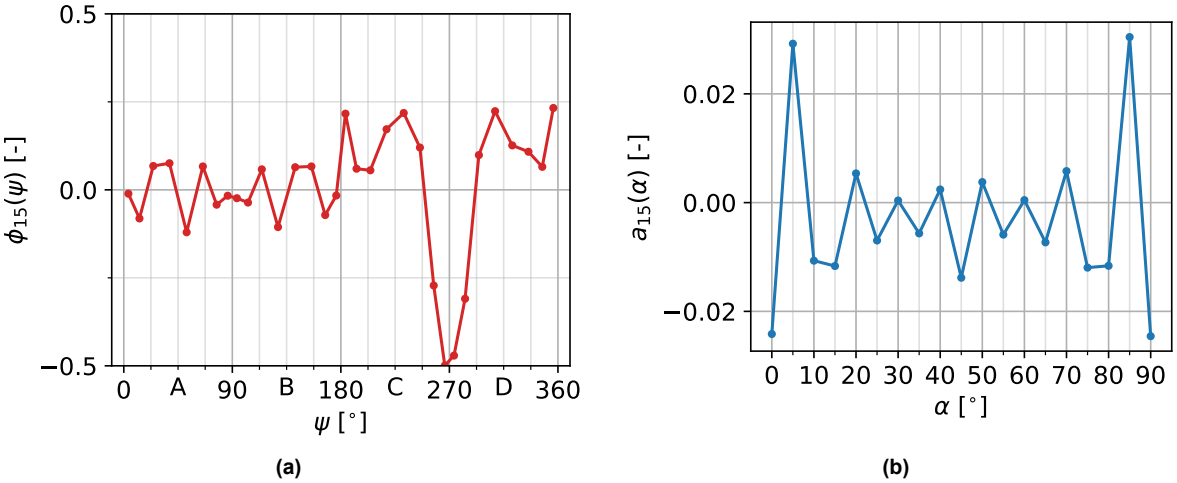
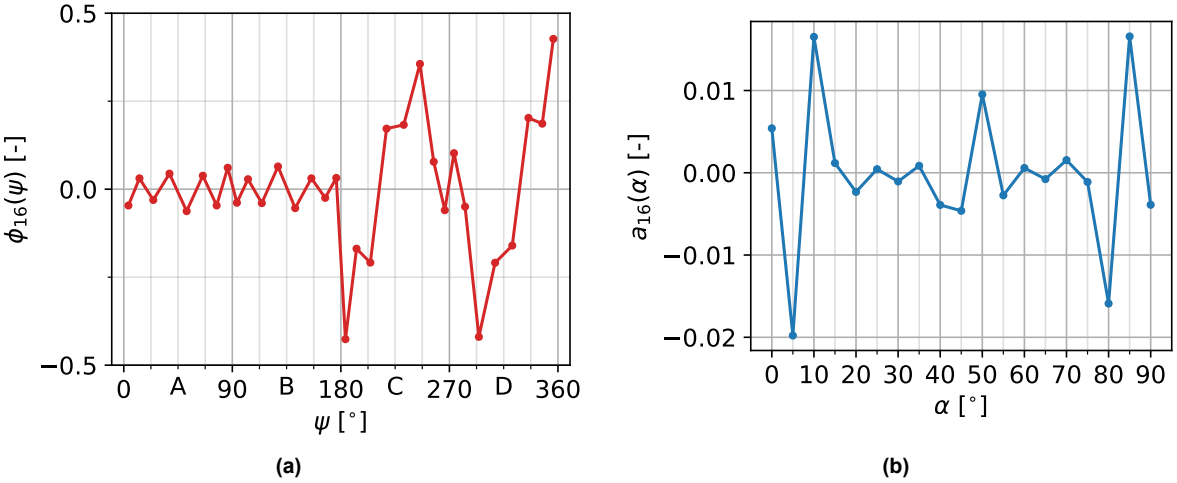


Figure B.14: Spatial POD mode and α-coefficients 14 from pressure distributions using even tap spacing



(a) (b)

Figure B.15: Spatial POD mode and α-coefficients 15 from pressure distributions using even tap spacing



