

IMPERIAL

**Mean flow reconstruction of turbulent
flows from sparse measurements using
Physics-Informed Neural Networks**

Yusuf Salahudeen Patel

February 27, 2025

Supervised by Dr. Georgios Rigas

Department of Aeronautics,
Imperial College London

Submitted in part fulfilment of the requirements for the degree of Doctor of
Philosophy in Aeronautics of Imperial College London and the Diploma of
Imperial College London

Abstract

Optimal aerodynamic design is guided by numerical simulation and experimental methods. For flows of engineering interest, numerical simulation is characterised by a trade-off between accuracy and tractability, whilst experimental methods are restricted by the inability to measure all spatio-temporal scales associated with turbulent dynamics. In this thesis, data assimilation (DA) methods, specifically Physics-Informed Neural Networks (PINNs), are explored, to augment low-fidelity numerical approaches (Reynolds-Averaged Navier-Stokes (RANS) equations) with sparse high-fidelity experimental data to accurately reconstruct turbulent flows.

Firstly, PINNs are benchmarked for mean-flow reconstruction using sparse mean velocity measurements against variational-DA. The PINN-DA-SA outperforms adjoint-based variational-DA, attributed to discretisation and regularisation errors in the variational formulation. Introducing an empirical closure model (Spalart-Allmaras) and corrective forcing with turbulence-model augmented PINNs, increase reconstruction accuracy in high velocity gradient and separation regions.

Secondly, a hierarchical PINN approach is proposed for constraining the flow physics with equations that systematically interpolate between direct numerical simulation of the Navier-Stokes (high-fidelity; resolved spectrum) and RANS (low-fidelity; modelled spectrum). We introduce PINNs constrained by the Navier-Stokes expressed in the frequency domain (Harmonic-Balanced Navier-Stokes (HBNS) equations) to resolve low-frequency harmonics and infer residual closure terms, without empirical turbulence models errors. PINN-DA-HBNS accurately reconstructs mean flow and predicts the skin-friction coefficient of the laminar circular cylinder better than RANS-constrained PINNs. Furthermore, use of PINN-DA is extended by building a neural network-based HBNS closure model, mapping low-frequency resolved quantities to residual closures.

Finally, the PINN-DA formulations are used to reconstruct a turbulent, axisymmetric bluff body wake from sparse experimental velocity data. Misaligned velocity measurements significantly affect flow reconstruction accuracy, leading to the proposal of transformation-embedded PINNs, to correct spatially-misaligned measurements. Using transformation-embedded PINN-DA, the bluff body wake flow is accurately reconstructed from sparse velocity measurements, and pressure is inferred away from the wall, where it cannot be measured non-intrusively.

Publications

The following work was published in

- Patel, Y., Mons, V., Marquet, O. & Rigas, G. Mar. 2024 Turbulence model augmented physics-informed neural networks for mean-flow reconstruction Phys. Rev. Fluids **9**, 034605 (Published)
- Patel, Y. & Rigas, G. Harmonic-Balanced Navier-Stokes constrained PINNs for accurate mean-flow reconstruction and HBNS closure modelling (To be submitted)
- Patel, Y. & Rigas, G. Mean-flow reconstruction of turbulent bluff body wake from spatially-misaligned, sparse experimental data using transformation-embedded PINNs (In preparation)

The following work was presented at:

- IUTAM Symposium on Data-driven modelling and optimization in fluid mechanics, Aarhus, 15-17 June 2022
- ICCS Machine Learning and Data Assimilation for Dynamical Systems, London, 21-23 June 2022
- APS 75th Annual Meeting of the Division of Fluid Dynamics, Indianapolis, 20-22 November 2022
- 14th International ERCOFTAC Symposium on Engineering, Turbulence Modelling and Measurements, Barcelona, 06-08 September 2023
- ERCOFTAC Workshop SIG47 on 3D Wakes, London, 10-12 April 2024

Statement of originality and copyright declaration

I hereby declare that all research presented in this thesis is the author's. All other contributions have been acknowledged and appropriately referenced.

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

Firstly, I would like to express my gratitude to my PhD supervisor, Dr Georgios Rigas. After encouraging me to pursue a PhD four years ago, he has provided support, ideas and guidance over the course of my PhD. Under his supervision, I have learnt much about how to approach and solve problems and effectively communicate my ideas.

I would also like to thank my corroborators from ONERA, Dr Olivier Marquet and Dr Vincent Mons, for providing their expertise and performing the variational simulations used in this research. Their insight and knowledge helped contribute to many interesting findings. I would also like to give additional thanks to Dr Taihang Zhu for providing data from his bluff body wake experiments.

Finally, and most importantly, I would like to thank my family for their support both during my PhD and growing up. Their sacrifices and guidance have provided me with the best possible platform to achieve my ambitions.

Nomenclature

$2D2C$	Two dimensional, two component
$\tilde{\cdot}$	Reconstructed or predicted flow field
ΔL	Spacing between measurements
$\gamma_{n_1}^{n_2}$	Coefficient where $\gamma_{n_1}^{n_2} = 0.5$ if $n_1 = n_2$, else 1
$\hat{\cdot}$	Harmonic, Fourier mode
$\hat{p}_0 = P$	Mean pressure field
$\lambda^D, \lambda^P, \lambda^B, \lambda^R$	Loss weighting for data, PDE, boundary condition and regularisation loss respectively
$\hat{\mathbf{u}}_0 = \mathbf{U} = U_i = (U, V, W)$	Mean velocity vector
$\hat{\mathbf{u}}_l, \hat{\mathbf{u}}_h$	Velocity modes divided into low-frequency resolved and high-frequency unresolved components respectively
$\hat{\mathbf{w}}_0 = [\hat{\mathbf{u}}_0, \hat{p}_0] = [\mathbf{U}, P]$	Mean flow field vector
$\hat{\mathbf{w}}_n = [\hat{\mathbf{u}}_n, \hat{p}_n]$	n -th harmonic frequency mode for flow field vector
$\nabla_s = (\nabla + \nabla^T)$	Strain rate tensor operator
$\nabla\phi$	Scalar, potential forcing field
$\tau = \tau_{ij} = \overline{u_i' u_j'} = \overline{\mathbf{u}' \mathbf{u}'}$	Reynolds stress tensor

θ	Neural network weights and biases
$\mathbf{f}_c$	Corrective forcing field
$\mathbf{f}_s$	Divergence-free, solenoidal, vectorial forcing field
$\mathbf{f} = f_i \equiv \nabla \cdot \tau$	Reynolds forcing vector
$\mathbf{L}$	Stokes operator
$\mathbf{M}$	Mass matrix
$\mathbf{N}$	Symmetrised non-linear convection operator
$\mathbf{R}_{n,s}$	Divergence-free, solenoidal, residual closure term for the n -th harmonic momentum equation
$\mathbf{R}_n$	Residual closure term for the n -th harmonic momentum equation
$\mathbf{T}_v$	Transformation matrix for velocity fields
$\mathbf{T}$	Transformation matrix for measurement locations
$\mathbf{u}'$	Fluctuating velocity vector from Reynolds decomposition, such that $\mathbf{u} = \mathbf{U} + \mathbf{u}'$
$\mathbf{U}_m = U_{m,i}$	Mean velocity measurements corresponding to $\mathbf{x}_m$
$\mathbf{u} = u_i = (u, v, w)$	Instantaneous velocity vector
$\mathbf{w}_l, \mathbf{w}_h$	Flow field divided into a low-frequency and high-frequency components respectively
$\mathbf{w} = [\mathbf{u}, p]$	Instantaneous flow field vector
$\mathbf{x}_m$	Location of data velocity measurements
$\mathbf{x} = (x, y, z)$	Spatial coordinates
$\mathcal{L}$	Lagrangian operator

$\mathcal{L}_P, \mathcal{N}_P$	Projected linear and non-linear operators respectively
$\mathcal{P}_l, \mathcal{P}_h$	Projection operators (sharp bypass filter) to project flow onto a low- and high-frequency subspace respectively
HBNS _N	Harmonic-Balanced Navier-Stokes operator for N resolved harmonics
ν	Kinematic viscosity
ν_t	Eddy viscosity
ω	Fundamental frequency
$\tilde{\nu}$	Pseudo eddy viscosity
$\mathcal{C}$	Cost function for optimising neural network
C_d	Drag coefficient
C_f	Skin friction coefficient
D	Diameter
f	Frequency
H	Height
J	Data loss term - error to measurements
N	Number of low-frequency harmonic modes
N_B	Number of boundary points
N_C	Number of collocation points
N_M	Number of high fidelity measurements
N_P	Number of points in domain

p'	Fluctuating pressure field from Reynolds decomposition, such that $p = P + p'$
p	Instantaneous pressure
$Re = \frac{u_\infty D}{\nu} = \frac{u_\infty H}{\nu}$	Reynolds number
S_p, S_{diff}, S_c, S_d	Production, diffusion, cross-diffusion and destruction terms from Spalart-Allmaras transport equation
$St = \frac{fD}{u_\infty}$	Strouhal number
t	Time
u_∞	Free-stream velocity
$\cdot^*$	Adjoint variable

Contents

Abstract	i
Publications	ii
Statement of originality and copyright declaration	iii
Acknowledgements	v
Nomenclature	vi
1 Introduction	1
1.1 Background	1
1.1.1 Problem	1
1.1.2 Goals	3
1.2 Outline	4
2 Literature Review	5
2.1 Current methods	5
2.1.1 Numerical methods	6

2.1.2	Experimental methods	10
2.1.3	Machine Learning and Neural networks	12
2.2	Data Assimilation and Flow Reconstruction	14
2.2.1	Variational Data Assimilation	14
2.2.2	Physics-Informed Neural Networks	19
2.3	Data-driven Turbulence Modelling	24
2.4	Objectives	27
3	Research Methodology	28
3.1	Governing Equations	29
3.1.1	RANS	30
3.1.2	Harmonic Balanced Navier-Stokes Equations	33
3.2	Data assimilation methods	41
3.2.1	PINN	41
3.2.2	Variational	44
3.2.3	Comparison: PINNs vs variational assimilation	45
3.3	Summary of Equations and Data	46
4	Flow Test Cases	47
4.1	2D laminar circular cylinder	47
4.1.1	Numerical Simulation and set-up	49
4.1.2	Dynamic Mode Decomposition	50

4.2	Periodic Hill	52
4.2.1	The periodic hill set-up	52
4.2.2	Comparison with a RANS-SA simulation	53
4.3	Axisymmetric bluff body wake	53
4.3.1	Experimental setup	55
4.3.2	Spectral Proper Orthogonal Decomposition	57
5	Correcting spatially misaligned experimental measurements with PINNs	59
5.1	Methodology	61
5.1.1	Defining the transformation matrix	61
5.1.2	Applying and learning the transformation matrix using PINNs	63
5.2	Validating transformation embedded PINNs	64
5.2.1	Flow reconstruction from measurements with individually applied transformations (rotation, scaling, translation)	65
5.2.2	Flow reconstruction from measurements with multiple combined transformations applied	68
5.2.3	Concluding Remarks	70
6	Turbulence model augmented PINNs versus variational data assimilation	72
6.1	Benchmark laminar cylinder flow comparison	74
6.1.1	Final Remarks on benchmark laminar comparison	77
6.2	Numerical and Model Setup for turbulent case	78
6.2.1	PINN set-up	78

6.2.2	Variational assimilation set-up	81
6.3	Model Optimisation	82
6.4	Results	83
6.4.1	PINN-DA-Baseline	84
6.4.2	Reconstructing pressure with PINN-DA-Baseline	85
6.4.3	PINN-DA-SA	87
6.5	Conclusion	99
7	HBNS constrained PINNs for mean-flow reconstruction and closure models	101
7.1	Numerical and PINN setup	104
7.1.1	Case and data sampling	104
7.1.2	PINN Setup and hyperparameter optimisation	105
7.2	Comparison of flow reconstruction with PINN-DA-HBNS	107
7.2.1	Problem formulation	108
7.2.2	Optimisation	109
7.2.3	Reconstruction error versus data resolution	110
7.2.4	Skin friction coefficient	112
7.2.5	Reconstruction of the mean flow using a RANS ($N = 0$), a quasi-linear ($N = 1$) and higher order ($N = 2$) approach.	115
7.3	Expanding flow reconstruction with PINN-DA to build HBNS closure models . .	122
7.3.1	Case setup and neural network configuration	123
7.3.2	Optimisation	125

7.3.3	Closure model using only velocity field input features	127
7.3.4	Closure models using resolved velocity gradient input features	129
7.4	Conclusion	133
7.4.1	Future Work	134
8	Reconstructing a 3D turbulent wake from 2D2C PIV data with PINNs	135
8.1	Numerical setup	137
8.1.1	Selection of harmonic modes for HBNS constrained data assimilation . .	137
8.1.2	Data sampling	140
8.1.3	Cubic spline interpolation	141
8.1.4	PINN Setup	142
8.2	Results	144
8.2.1	Use of transformation embedded PINN-DA on spatially misaligned ex- perimental measurement data	145
8.2.2	Flow Reconstruction using embedded transformation for spatially mis- aligned data	151
8.2.3	Pressure reconstruction of three-dimensional bluff body wake using PINN- DA	167
8.3	Conclusion	168
8.3.1	Future Work	170
9	Conclusion	171
9.1	Summary	171

9.2	Applications	173
9.3	Future Work	174
	Bibliography	176
A	Additional equations and derivations	190
A.1	Spalart-Allmaras equations	190
A.1.1	Formulation of Spalart-Allmaras equations	190
A.1.2	Wall-Distance multiplied Spalart-Allmaras Equations	191
A.2	Harmonic-Balanced Navier-Stokes equations	192
A.2.1	Derivation of HBNS equations	192
A.2.2	Division of low-frequency resolved and high-frequency unresolved scales for one resolved harmonic	194
B	Additional results for use of turbulence model augmented PINN-DA-SA	197
B.1	Hyper-parameter and loss weight selection	197
B.1.1	Tuning the model with a grid search approach	197
B.1.2	Alternative tuning methodologies	199
B.2	Modification to PINNs for Spalart-Allmaras augmentation	200
B.2.1	Singularity of the SA transport equation at the wall	200
B.2.2	Clipping negative values of $\tilde{\nu}$	202
B.3	Effect of regularisation on data assimilation techniques	202
B.3.1	Effect of solenoidal forcing regularisation on PINNs	203

B.3.2	Importance of regularisation for variational DA	204
C	Additional information for HBNS-constrained PINN-DA	207
C.1	Hyper-parameter selection	207
C.1.1	Activation Functions	207
C.1.2	PDE weighting	208
C.1.3	Architecture	208
C.2	PINN-DA-HBNS reconstructed fields	209
C.3	$N = 2$ HBNS closure model	209
D	Additional results for bluff body wake	213
D.1	Flow reconstruction without embedded transformation approach	213
D.2	Flow reconstruction of bluff body wake for $\Delta L = 0.5$	215
D.3	PINN-DA-HBNS-SA with fluctuating eddy viscosity field	219