(a) (b)

Figure B.16: Spatial POD mode and α-coefficients 16 from pressure distributions using even tap spacing

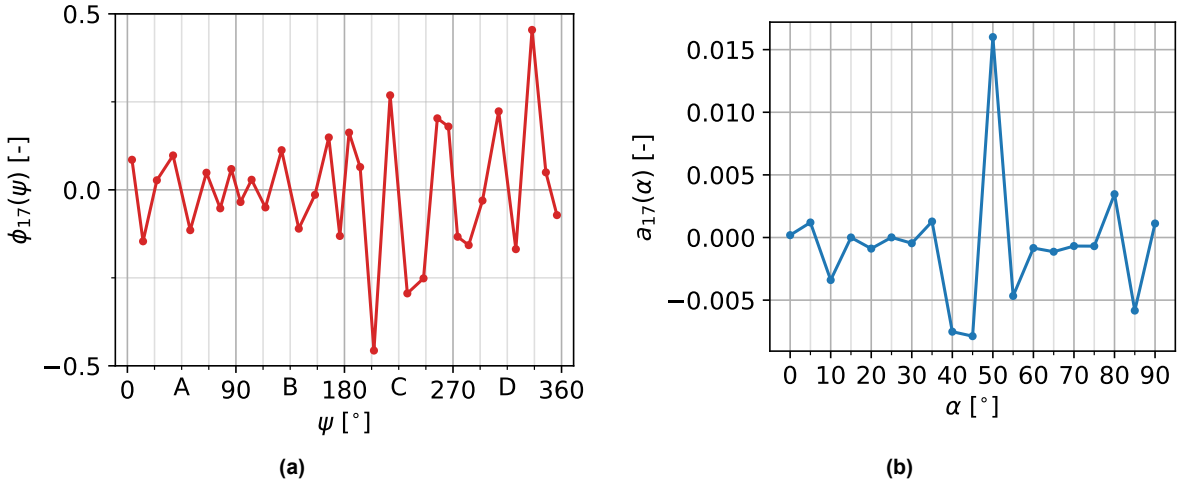


Figure B.17: Spatial POD mode and α-coefficients 17 from pressure distributions using even tap spacing

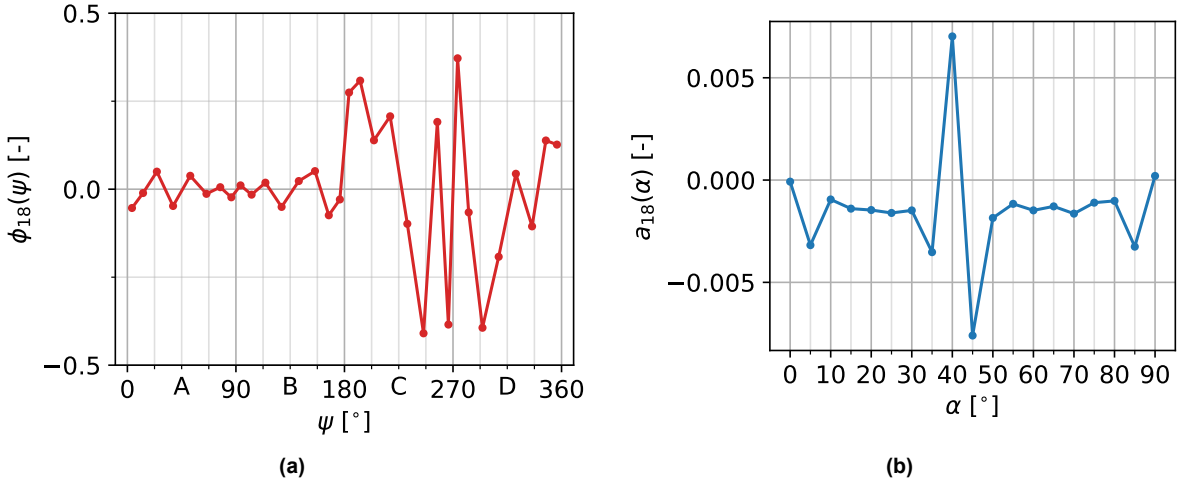


Figure B.18: Spatial POD mode and α-coefficients 18 from pressure distributions using even tap spacing

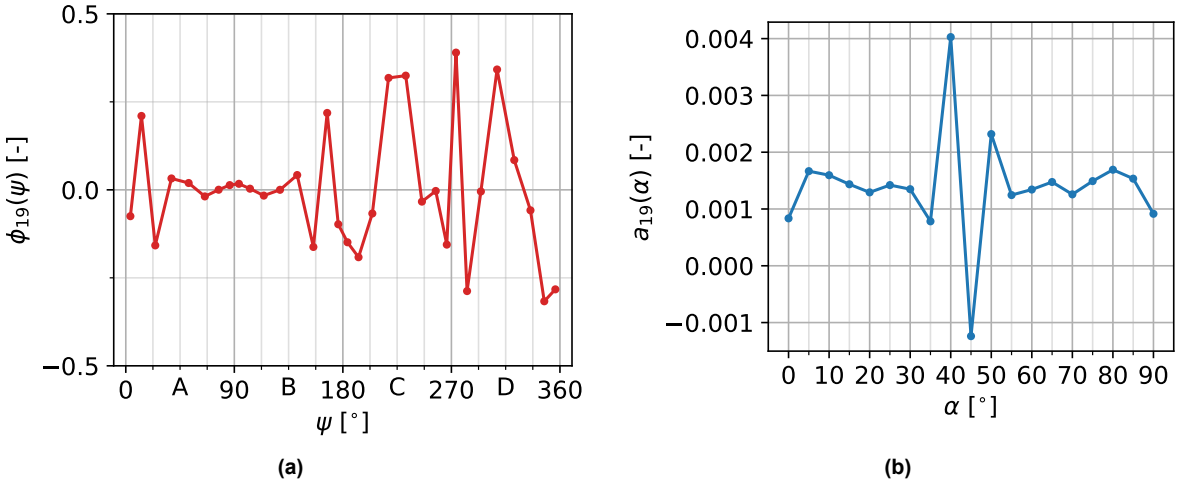


Figure B.19: Spatial POD mode and α-coefficients 19 from pressure distributions using even tap spacing

C

GPOD source code

```
import numpy as np
import matplotlib.pyplot as plt


#Snapshot POD function, input data shape m by n, return coefficients, bases and eigenvalues
def Snapshot_POD(training_data):
    m_t, n_t = np.shape(training_data)

    training_data_T = np.transpose(training_data)
    Cs = (1/(m_t - 1))*(training_data @ training_data_T) #Correlation matrix
    Eigvals, Eigvecs = np.linalg.eigh(Cs) #Eigenvalue and vectors
    Eigvals = np.real(Eigvals)
    Eigvecs = np.real(Eigvecs)

    #Sort eigenvalues and vectors based on magnitude
    idx = Eigvals.argsort()[::-1]
    Eigvals = Eigvals[idx]
    Eigvecs = Eigvecs[:,idx]

    #Compute POD coefficients and modes
    As = Eigvecs
    Phis = training_data_T @ As

    #Optional: normalize modes and switch temporal/spatial modes
    Norms = np.linalg.norm(Phis,2,axis = 0)
    Phis /= Norms
    As = training_data @ Phis

    return As, Phis, Eigvals

#GPOD reconstructions, input data to be used for training, reference data from which data originates, number of sensors and rank of the reconstructions
def GPOD(training_data, reconstruction_data, p, k):
    m_r, n_r = np.shape(reconstruction_data)

    As, Phis, Eigvals = Snapshot_POD(training_data) #Obtain POD basis from POD function on training data