List of Tables

3.1	Summary of different formulations of the governing equations for mean-flow reconstruction, the required training data and the chapters, in which they are applied	46
6.1	Loss function weights used in PINN optimisation. For PINN-DA-Baseline, the $\tilde{\nu}$ and λ^R weights can be ignored. Additionally, this regularisation weight is only valid for the $\Delta L = 0.5$ data resolution. For other resolutions, see Figure B.4. . .	82
6.2	Weights used in PINN optimisation when both mean velocity and Reynolds stress data are provided	85
6.3	Summary table describing the advantages and disadvantages of using PINN-DA and variational-DA for mean flow reconstruction.	98
7.1	Summary of different input features used to build data-driven HBNS closure model, $\mathbf{R}_{\mathbf{n},\mathbf{s}} = G(\bullet)$, mapping low-frequency resolved harmonics to low-frequency residual terms.	126
B.1	Sensitivity of final model loss function, C , to activation functions and learning rate on a model with 7 layers and 50 nodes/layer relative to final choice (tanh and 5×10^{-4} learning rate)	198
B.2	Sensitivity on final model loss function, C , to architecture parameters (number of layers and nodes) relative to final architecture (7 layers, 50 nodes).	198

B.3 Sensitivity on final model loss function, C , to perturbations on weights from

Table 6.1 198

List of Figures

2.1	Turbulent energy cascade showing how largest turbulent scales transfer energy to finer scales before dissipation as heat. Different numerical approaches resolve a different number of scales.	7
2.2	Comparison of the different turbulent scales being resolved and modelled for different simulation approaches with computational cost. Taken from (Deck et al. 2014).	8
2.3	Velocity distribution comparison of a jet at $Re = 3220$ using different simulation approaches (DNS, LES, RANS) from Rodriguez (2019).	9
2.4	Depiction of the basic Particle Image Velocimetry setup from McKeon et al. (2007).	11
2.5	The structure and training of an artificial neural network.	13
3.1	Structure of the Physics-Informed Neural Network as used in Chapter 6. A continuous mapping between spatial coordinates and flow variables is approximated by a deep neural network, denoted by N , defined by a set of weights and biases, θ . High-fidelity coarse data measurements and the (turbulence model augmented) RANS equations constrain the training of the network. k, m subscript the physics and data loss, respectively. i represents the i -th component of a vector.	42

4.1	The effect of number of resolved harmonics on the converged Strouhal number of the laminar circular cylinder for different Reynolds numbers. Resolution of six harmonic scales is sufficient to achieve convergence of Strouhal number. Taken from Rigas & Schmid (2022).	48
4.2	Numerical domain and boundary conditions for simulation of the laminar circular cylinder at $Re = 200$. The subdomain for flow reconstruction is also shown. Note that the cylinder body is circular and not an ellipsis as it appears, due to the axis scaling.	49
4.3	The mean mode and first three harmonic modes (extracted with DMD) of the streamwise component of $\hat{\mathbf{u}}$ for laminar circular cylinder flow at $Re = 200$. Real (top) and imaginary components (bottom).	50
4.4	Periodic hill domain and boundary conditions simulated at $Re = 5600$	53
4.5	High-fidelity (DNS) solution of the flow over a periodic hill compared with the RANS solution using an SA model. The dividing streamline between the recirculation region and rest of the flow is shown. (a) Streamwise mean velocity - DNS; (b) streamwise mean velocity - RANS; (c) absolute error of streamwise velocity relative to DNS solution ($U_{err} = U_{DNS} - U_{RANS}$). For comparison, the DNS dividing streamline is shown in grey.	54
4.6	Experimental setup for axisymmetric bluff body wake as performed by Zhu (2023), including the position of the 2D2C PIV plane, aft of the body, and the pressure taps on the rear surface.	55
4.7	PIV mean velocity fields for the axisymmetric bluff body wake at $Re = 1.88 \times 10^5$ for (a) streamwise and (b) vertical components.	56

5.1	Transformation embedded PINN architecture used to learn and correct spatial measurement misalignment. Importantly, one can see the introduction of the embedded transformation block of learnable parameters α , such that they correct the measurement locations.	63
5.2	Error (ε_T) of the estimated transformation matrix, $\mathbf{T}_{\text{est}}$ with respect to the magnitude of the true transformation. Tested on a) rotation only, b) scaling only, c) translation only.	65
5.3	Reconstruction of mean laminar circular cylinder flow from sparse misaligned measurements (left) without applied rotation, (right) with applied rotation. (Top) the misaligned measurement domain relative to the real location of measurements, (middle) reconstructed streamwise mean velocity, (bottom) reconstruction error.	67
6.1	High-fidelity DNS data of U (x-component) mean velocity component for laminar circular cylinder at $Re = 150$. (left) Full, true velocity field. (right) Sparse, high-fidelity field sampled at $\Delta L = 0.5$ data resolution used to reconstruct full flow field.	75
6.2	Comparison of the U (x-component) mean velocity component reconstructions for $d = 0.5$ data resolution with DNS (top). The variational-DA approach (left) and the PINN-DA-Baseline (right) figures are taken from (Foures et al. 2014) and (Sliwinski & Rigas 2022), respectively. (a,b) - U prediction. (c,d) - Absolute U error.	76
6.3	Turbulent periodic hill domain used for data assimilation, with $\Delta L = 0.5$ data resolution used to compare the variational and PINN approaches. (Left) The sparse velocity field at selected data resolution. (Right) The collocation points, where the PINN evaluates residual error (red) sampled across the domain using a Hammersley distribution and measurement locations (green crosses).	77

- 6.4 Structure of the Physics-Informed Neural Network. A continuous mapping between spatial coordinates and flow variables is approximated by a deep neural network, denoted by N , defined by a set of weights and biases, θ . High-fidelity coarse data measurements and the (turbulence model augmented) RANS equations constrain the training of the network. k, m subscript the physics and data loss, respectively. i represents the i -th component of a vector. 79
- 6.5 Residual loss during optimisation for (a) PINN-DA-Baseline, (b) PINN-DA-SA and (c) Variational-DA-SA. For PINN optimisation (training), there are two distinct phases: the ADAM (stochastic gradient descent based) optimisation for 150000 iterations followed by the L-BFGS-B phase. 82
- 6.6 Streamwise velocity profiles for DNS, RANS-SA and PINN-DA-Baseline. . . . 84
- 6.7 PINN-DA-Baseline mean-flow reconstruction.(a) Comparison between the DNS, PINN-DA-Baseline and PINN-DA-Baseline (with Reynolds stress data) velocity profiles for the U component. (b) True pressure field from DNS. (c) Pressure field inferred from the PINN-DA-Baseline approach from sparse mean velocity and Reynolds stress data. (d) Absolute mean-pressure error from PINN-DA-Baseline pressure reconstruction. 86
- 6.8 Reconstructed U velocity component for the PINN-DA-Baseline (top), PINN-DA-SA (middle), variational-DA-SA (bottom) approaches. (a,b,c) - U reconstruction. (d,e,f) - Absolute U error. PINN-DA-SA reconstructs the flow most accurately, significantly reducing the separation error seen in other approaches, whilst also reducing the near wall error seen in PINN-DA-Baseline. The dividing streamline (separating the recirculation region) is shown in (a,b,c) and compared with the DNS equivalent (in grey). Whilst the PINN-DA-SA recirculation region is indistinguishable, there is a small but noticeable difference for the PINN-DA-Baseline and variational-DA-SA solutions. 88

6.9	PDE residual error (x-momentum and y-momentum). (a) PINN-DA-Baseline and (b) PINN-DA-SA. The residual error between approaches are comparable despite the considerable improvements in mean velocity reconstruction error demonstrated by the PINN-DA-SA.	89
6.10	mean-flow reconstruction of streamwise velocity profiles for DNS, RANS-SA, PINN-DA-SA and variational-DA-SA.	92
6.11	A breakdown of the Reynolds forcing (excluding the potential field) f_1 (left) and f_2 (right) from the different methodologies: $f_i = \frac{\partial(2\nu_t S_{ij})}{\partial x_j} + f_{s,i}$. For PINN-DA-Baseline, $\nu_t = 0$. These are (a,b) - PINN-DA-Baseline, (c,d) - PINN-DA-SA, (e,f) - Variational-DA-SA.	93
6.12	$\nabla \times \mathbf{f}$. Comparison of curl($\mathbf{f}$) between the ground truth (DNS) and the different data assimilation techniques presented in this paper. (a) DNS, (b) PINN-DA-Baseline, (c) PINN-DA-SA, (d) Variational-DA-SA.	95
6.13	Volume-weighted L_2 mean velocity reconstruction error (as defined in Eq. (6.1)) versus data resolution for the three reconstruction methodologies (solid lines with symbols). All three approaches improve the RANS-SA predictions (dashed line).	96
7.1	Laminar circular cylinder at $Re = 200$ with data sampling showing (a) the coarse, measured velocity field at $\Delta L = 0.5$ (b) a sample collocation point distribution.	104
7.2	PINN-DA-HBNS architecture used in this chapter. The velocity fields, pressure fields and closure fields are separated into 3 separate sub-models. These become coupled through enforcement of the governing equations.	106
7.3	Convergence of loss during optimisation of PINN-DA-HBNS for differing number of resolved harmonics.	109
7.4	Velocity reconstruction accuracy versus data resolution for differing number of resolved harmonics, N . (a) $n = 0$ harmonic (mean) velocity reconstruction error (b) $n = 1$ harmonic velocity reconstruction error.	111

7.5	Skin-friction coefficient profile for DNS and PINN reconstructions for varying number of harmonics using data resolution $\Delta L = 0.5$	113
7.6	Flow reconstruction error for streamwise component of (a,b,d) mean flow ($\hat{\mathbf{u}}_0$) and (c,e) first velocity harmonic ($\hat{\mathbf{u}}_1$) for different number of resolved low-frequency harmonics, $N = 0, 1, 2$	114
7.7	Residual forcing on $n = 1$ x-momentum equations when resolving (a,b) $N = 1$ harmonics and (c,d) $N = 2$ harmonics	119
7.8	Curl of forcing, $\nabla \times \mathbf{R}_0$ for (a,c,e) the high-fidelity DNS data and (b,d,f) the PINN-DA-HBNS reconstruction. This compares (a,b) $N = 0$, (c,d) $N = 1$, (e,f) $N = 2$ resolved harmonics.	121
7.9	Schematic to show how sparse high-fidelity data can be leveraged using PINN-DA to produce a dataset, which can then subsequently be used to construct closure models.	123
7.10	Schematic of neural network closure model to find $\mathbf{R}_n = G(\hat{\mathbf{u}}_1)$. The neural network represents a pointwise mapping, in which the low-frequency resolved velocities (and gradients) at a single (spatial) point are used to determine the low-frequency residual closures at the equivalent location.	124
7.11	Convergence of training loss during optimisation of closure model for differing number of resolved harmonics (a) $N = 0$ (b) $N = 1$	126
7.12	True (PINN-DA-HBNS inferred) (a,b) and closure model predictions for $R_{0,u}$ field when trained with (c,d) velocity variables ($\mathbf{R}_n = G(\hat{\mathbf{u}}_1)$), (e,f) velocity gradients ($\mathbf{R}_n = G(\hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_1)$), (g,h) velocity laplacian ($\mathbf{R}_n = G(\hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_1, \nabla^2 \hat{\mathbf{u}}_1)$). This is compared for resolved harmonics (left) $N = 0$, (right) $N = 1$. Predictions are for $Re = 300$ after training on $Re = 150, 200$	128

7.13	Streamwise error distribution for predicted residual closure terms $\mathbf{R}_n$, for HBNS closure model resolved with different numbers of low-frequency harmonics. (a) $N = 0$ (b) $N = 1$. This is evaluated for the testing case, $Re = 300$	131
8.1	Relative magnitude of SPOD mode eigenvalues, performed on time resolved PIV data for the axisymmetric bluff body wake.	138
8.2	Streamwise SPOD velocity modes (real (top) and imaginary (bottom) components), extracted from experimental bluff body wake PIV data, representing two choices of first discrete resolved frequency, for flow reconstruction using PINN-DA-HBNS. (a,c) Low-frequency mode at $St = 0.04$. (b,d) Vortex-shedding frequency $St = 0.22$	139
8.3	Coarse pointwise measurements of the streamwise component of the mean velocity field used for flow reconstruction with PINNs, sampled at two different resolutions.	140
8.4	Cubic spline interpolation of streamwise PIV velocity data from two different data sampling resolutions ($\Delta L = 0.2, 0.5$). (Top) cubic spline interpolated reconstructions. (Bottom) reconstruction error.	141
8.5	Optimisation domain used to train the PINNs for flow reconstruction of the bluff body wake from sparse PIV velocity measurements. (a) Truncated domain for reconstruction. (b) Three-dimensional domain showing the data sampling (at $\Delta L = 0.2$) and a single sampled mini-batch of collocation and boundary points.	143
8.6	Reconstruction error averaged across all the PINN-DA formulations with and without the use of the embedded transformation approach for mitigating spatial measurement misalignment across different data resolutions. (a) L_2 -based error and (b) L_∞ -based error.	146

- 8.7 Reconstructed streamwise mean velocity error profiles for different PINN-DA formulations, evaluated at two different streamwise locations. (Top) without transformation embedded PINN-DA. (Bottom) with transformation embedded PINN-DA. 147
- 8.8 True PIV field in the vicinity of the bluff body trailing edge, where reconstruction error is highest. This shows the spatial misalignment between the PIV field and the trailing edge (and thus boundary conditions). 148
- 8.9 Nominal and transformed measurement coordinates showing the effect of using embedded transformation (restricted to the xy plane) during training of the PINNs when data spacing is $\Delta L = 0.2$ 150
- 8.10 3D velocity streamlines (coloured with streamwise mean velocity, as in Pavia et al. (2019)) for mean-flow reconstruction using PINN-DA-Baseline trained at data spacing $\Delta L = 0.2$, showing the annular recirculation of the full 3D flow. . . 152
- 8.11 PINN reconstruction using different methodologies for data spacing $\Delta L = 0.2$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-HBNS resolved at $St=0.22$ 153
- 8.12 Resolved forcing components ((left) streamwise, (right) vertical), $\mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1) = \overline{\hat{\mathbf{u}}_1} \cdot \nabla \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1}$, from PINN-DA-HBNS flow reconstructions, resolved at two different low-frequency harmonic scales. (a,b) - $St = 0.04$, (c,d) - $St = 0.22$. . . 155
- 8.13 Composite forcing terms for PINN-DA reconstructed flows using (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St = 0.04$. (e,f) PINN-DA-HBNS resolved at $St = 0.22$. The composite forcing is defined as $\mathbf{f} = \mathbf{R}_{0,s} + \mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1)$, where $\mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1) = \overline{\hat{\mathbf{u}}_1} \cdot \nabla \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1} = 0$ for PINN-DA-Baseline. 156

8.14 PINN reconstruction using different methodologies for data spacing $\Delta L = 0.2$, when embedded transformation method is being used. (a,b) PINN-DA-SA, (c,d) PINN-DA-HBNS-SA resolved at $St=0.04$, (e,f) PINN-DA-HBNS-SA resolved at $St=0.22$	159
8.15 Composite forcing terms for PINN-DA reconstructed flows using (a,b) PINN-DA-SA, (c,d) PINN-DA-HBNS-SA resolved at $St = 0.04$. (e,f) PINN-DA-HBNS-SA resolved at $St = 0.22$. The composite forcing is defined as $\mathbf{f} = \mathbf{R}_{0,s} + \nabla(\nu_t \nabla_s \hat{\mathbf{u}}_0) + \mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1)$, where $\mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1) = \overline{\hat{\mathbf{u}}_1} \cdot \nabla \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1} = 0$ for PINN-DA-SA.	161
8.16 Breakdown of relative contribution of corrective ($\mathbf{R}_{0,s}$), resolved ($\hat{\mathbf{u}}_1 \cdot \nabla \hat{\mathbf{u}}_1$) and modelled components $\nabla \cdot (\nu_t \nabla_s \hat{\mathbf{u}}_0)$ to total forcing for PINN-DA-SA and PINN-DA-HBNS-SA.	162
8.17 Reconstruction error for different PINN formulations using the embedded transformation approach showing (a) L_2 error and (b) L_∞ across data resolutions. . .	163
8.18 Mean streamwise velocity reconstruction error for PINN-DA-SA using data spacing $\Delta L = 0.5$, compared with RANS-SA simulation. Note how (a) the PIV data is directly subtracted from the RANS-SA prediction (i.e. both terms have underlying error), whilst (b) the PINN-DA-SA prediction is first transformed into measured coordinate space (corresponding to the measured PIV field).	166
8.19 Reconstructed pressure field using PINN-DA-Baseline with non-sparse velocity and Reynolds stress data: a) Pressure prediction at wall. b) Pressure across PIV plane. c) 3D iso-pressure surface at $C_p = -0.19$	168
B.1 Comparing the effect of convergence for PINN-DA-SA by applying (a) layerwise locally adaptive activation functions (L-LAAF) (b) an adaptive loss weighting algorithm.	199

B.2	Optimisation of PINN-DA-SA when using (middle) normal formulation of the SA transformation and (bottom) wall-scaled formulation (i.e. $\times d^2$). (a,b) convergence of x -momentum and SA transport loss using both formulations of the SA transport equation. (c,e) x -momentum residual field after 1000 iterations, using either the normal or wall-scaled SA respectively. (d,f) SA transport residual field after 1000 iterations, using either the normal or wall-scaled SA respectively.	201
B.3	Effect of regularisation parameter (solenoidal forcing magnitude penalisation), λ^R on PINN-SA solution for $\tilde{\nu}$. $\tilde{\nu}$ for increasing λ^R compared with the RANS-SA $\tilde{\nu}$ shows that increasing regularisation causes eddy viscosity to converge towards the RANS solution. The presented values of λ^R are 10^{-8} , 10^{-4} , 10^0 left to right.	203
B.4	The optimised regularisation parameter, $\lambda^{R,*}$, with respect to data resolution.	204
B.5	Absolute U error (a,c,e) and corrective solenoidal forcing, $\ f_{s,i}\ $, (b,d,f) for variational-DA-SA method using different regularisation approaches: No regularisation (top), L_2 regularisation (middle) and H_1 regularisation (bottom).	205
C.1	Effect of hyper-parameter selection on convergence of PINN-DA-HBNS. (a) Effect of activation function choice comparing tanh, swish and swish with layer-wise locally adaptive activation functions (L-LAAF). (b) Effect of spatial (SDF) weighting of residual PDE loss.	208
C.2	The effect of PINN-DA-HBNS architecture on loss convergence showing the impact of (a) split parallel branches (as used in Chapter 7) and (b) different number of nodes and layers.	209
C.3	Flow reconstruction for streamwise component of (a,b,d) mean flow ($\hat{\mathbf{u}}_0$) and (c,e) first velocity harmonic ($\hat{\mathbf{u}}_1$) for different number of resolved low-frequency harmonics, $N = 0, 1, 2$.	210
C.4	Convergence of training loss during optimisation of closure model for $N = 2$ resolved harmonics.	211

C.5	True (PINN-DA-HBNS inferred) (a) and closure model predictions for $R_{0,u}$ field when trained with (b) velocity variables ($\mathbf{R}_n = F(\hat{\mathbf{u}}_l)$), (c) velocity gradients ($\mathbf{R}_n = F(\hat{\mathbf{u}}_l, \nabla \hat{\mathbf{u}}_l)$), (c) velocity laplacian ($\mathbf{R}_n = F(\hat{\mathbf{u}}_l, \nabla \hat{\mathbf{u}}_l, \nabla^2 \hat{\mathbf{u}}_l)$). This is compared for $N = 2$ resolved harmonics. Predictions are for $Re = 300$ after training on $Re = 150, 200$	212
C.6	Streamwise error distribution for predicted residual closure terms $\mathbf{R}_n$, for HBNS closure model resolved with $N = 2$ low-frequency harmonics. This is evaluated for the testing case, $Re = 300$	212
D.1	PINN reconstruction using different methodologies for data spacing $\Delta L = 0.2$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-SA.	214
D.2	PINN reconstruction using different methodologies for data spacing $\Delta L = 0.5$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-SA.	215
D.3	Nominal and transformed measurement coordinates showing the effect of using embedded transformation (restricted to the xy plane) during training of the PINNs when data spacing is $\Delta L = 0.5$	216
D.4	Reconstructed streamwise mean velocity error profiles for different PINN-DA formulations, evaluated at two different streamwise locations (sampled at $\Delta L = 0.5$. (Top) without transformation embedded PINN-DA. (Bottom) with transformation embedded PINN-DA.	217
D.5	PINN reconstruction using different methodologies for data spacing $\Delta L = 0.5$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-SA.	218

D.6	PINN reconstruction of resolved forcing at $St = 0.22$ for data spacing $\Delta L = 0.5$.	
	(a,b) HBNS-SA with mean eddy viscosity only as in Chapter 8, (c,d) HBNS-SA with mean and fluctuating eddy viscosity.	219

Chapter 1

Introduction

1.1 Background

During the aerodynamic design process, new designs, such as different aerodynamic component shape topologies, are evaluated through analysis of their resultant flow fields. Various tools and approaches can be used to predict these flow fields including the use of numerical methods, such as computational fluid dynamics tools (CFD), or experimental methods, such as wind tunnel experiments. However, these approaches often have drawbacks. Given the central importance of the flow field to making key design choices, the primary motivation for this research is to explore and develop new tools that improve upon existing methodologies and can be used to enhance the aerodynamic design process. The explosion in growth of data-driven and specifically machine learning methods has opened up new possibilities for their use across many industries. This research aims to capitalise on this potential, by developing machine learning based tools for use in the aerodynamic design process.

1.1.1 Problem

To guide the aerodynamic design process, numerical simulations (CFD) are completed to compute the flow fields, by solving the governing equations of fluid motion. The Navier-Stokes

equations describe accurately the evolution of fluid flow in space and time. However, their direct numerical simulation (DNS) is computationally intractable in turbulent regimes, due to the resolution of many different length and time scales. This is problematic as most flows of engineering interest are turbulent. As a compromise, industrial simulation is dominated by Reynolds-Averaged Navier-Stokes simulations (RANS), which govern the time-averaged flow quantities instead of fully resolved time-varying values. However, the RANS equations are not in closed form and solving for these first-order statistics (mean-flow quantities) requires knowledge of the second-order velocity statistics, known as the Reynolds stresses. This closure problem is tackled by use of turbulence models, most commonly Boussinesq linear eddy-viscosity models (LEVM), including the Spalart-Allmaras (SA) model (Spalart & Allmaras 1992), $k - \omega$ (Wilcox 1988) and $k - \varepsilon$ (Launder & Sharma 1974) models. Whilst solutions of the RANS equations are computationally tractable, they are also less accurate than scale-resolving simulations (SRS). Although the turbulence models fully close the system of equations for the mean-flow quantities, they typically rely on empirical estimates and parameter tuning, lacking accuracy and generality, as quantified in a review by Xiao & Cinnella (2019). Alternatively, resolving the larger turbulent scales with Large Eddy Simulations (LES) simulations can be used. However, these are still costly and thus are not yet the industry standard.

Alternatively, engineers can use experimental methods to determine the flow field. A common experimental method is to carry out wind tunnel experiments. Thereafter, tools, such as Particle Image Velocimetry (PIV) and pressure tappings, for example, can be used to measure flow velocity or wall pressure respectively. Whilst experimental methods may provide valuable information about real-world flows and do not suffer from the issues of intractability, providing a high-fidelity representation of the flow, they have several other limitations. Firstly, experimental methods are often restricted in terms of spatio-temporal resolution as the flow can not be measured in the entire domain. PIV may be hardware restricted and thus velocity measurements are restricted to specific planes of interest (2D planes in a full 3D field). Likewise, whilst pressure taps can be used to measure wall pressure, it is difficult to obtain pressure measurements in the bulk flow without being intrusive. Secondly, there is often a greater financial cost when performing experiments, as expensive facilities and tools are required and test parts must

be manufactured. According to a report by NASA (2015), the annual operating cost of a wind tunnel is 1.35 million USD per annum. Finally, even as experimental methods capture the true physics of the system, they are also associated with experimental-specific errors. This includes imperfect test parts and conditions, as well as sensor noise and error impacting measurement accuracy.

1.1.2 Goals

The current limitations in both numerical and experimental methods create opportunities for the development of new tools to augment the aerodynamic design process. Providing a full and accurate representation of the flow field (typically unfeasible with either numerical or experimental methods) would have a high impact on the engineering design process. Thus, in this research, the principal goal is to develop such a tool that can, in a realistic engineering scenario, generate a full flow field, describing all flow variables (e.g. velocity and pressure), everywhere in the domain.

Growth in the development of data-driven methods and the availability of high-fidelity fluid data makes use of a data-based framework a good choice for such a tool. These data-driven methods have already been used to tackle wide range of fluid-based problems (Brunton et al. 2020, Duraisamy et al. 2019). One such example is the adoption of data assimilation methods to augment low-fidelity numerical simulations with partial high-fidelity experimental data. This can be used to reconstruct the full and accurate mean flow field, whilst still obeying the RANS equations, without the aforementioned limitations of existing numerical and experimental methods. This research aims to develop a tool that leverages high-fidelity measurement data to extract information not present in the original dataset: interpolate missing data by super-resolving the flow fields (such as coarse PIV fields), and infer missing fields (such as pressure and Reynolds stresses).

Amongst the rise of data-driven methods, machine learning tools, in particular, are an area of growing interest. Machine learning methods, like neural networks, have been used across a range of problems, such as computer vision, control and large language models. Furthermore,

in fields related to physics systems, the sub-field, Physics-Informed Machine Learning (PIML), has developed. In this approach, development of machine learning models is constrained with the knowledge that certain systems obey governing equations, in addition to high-fidelity data. Given that flows must obey the governing fluid equations (e.g. Navier-Stokes or RANS), this research will also implement a PIML methodology, when constructing a tool for aerodynamic design. In this work, the use of data assimilation with PIML and in particular Physics-Informed Neural Networks (PINNs), to reconstruct mean flow fields will be explored and evaluated.

1.2 Outline

The thesis is outlined as follows: Chapter 2 contains a literature review, explaining the limitations of current methods of obtaining flow fields and current developments in the use of data assimilation for flows. This is followed by a description of the specific objectives of this thesis, in order to meet the goals described in Section 1.1.2. In Chapter 3, the methodology is presented, including presentation of the key governing flow equations and the formulation of the data assimilation problem. Chapter 4 describes the flow test cases (laminar circular cylinder flow, turbulent periodic hill and experimental axisymmetric turbulent bluff body wake) used in the remaining sections of this work. In Chapter 5, an approach to correct biased experimental measurements (specifically where there is sensor miscalibration and misalignment) is proposed and validated on synthetic data. The key results of this research are then presented in Chapters 6-8. Chapter 6 demonstrates the use of turbulence-model augmented PINNs and directly compares them with variational data assimilation. Chapter 7 explores the augmentation of flow reconstruction with the resolution of low-frequency harmonic flow scales by performing data assimilation constrained by the Harmonic-Balanced Navier-Stokes. Chapter 8 applies the previously tested approaches to real experimental PIV data to validate use on a realistic engineering scenario (turbulent three-dimensional flow reconstructed from sparse experimental data). Finally, Chapter 9 contains concluding remarks and future areas of development.

Chapter 2

Literature Review

In this chapter, a description of state-of-the-art methods (numerical and experimental), along with their limitations, for generating flow fields will be discussed (Section 2.1) to highlight the need and subsequent development of data-driven tools. This is followed by an exploration into how data assimilation methods have been used to reconstruct flows (Section 2.2) or build models to improve existing numerical methods (Section 2.3). Finally, given the context of the literature review, the key objectives of this thesis are presented in Section 2.4, expanding on the goals set out in Section 1.1.2.

2.1 Current methods

As mentioned prior, generating flow fields accurately is a key aspect of the aerodynamic design process. Currently, this falls into two main categories: numerical simulations and experimental methods. Both approaches have advantages and drawbacks, which will be discussed in this section.

2.1.1 Numerical methods

The motion of fluids in both space and time can be described by the Navier-Stokes equations. In turbulent regimes, this flow is characterised by many length and time scales, as represented in the turbulent power spectrum seen in Figure 2.1. One can observe the interaction between different scales as the energy dissipates with decreasing scale size. At the largest scales, large eddies carry the most energy, which is transferred to smaller scales (see inertial subrange), before dissipating as heat in the smallest scales. These smallest scales are known as the Kolmogorov scales. The Kolmogorov length scale is defined as $\eta \equiv (\nu^3/\varepsilon)^{1/4}$, whilst the time scale is defined as $\tau_\eta \equiv (\nu/\varepsilon)^{1/2}$, as detailed in Pope (2000). Here, ε represents the dissipation rate. The ratio between the smallest Kolmogorov scales and the largest flow scales l_0, τ_0 (length and time respectively) can be related to the Reynolds number such that $\eta/l_0 \sim Re^{-3/4}$ and $\tau_\eta/\tau_0 \sim Re^{-1/2}$. Turbulent flows contain a mix of coherent flow structures, typically occurring at large scales, and incoherent, stochastic flow features. Whilst the incoherent, stochastic flow features are characterised by diffusive and dissipative effects, the larger coherent structures generally cause advective stirring. Therefore, to accurately simulate the flow and predict both coherent and incoherent flow features, one must resolve all of the length and time scales, by completing a direct numerical simulation (DNS).

In order to resolve all of these turbulent scales, DNS requires an extremely fine mesh and small time steps. This includes resolution of the smallest Kolmogorov length scale. Therefore, accurate simulation of most flows of interest (turbulent flows) with DNS is intractable, even without considering the resources available in industry and the desired turnaround time for aerodynamic design.