```

```
37      #QR-CP algorithm for the determination of sensor locations
38      if p > k:
39          Q,R,P = linalg.qr(Phis[:,0:k] @ np.transpose(Phis[:,0:k]),
                pivoting = True)
40      else:
41          Q,R,P = linalg.qr(np.transpose(Phis[:,0:k]), pivoting = True)
42
43      sensor_locs = P[0:p]
44      sensor_locs = sensor_locs[sensor_locs.argsort()]
45
46      reconstructed_data = np.zeros(np.shape(reconstruction_data))
47
48      nPhis = Phis[sensor_locs, 0:k]
49      nPhisT = np.transpose(nPhis)
50      M = nPhisT @ nPhis #GPOD matrix
51
52      #Reconstruct all data provided as original reference data using GPOD
53      for i in range(m_r):
54          g = reconstruction_data[i, sensor_locs][:,None]  #Sparse pressure
                distribution
55          f = nPhisT @ g #RHS linear system
56          b = np.linalg.solve(M, f) #GPOD coefficients from linear system Mb
                = f
57          g_tilde = Phis[:,0:k] @ b #Approximation of original pressure
                distribution
58
59          reconstructed_snapshot = g_tilde
60          reconstructed_snapshot = np.transpose(reconstructed_snapshot)
61          reconstructed_data[i,:] = reconstructed_snapshot
62
63      return reconstructed_data, sensor_locs, As, Phis, Eigvals, np.linalg.
            cond(M)
64
65  #Pressure drag coefficient square cylinder function, input x, y
        coordinates of pressure taps, surface pressure measurements and angle
        of attack, output Cd and Cl
66  #Integrates pressure per cylinder face using trapezoidal rule, assuming
        constant pressure from taps closest to the corners untill the corners
67  #Some additional modifications added to be able to function when sparse
        pressure distributions are provided
68  def integ_V1(x, y, Cp, alpha):
69      angles = np.rad2deg(np.arctan2(y,x))  + 180 + 45
70      angles[np.where(angles >= 360)] -= 360
71
72      plate_1 = np.where(angles < 90)
73      plate_2 = np.where(np.logical_and(angles > 90, angles < 180))
74      plate_3 = np.where(np.logical_and(angles > 180, angles < 270))
75      plate_4 = np.where(angles > 270)
76
77      if len(plate_1[0]) == 0:
78          x_1 = np.zeros(len(x) + 2)
79          y_1 = np.zeros(len(x) + 2)
80          Cp_1 = np.zeros(len(x) + 2)
81      else:
82          x_1 = np.hstack((-(L/2), x[plate_1], -(L/2)))
83          y_1 = np.hstack(((L/2), y[plate_1], -(L/2)))
```

```
84          Cp_1 = np.hstack((Cp[plate_1][0], Cp[plate_1], Cp[plate_1][-1]))
85
86      if len(plate_2[0]) == 0:
87          x_2 = np.zeros(len(x) + 2)
88          y_2 = np.zeros(len(x) + 2)
89          Cp_2 = np.zeros(len(x) + 2)
90      else:
91          x_2 = np.hstack((-(L/2), x[plate_2], (L/2)))
92          y_2 = np.hstack((-(L/2), y[plate_2], -(L/2)))
93          Cp_2 = np.hstack((Cp[plate_2][0], Cp[plate_2], Cp[plate_2][-1]))
94
95      if len(plate_3[0]) == 0:
96          x_3 = np.zeros(len(x) + 2)
97          y_3 = np.zeros(len(x) + 2)
98          Cp_3 = np.zeros(len(x) + 2)
99      else:
100         x_3 = np.hstack(((L/2), x[plate_3], (L/2)))
101         y_3 = np.hstack((-(L/2), y[plate_3], (L/2)))
102         Cp_3 = np.hstack((Cp[plate_3][0], Cp[plate_3], Cp[plate_3][-1]))
103
104     if len(plate_4[0]) == 0:
105         x_4 = np.zeros(len(x) + 2)
106         y_4 = np.zeros(len(x) + 2)
107         Cp_4 = np.zeros(len(x) + 2)
108     else:
109         x_4 = np.hstack(((L/2), x[plate_4], -(L/2)))