To make fluid simulations computationally tractable, at least some of the turbulent scales must be modelled. The most common approach in industry is to apply the Reynolds decomposition to the flow and time-average the Navier-Stokes equations, producing the RANS equations. Whilst their simulation is significantly cheaper computationally (mesh grid only needs to be small enough to resolve the mean flow gradients and not the small turbulent scales), there is a fundamental challenge with the RANS equations. The equations are unclosed and the first

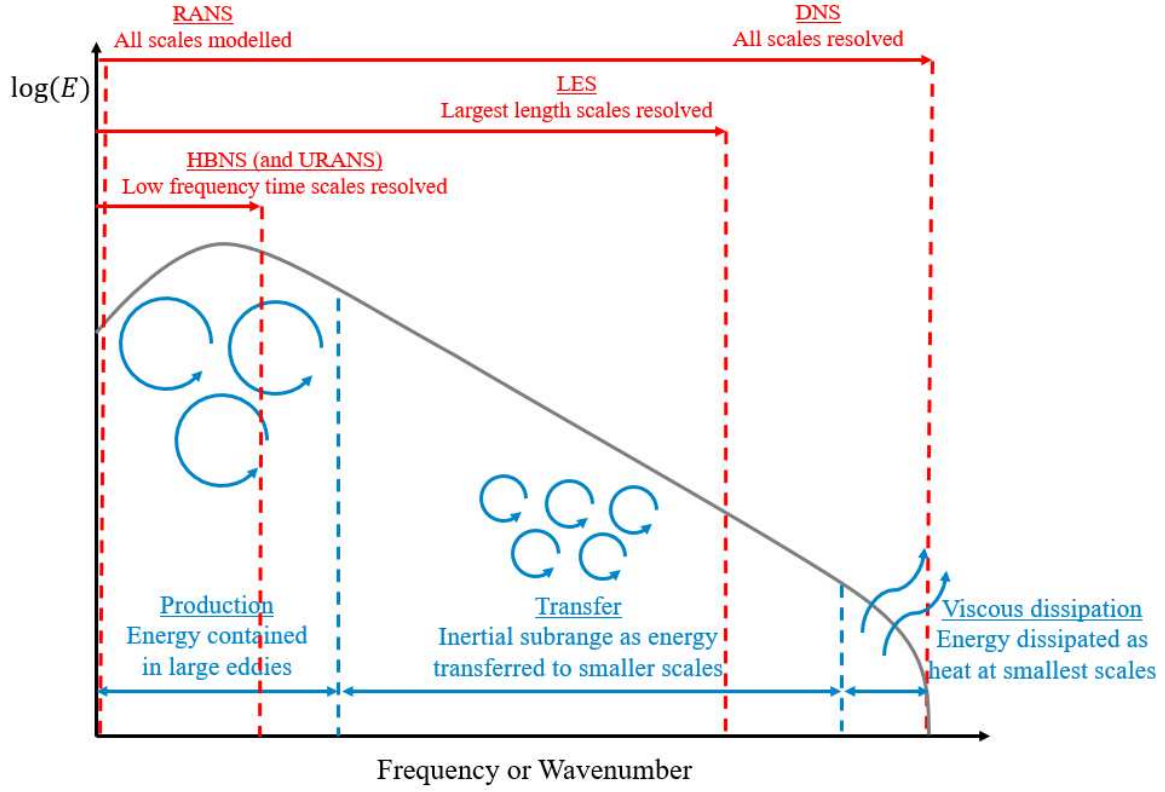


Figure 2.1: Turbulent energy cascade showing how largest turbulent scales transfer energy to finer scales before dissipation as heat. Different numerical approaches resolve a different number of scales.

order statistics (mean-flow quantities) depend upon the second order velocity statistics (the Reynolds stresses). Therefore, to solve this closure problem, turbulence models are used. They model the unknown Reynolds stresses and thus all of the turbulent scales (as seen in Figures 2.1-2.2). By far the most common class of turbulence models rely upon the Boussinesq linear eddy-viscosity model (LEVM). These include the Spalart-Allmaras (SA) model (Spalart & Allmaras 1992), $k - \omega$ (Wilcox 1988) and $k - \varepsilon$ (Launder & Sharma 1974) models. These models add additional equations, fully closing the system of equations for the mean-flow quantities. Use of these turbulence models makes simulation tractable and computation time is relatively small (compared with SRS approaches). However, these eddy-viscosity models have been constructed with empirical estimates and parameter tuning, and thus are often inaccurate or are not generalised for all flow regimes. The quantification of these errors is shown in a review by Xiao & Cinnella (2019). Non-linear eddy-viscosity models (NEVM) and Reynolds stress transport models (RSTM), which aim to formulate equations which directly model the individual Reynolds stress terms, have been developed, such as in Launder et al. (1975) and Wallin

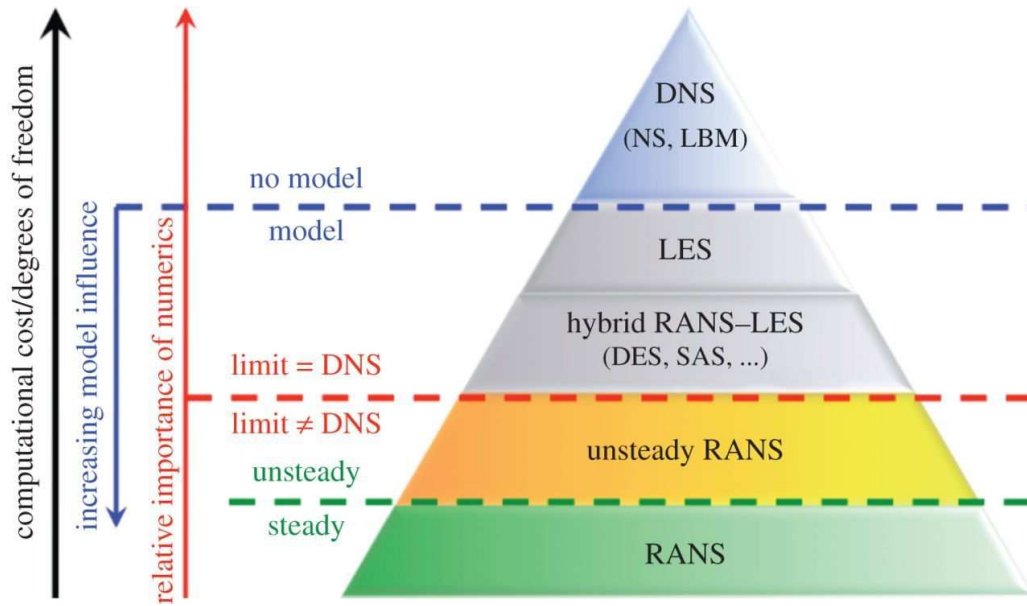


Figure 2.2: Comparison of the different turbulent scales being resolved and modelled for different simulation approaches with computational cost. Taken from (Deck et al. 2014).

& Johansson (2000). However, their improvement has only been demonstrated under certain conditions, such as those discussed in Gatski & Speziale (1993) (e.g. flows with sufficiently small strain rates), along with poorer convergence, and thus their use is limited in industry. By assuming that there are unsteady, slow moving fluid processes in the mean flow, one can reintroduce the time derivative to the RANS equations. These unsteady RANS (URANS) simulations average out high-frequency behaviour (across the time step), whilst modelling the transients. Despite enabling the calculation of some unsteady flows, URANS still relies heavily on modelling of the turbulent scales and there is debate about the validity and accuracy of this approach (Israel 2023).

Alternatively, instead of modelling all turbulent scales (as in RANS), one can choose to model only the smallest length scales, through LES simulations, as depicted in Figures 2.1-2.2. Using a low pass spatial filter, one can derive filtered Navier-Stokes equations and resolve the largest turbulence length scales. The effect of the filtered small length scales are then modelled using a sub-grid scale (SGS) model, such as the Smagorinsky–Lilly model (Smagorinsky 1963). As with DNS, the size of the smallest resolved scale depends on the mesh size, and thus often fine grids and small time-steps are still required to accurately capture small eddies, albeit larger than DNS (and the Kolmogorov length scale). Whilst cheaper than a DNS simulation, the cost

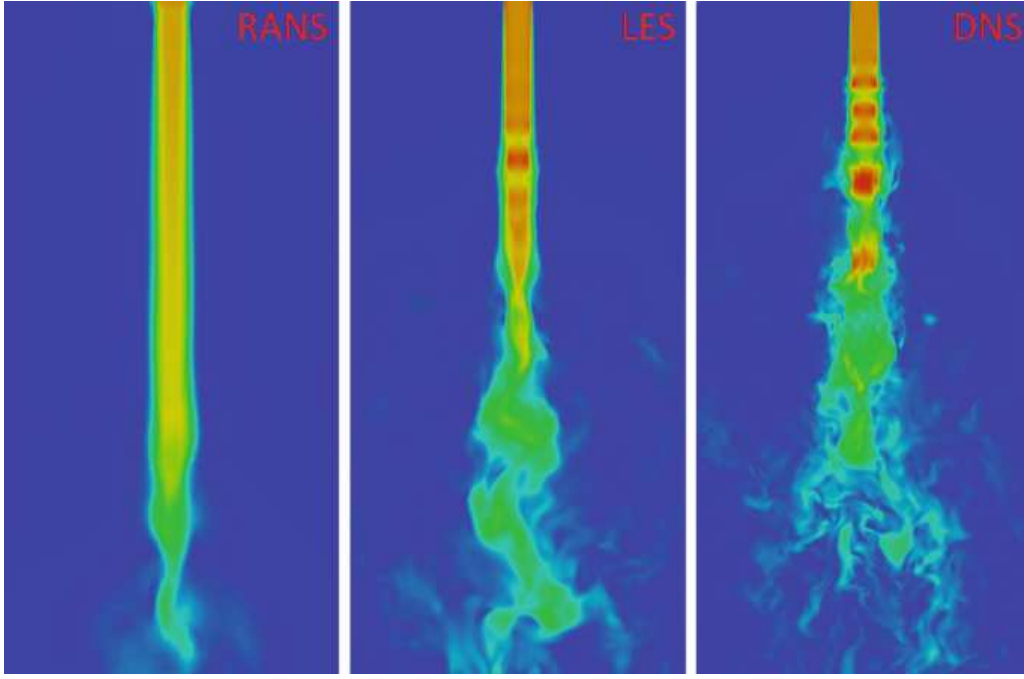


Figure 2.3: Velocity distribution comparison of a jet at $Re = 3220$ using different simulation approaches (DNS, LES, RANS) from Rodriguez (2019).

of LES is still higher than a RANS simulation. Therefore, despite an increase in accuracy, the computational cost of LES is still prohibitively high and thus is still not the most common simulation methodology for industrial application.

Instead of performing filtering in space, by transforming the governing fluid equations from the time domain to the frequency domain, one can choose to resolve the lowest frequency time scales as shown in Fabre et al. (2018) and Hernández et al. (2022). Using the Fourier-Galerkin weighted residual method and expanding the flow solution as a Fourier series (in time), one can derive the Harmonic Balanced Navier-Stokes (HBNS) and solve for some truncated number of low-frequency harmonic scales. In laminar-to-turbulence transitional regimes, solving the truncated frequency domain equations is an apt alternative to time domain methods. However, as flows become turbulent, the required number of harmonics to achieve sufficient convergence increases, as more energy transfers to the higher frequency. This approach can also become computationally intractable.

The choice of simulation methodology has large implications on the simulated flow solution, as demonstrated by Rodriguez (2019) and presented in Figure 2.3. Here RANS, LES and

DNS were all used to simulate a jet at $Re = 3220$, showing the effect of resolving and modelling different amounts of turbulent scales has on the flow solution. One can observe both similarities and differences between approaches. RANS, due to the large mesh grid and modelling of all scales, is not able to capture the dynamics of the jet flow as all the fluctuations in the flow are (time-)averaged out. Conversely, examination of the DNS jet flow, shows the fine chaotic flow fluctuations not present in the RANS simulation, as a result of the resolution of all turbulent flow scales down to the smallest Kolmogorov scales. The LES methodology is therefore an intermediate approach with a large fraction of the turbulent scales resolved, resulting in greater detail and accuracy than the RANS simulation, but less than the corresponding DNS. However, resolution of turbulent scales, requiring smaller mesh sizes, are associated with an increase in computational cost. Figure 2.3 highlights how performing numerical simulation of fluids is always a compromise between accuracy and tractability, which therefore impacts the engineering design process.

2.1.2 Experimental methods

Given the accuracy limitations of numerical methods, one can instead carry out laboratory experiments, most commonly using wind tunnels. However, whilst wind tunnel experiments should accurately represent the true physics of the flow (provided the experiment is well designed), wind tunnels are extremely expensive. According to a NASA report in 2015 (NASA 2015), the evaluated 23 wind tunnel facilities, ranging from subsonic to hypersonic tunnels, have an average current replacement value, at 2012 currency rates, of 126 million USD (or 80 million GBP) each. Furthermore, the annual operating costs (at 2012 rates) of these wind tunnels were on average 1.35 million USD (or 0.85 million GBP) each. Therefore building and operating or even renting one of these facilities can be incredibly expensive. Another consideration for wind tunnel (and other experimental methods) are the monetary and time costs when manufacturing test components. Given all of the above-mentioned costs, use of wind tunnels requires a large amount of investment and planning to maximise the information gathered during an experiment whilst reducing costs.

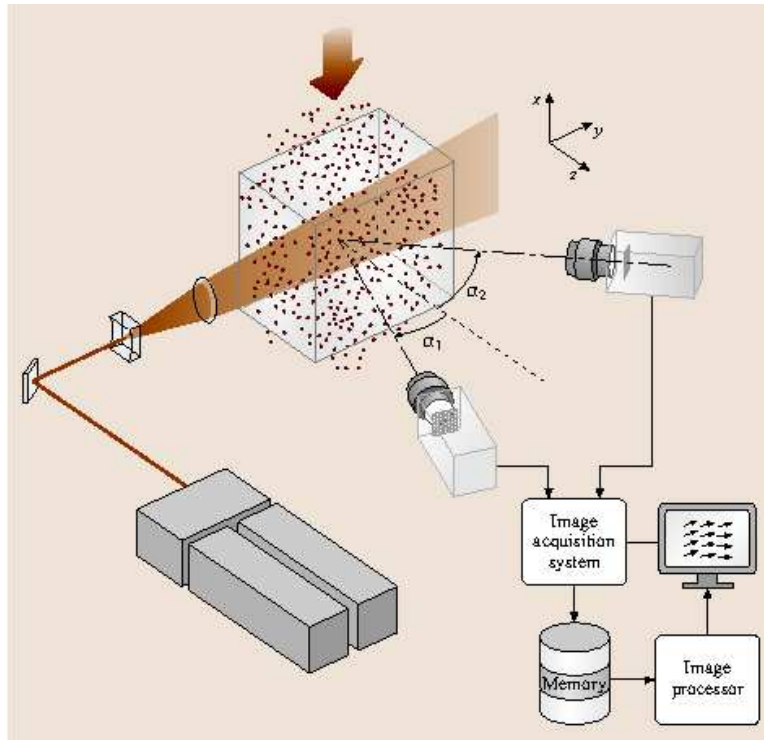


Figure 2.4: Depiction of the basic Particle Image Velocimetry setup from McKeon et al. (2007).

To experimentally measure flows, there exists many different measurement techniques, depending on the requirements. The aerodynamic forces exerted on a body immersed in a fluid can be attributed to the effects of the pressure field (pressure forces) and viscous effects (skin-friction forces). By placing sensors on the surface, one can determine the individual contribution to aerodynamic forces from these two components. Skin-friction forces are calculated using tools, such as hot wire anemometers and Preston tubes, to measure the velocity gradient at the wall. Pressure is typically measured with wall-mounted taps along the surface of the test body. However, whilst this enables one to measure surface pressure, measuring pressure in the bulk flow, requires more complicated tools such as a pressure rake. However, these devices also directly interfere with the flow and thus do not provide an accurate realisation of the flow.

Flow velocity can be measured with several approaches such as Laser Doppler Velocimetry (LDV) and Particle Image Velocimetry (PIV), as detailed in McKeon et al. (2007). PIV in particular is a common approach to measure velocity fields. By using a light sheet to illuminate a flow that has been seeded with tracer particles, one can use an imaging process algorithm to track their movement and thus measure the velocity of the particles and subsequently the flow.

Depending on the experimental setup, one can measure either two or three velocity components. The terminology XDYC (x-dimensions, y-components) is commonly used to describe the nature of the measured velocity field i.e. 2D2C are two velocity components measured on a 2D plane; 3D3C are all three velocity components measured in a 3D volume; 2D3C, where three components are measured on a 2D plane is called stereo PIV (due to use of two cameras). Therefore, depending on the facility and resources available, one may be hardware restricted and thus restricted to measuring certain fields (2D plane in a 3D flow field) or limited by measurement resolution. Additionally, due to the use of light, PIV and other velocity measurement methods are often unable to capture flow near the wall due to the interference from body reflections or even physical obstruction to the desired measurement regions.

Whilst these experimental methods enable an engineer to measure accurately a fluid flow, one must also consider error common with all experimental methods, such as noise, experimental uncertainty and sensor miscalibration, test part imperfection amongst other experimental limitations. Furthermore, the monetary costs of wind tunnel experiments can be high and thus may be used only sparingly. Therefore, use of experiments for engineering design can also be restricted.

2.1.3 Machine Learning and Neural networks

A key element of this thesis will be the use of data-driven methods and machine learning. The following section, therefore, briefly introduces relevant key concepts.

Machine learning describes the use of algorithms, designed to leverage data, in order to construct a model, which can be used across an array of problems, including representing (generalisable) mappings and decision making. Importantly, the aforementioned machine learning algorithms are used to “learn” a model from data. The most common approach to building machine learning models, is through the use of artificial neural networks, finding use across a range of problems with superior performance.

Neural networks are a model structure, presented in Figure 2.5, initially inspired by neurons and

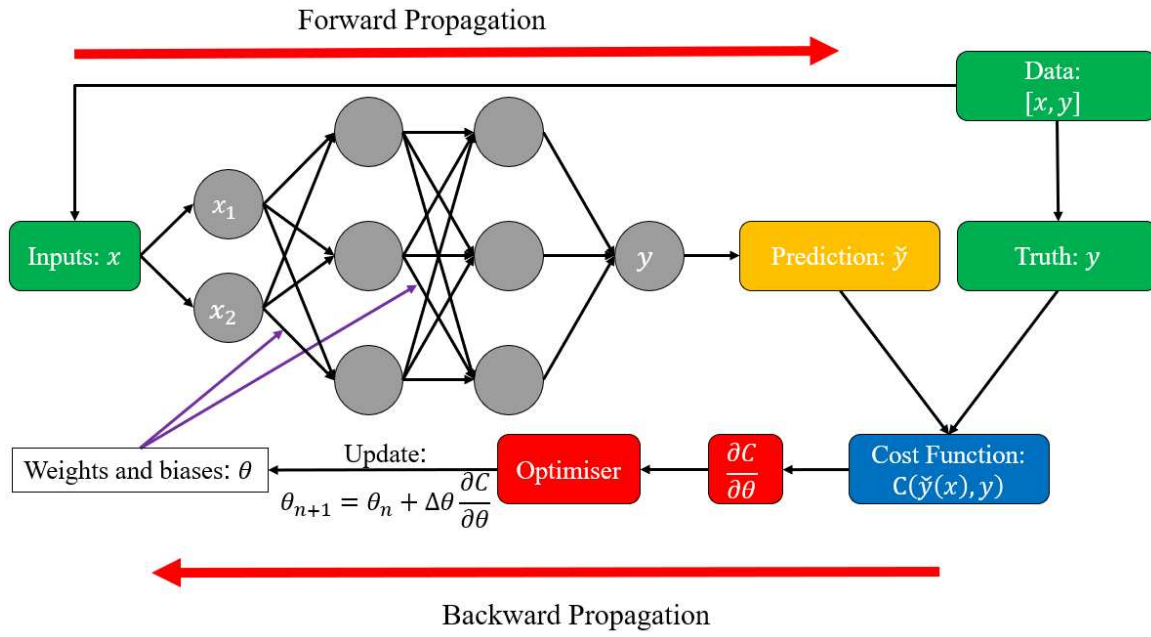


Figure 2.5: The structure and training of an artificial neural network.

synapses in biological brains. An artificial neural network (ANN) is constructed from nodes which are connected with weighted connections. According to the universal approximation theorem (Hornik et al. 1989), a neural network or multilayer perceptron (MLP) with at least one hidden, intermediate layer and non-linear activation functions can accurately depict any continuous, smooth function. In order to find the optimal weights and biases defining the network, θ , given some data mapping inputs $\mathbf{x}$ to outputs $\mathbf{y}$, a data-driven optimisation is set-up to minimise some cost or loss function such that

$$\arg \min_{\theta} C(\tilde{\mathbf{y}}(\mathbf{x}; \theta), \mathbf{y}). \quad (2.1)$$

Here, $\tilde{\mathbf{y}}$, represents the neural network output or prediction, given the input, $\mathbf{x}$, which corresponds to the true data value $\mathbf{y}$. The weights and biases are found using an optimisation algorithm, typically based on gradient descent, to iteratively update them.

2.2 Data Assimilation and Flow Reconstruction

In conjunction with the large growth in the use of data-based methodologies to tackle a wide range of fluid mechanics problems (Brunton et al. 2020, Duraisamy et al. 2019), data assimilation methods have enabled the augmentation of low-fidelity numerical simulations with experimental data, to reconstruct full and accurate mean-flow descriptions from the latter that obey the RANS equations. Such methods thus allow, among others, to extract information not present in the original dataset, interpolate missing data by super-resolving the flow fields (such as coarse PIV fields), and infer missing fields (such as pressure and Reynolds stresses). Different methods exist for tackling such a reconstruction problem: variational methods (Franceschini et al. 2020, Singh & Duraisamy 2016), Bayesian inference (Edeling et al. 2014, Magri & Doan 2020, Wu et al. 2016), real-state observer (Marquet et al. 2022, Saredi et al. 2021) and, amongst various machine-learning-based approaches, Physics-Informed Neural Networks (PINNs), which have recently been used in conjunction with RANS (Eivazi et al. 2022, Luo et al. 2020, Pioch et al. 2023, von Saldern et al. 2022). As further detailed in the following, both variational approaches and PINNs formulate data assimilation as an inverse problem, and more specifically as an optimisation one, where one aims to minimise the discrepancies between data and the reconstructed mean flow, in order to find a solution to the flow, which obeys both the governing flow equations and high-fidelity measurement data. Likewise, data assimilation can also be used to perform optimal control for fluids. In this case, the optimal control vector (and corresponding flow solution) is sought, such that some user-defined cost objective is minimised, which also obeys the governing equations. However, variational approaches and PINNs vastly differ in the solution method of such an optimisation problem, such that one might expect significant differences between their respective results.

2.2.1 Variational Data Assimilation

As above-mentioned, the variational approach formulates data assimilation as an optimisation problem, where the discrepancies between the reconstructed flow and data are minimised. This

minimisation is performed in conjunction with the strong imposition of the flow governing equations, (such as the RANS equations), through the adjoint method (other forms of variational data assimilation also exist, such as ensemble variational data assimilation). Both the RANS equations and their adjoint counterpart are solved based on classical Computational Fluid Dynamics (CFD) discretisation methods, such as the finite-element method that is employed here. Foures et al. (2014) notably applied the variational data assimilation method to super-resolve the mean flow for a laminar 2D circular cylinder. By providing sparse, synthetic experimental data (data that has been generated using high-fidelity computational methods but treated as experimentally generated) on a rectangular grid, the mean flow was successfully super-resolved. Notably, a Helmholtz decomposition was applied to the divergence of the Reynolds stress tensor, separating it into a potential component and a divergence-free component, the latter forming the control vector in the optimisation procedure. Ghosh et al. (2024) consider several loss functions for data enforcement (e.g. L_1 , L_2 , Huber, hybrid) when reconstructing a laminar flow from sparse velocity data that contains large outliers. It was shown that use of a Huber loss function was most robust to noisy data. This result also stresses that the effects of noise and experimental error must be considered when applying data assimilation methods to experimental data.

Mean-flow reconstruction of turbulent flows, which are often of greater interest, poses a greater challenge than the aforementioned laminar case, resulting in several developments for the variational data assimilation methodology. The variational data assimilation method by Foures et al. (2014) was expanded by Symon et al. (2017) by applying it to a turbulent mean flow around an aerofoil at $Re = 13500$ using experimental two-component planar PIV data. This required the introduction of a smoothing procedure to ensure smoothness of the control vector (or forcing vector), as well as care when defining the initial guess for the optimisation. The work by Franceschini et al. (2020) extends this work further. Instead of the Helmholtz decomposition, the Reynolds stress is broken down into an eddy-viscosity term and a non-eddy-viscosity forcing term. The eddy-viscosity term is approximated by including the SA model to the governing equations, whilst the variational data assimilation is used to find the non-eddy-viscosity forcing and also a corrective forcing within the SA equation itself. This is

intended to reduce the corrections required in regions where the LEVM is insufficient. A similar approach is considered by Brenner et al. (2024), applied to several 2D canonical flows. As with Franceschini et al. (2020), the divergence of the Reynolds stress tensor (or Reynolds forcing) is decomposed into a modelled eddy-viscosity component (solved with a $k - \epsilon$ model) and a non eddy-viscosity forcing term, which is itself decomposed using a Helmholtz decomposition, as in Foures et al. (2014), Symon et al. (2017). This formulation was shown to consistently outperform approaches where data assimilation was used to exclusively optimise the eddy viscosity. This is confirmed by Cato et al. (2023), showing that, whilst correcting and finding optimal eddy viscosity could correct mean-flow measurements, this was inferior to adding a corrective source term directly to the RANS momentum equations. The forcing corrections that are identified by these variational data assimilation methodologies may then be used to form a dataset to build machine-learning-based turbulence models, as performed in the field-inversion machine-learning approach (Holland et al. 2019, Singh 2018, Volpiani et al. 2021) (see Section 2.3). Instead of velocity measurements, Tayeh et al. (2023) explore the use of a similar framework as Franceschini et al. (2020) to reconstruct the turbulent mean flow of an aerofoil from limited wall observations (pressure or skin friction). It was shown that whilst use of skin friction measurements were no better than a RANS simulation, use of two wall pressure measurements was sufficient to vastly improve the turbulent mean-flow prediction (both mean flow and wall pressure profile reconstruction accuracy), on the condition that the pressure measurements are dissimilar to their corresponding RANS simulation values. However, whilst use of pressure measurements accelerated convergence relative to data assimilation with velocity data, it was not able to achieve the same mean velocity reconstruction accuracy.

Belligoli et al. (2021) explore a different approach to reconstructing mean flow with variational data assimilation. Instead two different types of perturbations were directly applied to the modelled Reynolds stress tensor when formulating the variational data assimilation problem. Although more expensive than the turbulence model production correction approach, used in Parish & Duraisamy (2016), the approaches by Belligoli et al. (2021) demonstrated higher accuracy.

Whilst the above-mentioned research has focused on mean two-dimensional flow, additional

work has been carried out to investigate use of variational data assimilation whilst considering unsteady flows and/or three dimensional flows. One example is by Chandramouli et al. (2020) with the 4DVar approach (three spatial dimensions and one temporal dimension). In their work, the three-dimensional, unsteady wake behind a cylinder flow at $Re = 3900$ is reconstructed from unsteady measurements at two perpendicular, synthetic 2D3C PIV planes. Mons et al. (2022) demonstrate an alternative variational data assimilation approach to infer instantaneous velocity and pressure fields from particle tracking velocimetry (PTV) data, on a numerical laminar cylinder flow case and an experimental turbulent jet case. Just as Brenner et al. (2024), Foures et al. (2014), Franceschini et al. (2020), Symon et al. (2017) assimilated the RANS equations treating the divergence of the Reynolds stress tensor from the RANS equations as a control vector to be optimised for, Mons et al. (2022) assimilated the Navier-Stokes equations treating the Eulerian acceleration term as a control vector. Using only instantaneous velocity measurements, this approach was shown to reconstruct the flow more accurately than the aforementioned 4DVar approach. Alternatively, Plogmann et al. (2024b) apply a similar formulation as Brenner et al. (2024) to perform variational data assimilation of the unsteady RANS equations (URANS) on two-dimensional, turbulent cylinder flow. Using sparse time-averaged velocity measurements, Plogmann et al. (2024b) was able to reconstruct the mean flow. Plogmann et al. (2024a) expanded this concept by transforming the URANS formulation into the Fourier domain in time (similar to the HBNS equations), removing the need of costly time stepping schemes (as in 4DVar). Whilst use of a URANS formulation does have drawbacks, such as the use of turbulence models designed for RANS flows and limited resolution of turbulent scales (Israel 2023), it was demonstrated that the dynamics of the unsteady flow solution (i.e. vortex-shedding frequency) was improved compared to a URANS simulation.

With data assimilation techniques, sensor placement becomes an important problem. Where data is available can have a big effect on the performance of the flow reconstruction. Symon et al. (2020) combines the variational data assimilation approach with a resolvent analysis to guide the placement of sensors, such that the maximum information can be extracted from the flow data, when reconstructing mean flow of a cylinder. By super-resolving the mean flow, the resolvent operator can subsequently be calculated more accurately than without data

assimilation, and using a single time-resolved velocity probe, the amplitude of the resolvent modes can be determined. Thereafter, the most critical regions of flow to place sensors can be determined. This approach was also applied to an aerofoil at $Re = 10000$ in Symon & McKeon (2018). In a similar framework, Mons & Marquet (2021) investigated the use of second-order adjoint techniques to determine optimal sensor arrangements.

Whilst this section primarily focuses on variational data assimilation for flow reconstruction, other classical numerical methods also exist. The other main category of data assimilation methods are sequential data assimilation techniques (Magri & Doan 2020). Whilst variational data assimilation takes into account future observations, sequential data assimilation methods only account for current and past observations. Sequential data assimilation is generally based on the principle of Bayesian minimum variance estimation, such as ensemble Kalman filter methods (Magri & Doan 2020, Nóvoa, Noiray, Dawson & Magri 2024, Nóvoa, Racca & Magri 2024). Magri & Doan (2020) applies sequential data assimilation methods to forecast a chaotic, turbulent reacting flow. Furthermore, the effect of biased observation data is considered in (Nóvoa, Noiray, Dawson & Magri 2024, Nóvoa, Racca & Magri 2024), using an ensemble Kalman filter-based approach to mitigate for the effects of bias.

There are also other alternative classical numerical methods. Agarwal et al. (2021) demonstrated the use of a constrained cost minimisation (CCM) technique to reconstruct velocity and pressure fields from both noisy and sparse data (synthetic particle tracks), by adding several physical constraints. Zauner et al. (2022) use a nudging-based data assimilation with the URANS equations to reconstruct the turbulent flow around a square cylinder, showing how this approach can mitigate some of the limitations in a standard URANS. Franceschini et al. (2021) applies a similar concept as in Symon et al. (2020), to demonstrate the use of resolvent analysis coupled with the mean-flow equations for mean and unsteady flow fields reconstruction, using several time-resolved measurements. Representing the Reynolds stresses with resolvent modes, both the mean-flow and unsteady flow components were captured. However, this approach was only demonstrated on low Reynolds number flows and a spatial dependence on measurement location was observed.