110         y_4 = np.hstack(((L/2), y[plate_4], (L/2)))
111         Cp_4 = np.hstack((Cp[plate_4][0], Cp[plate_4], Cp[plate_4][-1]))
112
113     X = np.hstack((x_1, x_2, x_3, x_4))
114     Y = np.hstack((y_1, y_2, y_3, y_4))
115     PSI = np.rad2deg(np.arctan2(Y,X))  + 180 + 45
116     PSI[np.where(PSI >= 360)] -= 360
117
118     force_1 = np.trapz(Cp_1, y_1)*-1
119     force_2 = np.trapz(Cp_2, x_2)
120     force_3 = np.trapz(Cp_3, y_3)
121     force_4 = np.trapz(Cp_4, x_4)*-1
122
123     force_x_prime = force_1 - force_3
124     force_y_prime = force_2 - force_4
125     force_x = force_x_prime*np.cos(np.deg2rad(alpha)) + force_y_prime*np.
            cos(np.deg2rad(90 - alpha))
126     force_y = -1*force_x_prime*np.cos(np.deg2rad(90 - alpha)) +
            force_y_prime*np.cos(np.deg2rad(alpha))
127     Cd = force_x/(L)
128     Cl = force_y/(L)
129
130     return Cd, Cl
131
132
133 filepath = 'C:\\Users\\LHendriksen\\Documents\\TUD\\Thesis\\Luuk_OJF_June'
        #Specify filepath to foldercontaining data
134 data = np.load(filepath + '\\Even_averaged.npy')    #Specify filename of
        file containing data
135
```

```
136 Exp_Cp = data[0:-2,::]  #First untill third to last row contain pressure
    distributions
137
138 # Last two rows contain x and y coordinates of pressure taps
139 x = data[-2,::]
140 y = data[-1,::]
141
142 L = 0.2     #Square cylinder sidelength
143
144 m,n = np.shape(Exp_Cp) #Data dimensions
145
146 #Computing tap locations in psi from x and y
147 psi = np.rad2deg(np.arctan2(y,x))  + 225
148 psi[np.where(psi >= 360)] -= 360
149
150 Cp_mean = np.mean(Exp_Cp, axis = 0)[None,::]    #Average pressure
    distribution over angles of attack
151
152 #All angles of attack and those to be used for training
153 alphas = np.linspace(0,90,37)
154 alphas_training = np.linspace(0,90,19)
155
156 alphas_training_i = np.zeros(len(alphas_training)).astype(int)
157
158 num_sensors = 10 #Specify number of sensors to be used for reconstructions
     (n_s)
159
160 for i in range(len(alphas_training)):
161     alphas_training_i[i] = np.where(abs(alphas - alphas_training[i]) <
        0.0001)[0][0]
162
163 reconstruction_rank = num_sensors #Reconstruction rank K is equal to
    number of sensors n_s
164
165 Exp_Cp_training = Exp_Cp[alphas_training_i,::] - Cp_mean #Subtract mean
    from training
166 Exp_Cp_reconstruct = Exp_Cp - Cp_mean #Subtract mean from all data
167
168 Exp_Cp_reconstructed, sensor_locs, As, Phis, Eigvals, kappa_M = GPOD(
    Exp_Cp_training, Exp_Cp_reconstruct, num_sensors, reconstruction_rank)
    #Perform GPOD reconstruction
169 Exp_Cp_projected = GPOD(Exp_Cp_training, Exp_Cp_reconstruct, n,
    reconstruction_rank)[0] #Perform POD projection
170
171 #Add mean back
172 Exp_Cp_reconstructed += Cp_mean #Exp_Cp_reconstructed now contains the
    GPOD approximation of all original data
173 Exp_Cp_reconstruct += Cp_mean
174 Exp_Cp_projected += Cp_mean
175
176 Cd_reconstructed = np.zeros(len(alphas))
177 Cd_exact = np.zeros(len(alphas))
178 Cd_projected = np.zeros(len(alphas))
179
180 #Compute pressure drag for exact, reconstructions and projections
181 for j in range(len(alphas)):
```

```
182     Cd_reconstructed[j] = integ_V1(x, y, Exp_Cp_reconstructed[j,::],
            alphas[j])[0]
183     Cd_exact[j] = integ_V1(x, y, Exp_Cp[j,::], alphas[j])[0]
184     Cd_projected[j] = integ_V1(x, y, Exp_Cp_projected[j,::], alphas[j])[0]
```