2.2.2 Physics-Informed Neural Networks

As an alternative to variational data assimilation, one can consider Physics-Informed Machine Learning (PIML) and specifically the PINN approach first presented by Lagaris et al. (1998) and expanded by Raissi et al. (2019). In conventional neural network approaches, models are optimised by minimising a cost function, which is purely driven by error to data. The PINN framework enforces weakly the governing equations, by adding the residual PDE errors to the cost function. This is similar to the aforementioned variational approaches, although PINNs only softly constrain the governing equations. Constraining the optimisation with fully or partially known governing equations (i.e. Navier Stokes or RANS), in addition to measurement data, ensures that solutions are physical, improving model performance, reduces the possible solution space during optimisation and the dependence on high volumes of data to train models. This approach has already shown great promise across a range of inverse flow problems (water wave problem in Jagtap, Mitsotakis & Karniadakis (2022) and high speed and supersonic flows in Mao et al. (2020) and Jagtap, Mao, Adams & Karniadakis (2022)).

PINNs have been used exhaustively to reconstruct the laminar cylinder flow. Sliwinski & Rigas (2022) approached the same mean flow problem as Foures et al. (2014), albeit using the PINN framework instead. Firstly, the Helmholtz decomposition was applied to the unknown divergence of the Reynolds stress tensor (forcing vector), as used in Foures et al. (2014), to reconstruct the flow from mean velocity data (first order statistics) alone. Secondly, by considering the full Reynolds stress tensor and providing both sparse first- and second-order velocity statistical data, the mean pressure field, which is otherwise difficult to measure experimentally, was also successfully extracted. Whilst the PINN approach demonstrated similar mean-flow reconstruction errors to the variational approach, no direct comparison has been completed. Xu et al. (2023) explore the sensitivity to data sparsity and location when reconstructing the unsteady laminar cylinder flow, demonstrating significant robustness across a range of conditions. Alternatively, Raynaud et al. (2022) represent the flow field as a truncated Fourier series (as in the HBNS equations) to reconstruct the unsteady laminar cylinder flow with ModalPINNs. Using unsteady velocity data to train the model, it was shown that ModalPINNs were able

to extract harmonic flow fields. Additionally, Raynaud et al. (2022) compared whether constraining the PINNs with the Navier-Stokes or the HBNS equations performed better. In their work, it was observed that using the time-resolved velocity data, enforcing the Navier-Stokes equations reconstructed the flow more accurately. One limitation of the trialled approach is the truncation of the Fourier series to some N temporal harmonics (here $N \leq 3$). Whilst sufficient for the laminar cylinder flow, neglecting all the unresolved, high-frequency harmonics introduces error into the enforced equations and thus for turbulent flows may neglect important flow physics.

In recent years, there has been a growing exploration of use of PINNs for turbulent mean-flow reconstruction. In the work by von Saldern et al. (2022), PINNs were combined with velocity data, to reconstruct the mean swirling turbulent jet flow and infer fields, such as density or eddy viscosity, that are not provided in the measured data. Eivazi et al. (2022) and Hasanuzzaman et al. (2023) used PINNs, constrained by the RANS equations, to reconstruct the flow field within a subdomain, by providing data at the domain boundaries for all flow fields (from experimental PIV data). As with the variational formulation, one can also use PINNs, constrained by the RANS equations with the addition of turbulence models, such as in Pioch et al. (2023). where several formulations of the Reynolds stress terms were trialled, including a normal $k - \omega$ model, a Prandtl mixing-length model and an optimal eddy-viscosity model. Pioch et al. (2023) was able to accurately reconstruct the mean flow over a backward-facing step from mean velocity measurements, with the optimal eddy viscosity (and pseudo Reynolds stress) approach performing best. However, as seen previously in Brenner et al. (2024), restricting the Reynolds stress terms to an isotropic LEVM formulation limits the potential of this approach due to neglect of anisotropy. Steinfurth & Weiss (2024) use PINNs to assimilate experimental data and reconstruct a mean three-dimensional turbulent flow. In their work, the PINNs are constrained by the RANS equations using an optimal eddy viscosity formulation, 2D3C mean velocity data from multiple PIV planes, in addition to wall pressure and wall shear stress data. Whilst it was shown that wall pressure data had little effect on the flow reconstruction, use of wall shear stress data, particularly in the absence of near-wall PIV velocity, improves the accuracy of the wall velocity gradient significantly and thus the boundary

layer profile. Although this work importantly extends the use of PINNs to three-dimensional turbulent mean flow, a large amount of mean velocity data was used (around 100 PIV planes), which may be unfeasible for industrial applications.

In addition to turbulent mean flow, research has been carried out, to investigate the use of PINNs for reconstructing unsteady flows. Using time-varying experimental 3D3C velocity data, Wang, Liu & Wang (2022) reconstructed instantaneous velocity and pressure fields. Furthermore, it was demonstrated that the PINN approach can be successfully used even with noisy and turbulent experimental data. This was confirmed in the work by Clark Di Leoni et al. (2023). In this work, PINNs were compared with the aforementioned numerical CCM approach (Agarwal et al. 2021), to reconstruct several flow cases at moderate Reynolds number, from particle track data. It was shown that even with noisy data, the PINN was able to reconstruct the instantaneous flow at least as accurately as CCM (and often more so). This was partly attributed to the often seen trend, that PINNs produce smoother solutions and thus better gradient predictions. Zhou & Grauer (2023) further explore the use of PINNs to improve processing experimental data, specifically Lagrangian particle tracking (LPT). LPT uses seed particles to follow and thus track three-dimensional flows. However, traditional algorithms assume that the seed particles exactly follow the flow trajectory, neglecting the effects of inertia. However, by embedding additional physical constraints on the PINN, it was able to account for inertial effects and predict the size of particles. This result is important as it demonstrates how, by including additional constraints, PINNs can be used to account for real-life effects from experimental methods (in this case particle inertia).

Several works, such as by Molnar et al. (2023) and Angriman et al. (2023), show implementation of PINNs within existing experimental frameworks or methodologies. Molnar et al. (2023) proposed a new PINN-based process for background-oriented Schlieren imaging (BOS), to reconstruct the density field from experimental data, with greater accuracy than traditional BOS image processing algorithms. Likewise, Angriman et al. (2023) combines data assimilation using PINNs to augment a nudging based numerical solver to generate turbulent states (such as a inflow boundary condition) without directly solving the Navier-Stokes equations, by assimilating various statistical flow data.

One of the greatest challenge when training PINNs is achieving sufficient convergence during optimisation. One avenue to increase the capability of PINNs is through the development of novel PINN architectures. The PINN framework has been expanded with Conservative PINNs (Jagtap, Kharazmi & Karniadakis 2020) and Extended PINNs (Jagtap & Karniadakis 2020), with further development by Shukla et al. (2021). These methods divide the domain into subdomains, each with its own PINN, increasing the representative capacity of the framework, as discussed in Hu et al. (2022). Additionally, Mishra & Molinaro (2022) and De Ryck et al. (2023) complete error analysis on the use of (extended) PINNs. Fourier feature networks, as proposed by Wang, Wang & Perdikaris (2021), embed spatio-temporal Fourier features in the input layer, such that multi-scale features can be better captured and the effects of spectral bias reduced. Bayesian PINNs (B-PINN), as proposed by Yang et al. (2021), may be particularly useful for reconstruction with noisy data. B-PINNs characterise data and model parameters as random variables with some probability function. Trained on noisy data, B-PINNs were a more accurate predictor than the normal PINN. Furthermore, B-PINNs also allow for quantification of uncertainty. This is demonstrated by Molnar & Grauer (2022), when used to reconstruct a flow from experimental measurements. The PINN framework can also be extended to convolutional neural networks. Gao et al. (2021) demonstrated the use of convolutional neural networks (CNNs) to super-resolve or denoise steady flow fields by enforcing weakly the steady Navier-Stokes equations during training. Kelshaw & Magri (2024) proposes physics-constrained CNNs (PC-CNN). PC-CNNs embed physical laws (unsteady Navier-Stokes) to the training of a CNN that maps between coarse measurements and high resolution fields. Applied to the Kolmogorov flow, the CNN model was able to super-resolve finer turbulent scales that would not be possible with classical interpolation methods, through successful augmentation of the governing laws. Furthermore, using PC-CNNs, Kelshaw & Magri (2024) was able to infer the true state using observation data contaminated by spatially varying systematic error (bias or epistemic uncertainty). Mo & Magri (2024) applies PC-CNNs to reconstruct the full flow field of a laminar, unsteady bluff body wake and a turbulent Kolmogorov flow from sparse, incomplete observation data. Additionally, the effect and choice of loss function was investigated, especially when considering robustness to noisy observation data.

In addition to model architecture, weighting of the composite loss terms for PINN optimisation can have a big effect on convergence. Tuning loss weights is a complicated task and can have a big effect on convergence of PINNs, and has thusly led to the development of loss weighting algorithms. A popular approach is through the use of weighting algorithms, as proposed by Wang, Teng & Perdikaris (2021). Here, a weighting algorithm, based on the relative loss (and loss gradients) was shown to provide better PINN convergence. Similarly, Sharma et al. (2023) suggested the use of spatially weighting PDE loss terms, particularly for stiff PDEs. Another approach to improve PINN training is the use of adaptive activation functions (Jagtap, Kawaguchi & Karniadakis 2020), which demonstrate acceleration of model convergence. A more unique approach to this problem was suggested by Zeng et al. (2022) with Competitive PINNs. In this approach a second, discriminator neural network is used, which produces a spatial weighting of the PINN loss terms. Thereafter, a min-max optimisation is constructed, in which the PINN aims to minimise the data and PDE loss and the discriminator network aims to maximise the loss. In this work, it is explained how this approach is better conditioned than the traditional loss approaches and thus has promise for future application.

Given the relative infancy of PINNs, an alternative approach to improving PINNs is by applying advancements in other data-assimilation methods to the existing PINN framework, in order to improve on previous studies. It is possible that similar strategies as proposed in the variational approaches (as discussed in Section 2.2.1) may also be developed for the PINN framework, to improve reconstruction of flows. Du et al. (2023) compare both these approaches (PINN and adjoint-variational) for a turbulent channel flow using sparse spatio-temporal velocity measurements, concluding that the variational approach was a better forecaster. However, these comparisons, such as Clark Di Leoni et al. (2023), Du et al. (2023), are rare and therefore there is limited understanding in how these approaches differ.

2.3 Data-driven Turbulence Modelling

Whilst solutions of the RANS equations are computationally tractable, they are also less accurate due to their under-determined nature and the lack of exact turbulence closure models (Xiao & Cinnella 2019). In addition to flow reconstruction, the rise of data-assimilation techniques has led to a growing number of alternate data-driven approaches to tackle the closure problem. Whilst early approaches often focused on improving existing turbulence models by tuning the constants within them, such as Yarlanki et al. (2012), a wide range of approaches to turbulence modelling have developed (Vinuesa 2021).

Despite their growing number, current data-driven turbulence models have still not yet reached mainstream application. Wu et al. (2019) showed that improved predictions of Reynolds stress tensors using data-driven turbulence models do not always improve mean velocity solutions of the RANS equations. Brener et al. (2021) show how different reformulations of the Reynolds stresses (such as the decomposition of the Reynolds stresses into eddy-viscosity and non eddy-viscosity based terms) result in different mean velocity errors. Furthermore, black-box machine learning turbulence models often suffer from a lack of generalisability, as tackled by Edeling et al. (2014) (Bayesian Model-Scenario Averaging) and Beetham & Capecelatro (2020) (Galilean invariance embedding), and interpretability, explored by He et al. (2022) (Shapley Additive Explanation analysis) and Weatheritt & Sandberg (2016) (Gene Expression Programming). Here interpretability is defined as the capacity to understand and explain how and why a data-driven turbulence model makes a certain prediction about a flow, under a given condition. One example, such as in Weatheritt & Sandberg (2016), is the formulation of equations, defining the turbulence model.

One approach to developing practical, generalisable and interpretable models is through the inclusion of physical constraints on turbulence model development (similar to PINNs). Beetham & Capecelatro (2020) constructed a sparse regression model that embeds Gallilean invariance, to produce a generalisable model that requires less data than historic neural network approaches, performing better than historical LEVM across different geometries and higher Reynolds number than the training data. Many data-driven turbulence models map flow features, such as in

Wang et al. (2017), showing success across a wider range of cases. Likewise, Cherroud et al. (2022) improved the LEVM by introducing an additional anisotropic term, mapped as a function of several flow tensors and invariants. After tuning this term with data, the resultant model demonstrated improved accuracy for velocity fields and skin friction coefficient. Similarly, the work by Volpiani et al. (2021) uses a neural network to model a corrective term to the RANS momentum equations (instead of corrections to the turbulence model equations) as a function of RANS outputs. This acts in combination with a SA turbulence model.

Selecting and weighting turbulence model or coefficients based on flow regime has also seen some success, such as in Edeling et al. (2014) and de Zordo-Banliat et al. (2024). In de Zordo-Banliat et al. (2024), a spatial weighting of several different LEVMs, dependent on the local flow, is used to form an aggregate turbulence model, with this resultant model outperforming each of the individual composite models

Holland et al. (2019) also apply machine learning techniques to find an additional corrective production term to the SA equation, by applying the field inversion process first proposed by Singh (2018). This corrective term is deduced from high-fidelity data using the field inversion process. This corrective term is then mapped from several flow features. It was found that the augmented Spalart-Allmaras model using this technique achieved a substantially better level of predictive accuracy than a traditional SA model, with the velocity field, pressure distribution and force plots matching experimental data better, even in the post-stall regime. Similarly, Grabe et al. (2023) apply the field inversion process with a specific focus on selection of flow features. It follows therefore, that a data-assimilation process to reconstruct the flow, as in Section 2.2, can be used to increase the volume of high-fidelity data, especially if one wants to reduce the volume of data required to build a data-driven turbulence model.

One limitation of most RANS turbulence model approaches (including traditional LEVM turbulence models) is the modelling of non-local dynamics with a local model. The Navier-Stokes are Markovian as the future state of the flow can be determined from the current state. However, by an averaging or filtering operation, the governing flow equations (e.g. RANS, LES or other SRS approaches) become non-local, as the evolution of the statistical or filtered flow state

depends on the history of previous flow states (as discussed in Girimaji (2023), Lumley (1979), Mishra et al. (2015)). The filtered flow physics appear as non-Markovian effects in the unclosed terms, and thus are significantly influenced by non-local spatio-temporal effects. In particular, coherent flow structures, driven by instabilities in the larger turbulent scales, have strong non-local spatio-temporal effects on the unclosed terms. Both LEVM and most data-driven turbulence models assume that the closure at a single point can be determined from a local tensor at the same point (locality assumption). However, modelling the non-local evolution of turbulence with a local model is fundamentally limited, leading to poor representation of the true turbulent dynamics (Agrawal & Koutsourelakis 2024, Mishra et al. 2015). To overcome this limitation, approaches, such as Han et al. (2023), use non-local input features (i.e. taking inputs from a point cloud) to build a RANS closure model, enabling better representation of the non-local effects in the Reynolds stress tensor.

Whilst RANS represents the extremity, in which all flow fluctuations and turbulent scales are averaged, resulting in the strongest influence from the non-local effects, resolution of some turbulent scales, using approaches such as LES and other SRS formulations, subsequently reduces the influence of non-local spatio-temporal effects. In particular, the resolution of coherent structures can significantly reduce the non-local effects. As a result, development of closure models for SRS approaches is less affected by the locality assumption (as used in many RANS turbulence model formulations). Rigas & Schmid (2022) develop a data-driven framework, using the laminar circular cylinder flow, to build a closure model for the frequency-truncated HBNS equations. Using a neural network, the low-frequency, coarse grained residual terms, were expressed as a function of the resolved low-frequency velocity harmonics (i.e. Fourier modes in time) for unseen Reynolds numbers. Application of this approach enables accurate solving of the low-frequency coarse grained dynamics at a reduced computational cost. However, whilst accurate modelling of the residual closure terms was observed, use of the trained closure model with a numerical harmonic solver has yet to be demonstrated. Similar to Han et al. (2023), Parish & Duraisamy (2017) discuss the use of non-local closure models albeit for LES (instead of RANS).

2.4 Objectives

In Section 1.1.2, the primary goal of this research were provided: design a data-driven tool for aerodynamic design that can generate a full and accurate flow representation (using PINNs). In particular, the aim of this work is to explore and evaluate the use of data assimilation methods, in particular PINNs, to reconstruct mean flow fields. Following the analysis of the literature, meeting the aforementioned goals is primarily achieved through the exploration of the following objectives:

- To directly compare PINNs with existing data assimilation methods, specifically variational data assimilation, on the same test cases and to evaluate whether techniques and approaches from these existing methods can be transferred successfully to PINNs.
- To investigate how augmenting and constraining the physics formulation with prior information and equations (e.g. inclusion of turbulence models, resolution of some scales, provision of different data) affects the capabilities of PINNs.
- To extend the application of PINNs (especially already trained PINNs) beyond their primary function (super-resolution and field inference) by leveraging their flow field representation to develop new models (such as closure models).
- To demonstrate use of PINNs for more realistic engineering use cases (turbulent, experimental, three-dimensional flow field), instead of “synthetic” data, and investigate, and tackle the challenges associated with laboratory measurements.

Chapter 3

Research Methodology

In this thesis, the inverse problem is solved, where a mean (time-averaged) flow field is reconstructed from a set of partial observations. The inverse problem is defined, such that a set of flow observations (data) is used to discover the factors which caused them. These factors may correspond to unknown information about the flow, such as closure terms or parameter correction of empirical closure models. This contrasts the forward problem, where factors (boundary conditions) are defined, and the resultant effects of the system are calculated. The flow observations are defined as N_M sparse high-fidelity mean velocity measurements at points $\mathbf{x}_m = (x_m, y_m, z_m)$, where the subscript m denotes the m -th sparse measurement. The mean velocity measurements are denoted as $\mathbf{U}_m = U_{m,i} = (U_m, V_m, W_m)$ and U_i is the i -th component of the mean velocity. The reconstructed flow field is sought by solving the constrained optimisation problem, such that error to measurements, J , is minimised, where

$$J = \sum_{m=1}^{N_M} \|\mathbf{U}_m - \check{\mathbf{U}}(\mathbf{x}_m)\|^2, \quad (3.1)$$

whilst simultaneously the reconstructed flow field satisfies the under-determined governing flow equations. Here, $\check{\mathbf{U}}(\mathbf{x}_m)$ is the reconstructed mean velocity at $\mathbf{x}_m$. Whilst the equations in this section are presented for general flows, they can also be applied to two dimensional flows (as in Chapters 6 and 7).

Firstly, the governing flow equations are presented in Section 3.1. Section 3.1.1 describes the governing mean-flow equations (RANS) both without (baseline) and with (SA) turbulence model augmentation. This is followed by Section 3.1.2, which describes the governing harmonic balanced equations (HBNS) and compares it with the mean-flow equations.

Secondly, in Section 3.2, the details of the two data assimilation approaches used in this work, to solve the aforementioned optimisation, are described - a PINN approach (Section 3.2.1) and an adjoint-based variational approach (Section 3.2.2).

3.1 Governing Equations

First, consider the incompressible Navier-Stokes equations, which describe a fluid flow

$$\nabla \cdot \mathbf{u} = 0, \quad (3.2a)$$

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p + \nu \Delta \mathbf{u} = \mathbf{0}, \quad (3.2b)$$

where $\mathbf{u} = (u, v, w)$ is the instantaneous velocity vector, p is the instantaneous pressure and ν is the kinematic viscosity of the fluid. Letting $\mathbf{w} = [\mathbf{u}, p] = [u, v, w, p]$, the Navier-Stokes equations can be written in compact form,

$$\mathbf{M} \partial_t \mathbf{w} + \mathbf{L} \mathbf{w} + \frac{1}{2} \mathbf{N}(\mathbf{w}, \mathbf{w}) = \mathbf{0}, \quad (3.3)$$

where $\mathbf{M}$ is a mass matrix, $\mathbf{L}$ denotes the Stokes operator and $\mathbf{N}$ represents the symmetrised non-linear convection operator. These are defined as:

$$\mathbf{M} = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & 0 \end{pmatrix}; \quad \mathbf{L} = \begin{pmatrix} \nu \Delta & \nabla() \\ \nabla \cdot () & 0 \end{pmatrix}; \quad \mathbf{N}(\mathbf{w}_1, \mathbf{w}_2) = \begin{pmatrix} \mathbf{u}_1 \cdot \nabla \mathbf{u}_2 + \mathbf{u}_2 \cdot \nabla \mathbf{u}_1 \\ 0 \end{pmatrix}. \quad (3.4)$$

3.1.1 RANS

By applying the Reynolds decomposition, the unsteady flow field can be separated into a mean $(\mathbf{U}, P)$ and fluctuating component $(\mathbf{u}' = (u', v', w'), p')$, i.e. $\mathbf{u} = \mathbf{U} + \mathbf{u}'$; $p = P + p'$. Through time-averaging of the unsteady Navier-Stokes equations (3.2), one can obtain the steady RANS equations. The data-driven techniques used in this study will be applied to mean-flow data, which satisfy the RANS equations

$$\nabla \cdot \mathbf{U} = 0, \quad (3.5a)$$

$$\mathbf{U} \cdot \nabla \mathbf{U} + \nabla P - \nabla \cdot (\nu \nabla_{\mathbf{S}} \mathbf{U}) + \nabla \cdot \tau = \mathbf{0}, \quad (3.5b)$$

where τ is the Reynolds stress tensor and $\tau_{ij} = \overline{u_i' u_j'}$. $\nabla_{\mathbf{S}} \mathbf{U}$ is the mean strain-rate tensor and the operator $\nabla_{\mathbf{S}} = (\nabla + \nabla^T)$. For a three-dimensional flow (3D), this results in a system of 4 equations with 10 unknowns (assuming no spanwise mean flow, this can be reduced down to 3 equations and 6 unknowns). Given the dependence of the mean-flow solution, $(\mathbf{U}, P)$, on the terms of the unknown Reynolds stress tensor, τ , determination of these terms is critical to the performance and capability of data-assimilation methods for reconstructing the mean-flow field. The remainder of Section 3.1.1 will present formulations of the mean-flow equations, such that data can be used to infer the Reynolds stress (closure) terms or their divergence.

Baseline formulation of mean-flow equations

To solve the inverse problem, an additional constraint is placed on the closure term to improve its inference, such that it can be inferred using only mean velocity data. Firstly, the divergence of the Reynolds stress tensor $\nabla \cdot \tau$ (as it appears in the RANS equations (3.5b)) is considered as a forcing vector, $\mathbf{f}$ (Reynolds forcing vector), where f_i can be used to denote the i -th component of the vector. This reduces the closure term from 6 individual Reynolds stresses to 3 individual forcing terms (or 2 forcing terms from 3 stresses with no spanwise mean flow). Subsequently, a Helmholtz decomposition is applied, as in Brenner et al. (2024), Foures et al. (2014), Plogmann

et al. (2024b), Sliwinski & Rigas (2022),

$$\mathbf{f} \equiv -\nabla \cdot \boldsymbol{\tau} = \nabla \phi + \mathbf{f}_s, \quad (3.6)$$

decomposing the forcing into a scalar, potential part (ϕ) and a divergence-free, solenoidal, vectorial component ($\mathbf{f}_s$). This latter condition (divergence-free) provides an additional equation. Introducing decomposition (3.6) into the RANS equations (3.5), gives

$$\nabla \cdot \mathbf{U} = 0, \quad (3.7a)$$

$$\mathbf{U} \cdot \nabla \mathbf{U} + \nabla (P - \phi) - \nabla \cdot (\nu \nabla_s \mathbf{U}) - \mathbf{f}_s = \mathbf{0}, \quad (3.7b)$$

$$\nabla \cdot \mathbf{f}_s = 0. \quad (3.7c)$$

In addition to the new equation, the Reynolds forcing vector now consists of 4 terms (3 in two dimensions), in the form of a solenoidal forcing vector and a scalar potential field. In this system, there are now 5 equations and 8 unknowns (4 equations and 6 unknowns in 2D). By combining the pressure and potential terms into a single scalar field ($P - \phi$), another unknown can be removed. The resultant flow field, $(\mathbf{U}, P - \phi, \mathbf{f}_s)$ is composed of the mean velocity components, combined pressure and potential forcing term and the solenoidal forcing components. Providing only mean velocity data, at discrete locations, is sufficient to close the system at those points. However, the $P - \phi$ term can not be separated into its constituent parts and the pressure field cannot be determined.

Turbulence Model - SA

Whilst the above approach combines data with a formulation of the governing equations to infer the unclosed divergence of the Reynolds stress tensor, the closure problem still poses a challenge away from data points, since an infinite number of potentially physically unrealisable solutions satisfy the under-determined governing equations. The following methodology is proposed to

augment the aforementioned approach, by enforcing additional physical constraints.

When performing numerical simulations of the RANS equations, turbulence modelling is utilised to approximate the Reynolds stresses and close the equations. Typically, the Boussinesq approximation is used with partial success across a range of flows, where the Reynolds stresses are assumed to be isotropic and depend on the eddy viscosity ν_t . However, neglecting the anisotropic component, as well as other modelling assumptions made in the calculation of ν_t , introduce errors which propagate to the mean-flow solution. Instead, as applied in Franceschini et al. (2020), one can decompose the Reynolds forcing into a modelled, isotropic part and a corrective component which can include the anisotropic part, in addition to any isotropic component not captured by the modelled part, such that

$$\mathbf{f} = \nabla \cdot (\nu_t \nabla_{\mathbf{S}} \mathbf{U}) + \mathbf{f}_{\mathbf{c}} = \nabla \cdot (\nu_t \nabla_{\mathbf{S}} \mathbf{U}) + \underbrace{\nabla \phi + \mathbf{f}_{\mathbf{s}}}_{\mathbf{f}_{\mathbf{c}}}. \quad (3.8)$$

The first term represents the modelled forcing component and the second term the corrective forcing component, such that the mean-flow solution matches the data (high-fidelity measurements). Additionally, the Helmholtz decomposition is applied to $\mathbf{f}_{\mathbf{c}}$. By augmenting the system with a turbulence model, one can solve for ν_t . Here, the one-equation Spalart-Allmaras turbulence model, which provides a governing equation for a pseudo eddy-viscosity variable $\tilde{\nu}$, will be used. The actual eddy viscosity ν_t may then be obtained from $\tilde{\nu}$ through an algebraic expression. The full SA equations are given in Appendix A.1.1. The decomposition (3.8) has a closed (modelled) part, ν_t , and unclosed (corrective) part, $\mathbf{f}_{\mathbf{c}}$. Equation (3.8) can then be substituted into Equation (3.5b) to derive the final SA augmented RANS equations,

$$\nabla \cdot \mathbf{U} = 0, \quad (3.9a)$$

$$\mathbf{U} \cdot \nabla \mathbf{U} + \nabla (P - \phi) - \nabla \cdot ((\nu + \nu_t) \nabla_{\mathbf{S}} \mathbf{U}) - \mathbf{f}_{\mathbf{s}} = \mathbf{0}, \quad (3.9b)$$

$$\nabla \cdot \mathbf{f}_{\mathbf{s}} = 0, \quad (3.9c)$$

$$\mathbf{U} \cdot \nabla \tilde{\nu} - S_p - S_{diff} - S_c - S_d = 0. \quad (3.9d)$$

Applying the turbulence model allows an isotropic component of Reynolds stresses to be approximated using the SA equation, leaving the anisotropic (and potentially an isotropic) component underdetermined. As a large portion of the Reynolds forcing is now modelled, data-assimilation techniques only need to apply smaller corrections, compared with Equation (3.7), to close the solution.

3.1.2 Harmonic Balanced Navier-Stokes Equations

In the previous section, it was shown how applying the Reynolds decomposition and time-averaging the Navier-Stokes equations produces the underdetermined RANS equations, which can then be used with data assimilation methods. In this section, an alternative approach is considered, by expressing the Navier-Stokes in the frequency domain, enabling the resolution of low-frequency harmonic flow scales (and consequently improving accuracy of data assimilation methods). This compares with the previous turbulence model augmented formulation. However, instead of adding prior physics knowledge to the data assimilation with an empirical turbulence model, which is based on several modelling assumptions, such as the Boussinesq approximation, and has been tuned for specific flow regimes, low-frequency flow scales are accurately resolved with the Navier-Stokes projected in the frequency domain (not unlike LES methods).

Firstly, the solution of the Navier-Stokes can be written as a Fourier series expansion (assuming the flow is at steady state), such that

$$\mathbf{w}(t) = \sum_{-\infty}^{\infty} e^{in\omega t} \hat{\mathbf{w}}_{\mathbf{n}}, \quad (3.10)$$

where $\hat{\mathbf{w}}_{\mathbf{n}}$ are spatial modes for the n -th frequency. Given that $\mathbf{w}$ is the flow field and therefore must be real, then $\hat{\mathbf{w}}_{-\mathbf{n}} = \overline{\hat{\mathbf{w}}_{\mathbf{n}}}$, where the overbar represents the complex conjugate. Furthermore, the $n = 0$ mode, $\hat{\mathbf{w}}_{\mathbf{0}}$, which is both real and does not vary in time, and represents the mean flow i.e $\hat{\mathbf{w}}_{\mathbf{0}} = [\mathbf{U}, P]$. One can apply the Fourier expansion (3.10) to the compact

form of the Navier-Stokes equations (3.3) and perform a harmonic balance (Fourier-Galerkin method), to express the equations in the frequency domain (see Appendix A.2). Furthermore, the Fourier series expansion is truncated to a large number, M , of discrete resolved harmonics, such that one can sufficiently represent the time-scales of the flow dynamics. This results in the Harmonic-Balanced Navier-Stokes equations (HBNS),

$$[in\omega\mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)]\hat{\mathbf{w}}_n + \sum_{\substack{n=n_1+n_2 \\ -M \leq n_1 \leq n_2 \leq M \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) = 0, \quad (3.11)$$

where ω is the fundamental frequency and the coefficient $\gamma_{n_1}^{n_2}$ is 0.5 for $n_1 = n_2$, otherwise 1. The solution to these equations are the complex modes, $\hat{\mathbf{w}} = [\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1, \dots]$.

Separating the low- and high-frequency dynamics

Additionally, one can consider decomposing the full flow solution, $\mathbf{w}$ into a low-frequency component, $\mathbf{w}_l$ and a high-frequency component, $\mathbf{w}_h$:

$$\mathbf{w}(t) = \mathbf{w}_l(t) + \mathbf{w}_h(t). \quad (3.12)$$

Using projection operators, $\mathcal{P}_l \equiv \mathbf{w}_l$ and $\mathcal{P}_h \equiv \mathbf{w} - \mathbf{w}_l = \mathbf{w}_h$, (sharp bypass filters), one can define the low- and high-frequency components,

$$\mathbf{w}_l(t) = \sum_{\|l\| \leq N} e^{il\omega t} \hat{\mathbf{w}}_l, \quad (3.13a)$$

$$\mathbf{w}_h(t) = \sum_{\|h\| > N} e^{ih\omega t} \hat{\mathbf{w}}_h. \quad (3.13b)$$

Here, N , represents the number of low-frequency scales. These projection operators have the following properties:

$$\mathcal{P}_l[\bullet] + \mathcal{P}_h[\bullet] = \mathcal{I}[\bullet], \quad (3.14a)$$

$$\mathcal{P}_l[\mathcal{P}_l[\bullet]] = \mathcal{P}_l[\bullet], \mathcal{P}_h[\mathcal{P}_h[\bullet]] = \mathcal{P}_h[\bullet], \quad (3.14b)$$

$$\mathcal{P}_l[\mathcal{P}_h[\bullet]] = \mathcal{P}_h[\mathcal{P}_l[\bullet]] = \mathbf{0}, \quad (3.14c)$$

where $\mathcal{I}[\bullet]$ is the identity operator. Likewise, using these operators, the HBNS equations (3.11) can also be projected onto the low- and high-frequency subspaces,

$$in\omega \mathbf{M} \hat{\mathbf{w}}_l = \mathcal{L}_P[\hat{\mathbf{w}}_l] + \mathcal{P}_l(\mathcal{N}_P[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_l] + \mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h] + \mathcal{N}_P[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_h] + \mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_l]), \quad (3.15a)$$

$$in\omega \mathbf{M} \hat{\mathbf{w}}_h = \mathcal{L}_P[\hat{\mathbf{w}}_h] + \mathcal{P}_h(\mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_l] + \mathcal{N}_P[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_h] + \mathcal{N}_P[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_l] + \mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h]), \quad (3.15b)$$

where $\mathcal{L}_P$ and $\mathcal{N}_P$ are the projected linear and non-linear operators respectively. This projection separates the low- and high-frequency dynamics.

Truncated Harmonic-Balanced Navier-Stokes

To compute the Fourier series expansion and subsequently solve the HBNS equations (see Eq. (3.11)), the large number of discrete resolved (time) scales, M must be computed using a discrete Fourier transform. However, for turbulent flows, resolution of a sufficient number of scales is computationally intractable. Instead, one can reduce the computational cost by resolving only the low-frequency scales (i.e. $\hat{\mathbf{w}}_l$ described in Eq. (3.13a)) by truncating the series to N low-frequency resolved harmonics, where $N \ll M$. As a consequence, the flow solution consists of $(2N + 1)$ complex modes $\hat{\mathbf{w}} = [\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_{-N}, \hat{\mathbf{w}}_N]$ and the HBNS equations are subsequently truncated as follows:

$$[in\omega \mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)] \hat{\mathbf{w}}_n + \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) = -\mathbf{R}_n, \quad (3.16)$$

where $\mathbf{R}_n$ is a residual term, that captures all the unresolved high-frequency scales. By neglecting the effect of all the unresolved high-frequency harmonic scales (i.e. $\hat{\mathbf{w}}_h$, such that $\mathbf{R}_n = 0$), the above system of equations becomes closed form and is non-linearly coupled. However, if the effects of these unresolved scales are significant (such as for a turbulent flow), neglecting $\mathbf{R}_n$ will result in an erroneous solution $\hat{\mathbf{w}}$. This can be observed by dividing the low-frequency projection, Equation (3.15a), into two parts: i) terms only including the resolved, low-frequency

harmonics and ii) terms which include the unresolved high-frequency harmonics.

$$\underbrace{in\omega\mathbf{M}\hat{\mathbf{w}}_1 + \mathcal{L}_P[\hat{\mathbf{w}}_1] + \mathcal{P}_l(\mathcal{N}_P[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1])}_{\text{HBNS}_N(\hat{\mathbf{w}}_1)} + \underbrace{\mathcal{P}_l(\mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h] + \mathcal{N}_P[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h] + \mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_1])}_{\mathbf{R}_n \equiv F(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h)} = \mathbf{0}, \quad (3.17)$$

where HBNS_N describes an operator for the HBNS equations, resolving only the N low-frequency harmonic scales (and is equivalent to the left-hand side of Eq. (3.16)) and $\mathbf{R}_n$ accounts for the effects of the high-frequency unresolved scales. For a turbulent flow, $\mathbf{R}_n$ has a large impact on the resolution of the low-frequency resolved scales, as they interact non-linearly with the high-frequency unresolved scales. In addition to the formulation in Eq. (3.17), $\mathbf{R}_n$ can also be expressed as

$$\mathbf{R}_n = \sum_{\substack{n=n_1+n_2 \\ -M \leq n_1 \leq n_2 \leq M \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) - \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}). \quad (3.18)$$

Whilst Eqs. (3.16) and (3.17) describes the flow field in terms of N resolved low-frequency harmonic terms, the solution depends on the closure term, $\mathbf{R}_n$, which itself depends on unresolved high-frequency components.

Comparing Truncated HBNS with RANS equations

Whilst the HBNS equations describe the flow in the frequency domain, one can also draw a comparison to the RANS equations from Section (3.1.1). Applying the $N = 0$ truncation of Eq. (3.16) results in

$$\nabla \cdot \hat{\mathbf{u}}_0 = 0, \quad (3.19a)$$

$$\hat{\mathbf{u}}_0 \cdot \nabla \hat{\mathbf{u}}_0 + \nabla \hat{p}_0 - \nu \Delta \hat{\mathbf{u}}_0 + \mathbf{R}_0 = 0. \quad (3.19b)$$

This is equivalent to the RANS equations, presented in Eq. (3.5), where $(\hat{\mathbf{u}}_0, \hat{p}_0) \equiv (\mathbf{U}, P)$, the residual term is the same as the divergence of the Reynolds stress tensor ($\mathbf{R}_0 \equiv \nabla \cdot \tau$). As a result, one can consider solving this problem and finding the residual term, $\mathbf{R}_0$, is the same as solving the RANS equations and finding the closure term. Therefore, applying data

assimilation to either of these problems is also equivalent.

Therefore, by resolving some N low-frequency harmonic scales, one can consider that a portion of the Reynolds stress terms, $\nabla \cdot \tau$, is now resolved and the “amount” of the residual term that is unclosed, capturing the interactions of the high-frequency, unresolved harmonics is reduced. This is shown in more detail in Appendix A.2.2, demonstrating how the truncated harmonics effect the low- and high-frequency dynamics.

Helmholtz decomposition on the residual term

In this research, data assimilation for flow reconstruction is performed using only resolved velocity measurements. Therefore, the Helmholtz decomposition is also applied to the $\mathbf{R}_n$ closure term, separating it into a potential and divergence-free, solenoidal component,

$$\mathbf{R}_n = -\nabla\phi_n + \mathbf{R}_{n,s}. \quad (3.20)$$

The potential forcing is thus absorbed into the pressure term of the momentum equations (i.e. $(\hat{p} - \phi)_n$) and the divergence-free condition is added to the full system

$$[in\omega\mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)]\hat{\mathbf{w}}_n + \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) + \mathbf{R}_{n,s} = 0, \quad (3.21a)$$

$$\nabla \cdot \mathbf{R}_{n,s} = 0, \quad (3.21b)$$

where $\hat{\mathbf{w}}$ now contains the composite pressure term $(\hat{\mathbf{p}} - \phi)$ i.e. $\hat{\mathbf{w}} = (\hat{\mathbf{u}}, (\hat{\mathbf{p}} - \phi))$. This system now contains in 3D, $5(N + 1)$ equations and $7(N + 1)$ unknowns ($4(N + 1)$ equations and $5(N + 1)$ unknowns in 2D). By providing both mean velocity data and harmonic resolved velocity fields (representing some of the unsteady fluctuating velocity component), the system can be closed at the discrete measurement locations and the (solenoidal) closure term found using data assimilation, although as before, $(\hat{\mathbf{p}} - \phi)$ can not be separated into constituent parts. This is the same as Plogmann et al. (2024a), where the inclusion of the residual closure

terms, accounting for the effect of high-frequency scales, requires the use of unsteady velocity component data, specifically its harmonic, Fourier components. An example of this is provided in Appendix A.2.2.

Turbulence model augmentation of HBNS equations - HBNS-SA

As with the RANS equations, by applying turbulence model augmentation to the HBNS equations, one can use eddy viscosity to model some of the unresolved scales. Firstly, similar to Eq. (3.10), the unsteady eddy-viscosity field is written as a Fourier series expansion,

$$\nu_t(t) = \sum_{-\infty}^{\infty} e^{in\omega t} \hat{\nu}_{tn}, \quad (3.22)$$

where $\hat{\nu}_{tn}$ are spatial eddy-viscosity modes for the n -th frequency. One can apply the above Fourier expansion to the unsteady modelled eddy-viscosity forcing, $\nabla \cdot (\nu_t(t) \nabla_{\mathbf{S}} \mathbf{u}(t))$, and perform a harmonic balance (Fourier-Galerkin method) as before, to express the modelled forcing in the frequency domain,

$$\sum_{\substack{n=n_1+n_2 \\ -N \leq n_1, n_2 \leq N}} \nabla \cdot (\hat{\nu}_{n_1} \nabla_{\mathbf{S}} \hat{\mathbf{u}}_{\mathbf{n}_2}). \quad (3.23)$$

Note that as with the HBNS equations, the solution has been truncated to N resolved harmonics, and that there is a modelled eddy-viscosity forcing component for each n -th frequency. Thereafter, the unresolved, residual forcing term can be decomposed into a modelled and corrective component, which itself is divided using a Helmholtz decomposition.

$$\mathbf{R}_{\mathbf{n}} = \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1, n_2 \leq N}} \nabla \cdot (\hat{\nu}_{n_1} \nabla_{\mathbf{S}} \hat{\mathbf{u}}_{\mathbf{n}_2}) + \mathbf{R}_{\mathbf{n},\mathbf{c}} = \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1, n_2 \leq N}} \nabla \cdot (\hat{\nu}_{n_1} \nabla_{\mathbf{S}} \hat{\mathbf{u}}_{\mathbf{n}_2}) + \underbrace{\nabla \phi_n + \mathbf{R}_{\mathbf{n},\mathbf{s}}}_{\mathbf{R}_{\mathbf{n},\mathbf{c}}}. \quad (3.24)$$

The turbulence model augmented HBNS equations can therefore be written as

$$\begin{aligned}
[in\omega\mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)]\hat{\mathbf{w}}_n + \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) \\
+ \mathbf{R}_{n,s} + \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1, n_2 \leq N}} \nabla \cdot (\hat{\nu}_{n_1} \nabla_{\mathbf{s}} \hat{\mathbf{u}}_{n_2}) = 0.
\end{aligned} \tag{3.25}$$

When truncating the solution to $N = 1$ resolved components (as in Chapter 8), the modelled eddy-viscosity terms added to the n -th momentum equations becomes

$$n = 0 : \quad \nabla \cdot (\hat{\nu}_{t,0} \nabla_{\mathbf{s}} \hat{\mathbf{u}}_0 + \hat{\nu}_{t,1} \nabla_{\mathbf{s}} \overline{\hat{\mathbf{u}}_1} + \overline{\hat{\nu}_{t,1}} \nabla_{\mathbf{s}} \hat{\mathbf{u}}_1), \tag{3.26a}$$

$$n = 1 : \quad \nabla \cdot (\hat{\nu}_{t,0} \nabla_{\mathbf{s}} \hat{\mathbf{u}}_1 + \hat{\nu}_{t,1} \nabla_{\mathbf{s}} \hat{\mathbf{u}}_0). \tag{3.26b}$$

The eddy viscosity is thus broken into a mean ($\hat{\nu}_{t,0}$) and harmonic ($\hat{\nu}_{t,1}$) component. However, in this work, the effect of the fluctuating, harmonic components of the modelled eddy-viscosity are neglected, due to their relatively small influence, as seen in von Saldern et al. (2023) (the effect of using the formulation of modelled eddy-viscosity without neglecting harmonic components, is evaluated in Appendix D.3). Using the SA equation to solve for the mean eddy viscosity, $\hat{\nu}_{t,0}$, the final HBNS-SA governing equations can be written as

$$n = 0 : [\mathbf{L} + \frac{1}{2} \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)]\hat{\mathbf{w}}_0 + \sum_{\substack{n_1+n_2=0 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) + \mathbf{R}_{0,s} + \nabla \cdot (\nu_t \nabla_{\mathbf{s}} \hat{\mathbf{u}}_0) = 0, \tag{3.27a}$$

$$n \geq 1 : [in\omega\mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)]\hat{\mathbf{w}}_n + \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) + \mathbf{R}_{n,s} = 0, \tag{3.27b}$$

$$\nabla \cdot \mathbf{R}_{n,s} = 0, \tag{3.27c}$$

$$\hat{\mathbf{u}}_0 \cdot \nabla \tilde{\nu} - S_p - S_{diff} - S_c - S_d = 0. \tag{3.27d}$$

One can observe that if N is sufficiently large (and approaches M), all non-linear interactions (i.e. $\mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2})$) are resolved and therefore the modelled forcing, ν_t , (and also the correc-

tive forcing, $\mathbf{R}_{n,s}$) tend to zero and are not required. It follows therefore that the HBNS-SA formulation is suitable for turbulent flows, in which only a few harmonics (producing coherent structures) are resolved, with the remaining high-frequency unresolved effects modelled with eddy viscosity (ν_t) or inferred with a corrective forcing component ($\mathbf{R}_{n,s}$) using data assimilation.

Finding the closure term, $\mathbf{R}_n$

In this work, in addition to flow reconstruction, the use of data assimilation methods to find HBNS closures will also be explored.

Examination of Equation (3.17) shows how accounting for the unresolved, high-frequency scales is equivalent to the HBNS equations with a closure term in equation (3.16). Specifically, the second part from Eq. (3.17), $F(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h)$, is directly analogous to the residual closure term, $\mathbf{R}_n$ (as defined in Eq. (3.18)) and captures the effects of interactions of high-frequency unresolved harmonic scales, $\hat{\mathbf{w}}_h$, on the low-frequency resolved harmonic scales, $\hat{\mathbf{w}}_1$. Furthermore, by solving Eq. (3.15b) to find $\hat{\mathbf{w}}_h$ as function of $\hat{\mathbf{w}}_1$, one can substitute it into $F(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h)$,

$$\mathbf{R}_n \equiv F(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h) = F(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h(\hat{\mathbf{w}}_1)) = G(\hat{\mathbf{w}}_1). \quad (3.28)$$

Consequently, the closure term $\mathbf{R}_n$ becomes exclusively a function of $\hat{\mathbf{w}}_1$. Furthermore, examining Eq. (3.15b) shows the presence of first- and second-order gradient terms in the linear and non-linear operators. It follows, therefore, that the residual term can finally be described as function of resolved velocity and resolved velocity gradient terms of the low-frequency harmonics:

$$\mathbf{R}_n \equiv G(\hat{\mathbf{w}}_1, \nabla \hat{\mathbf{w}}_1, \Delta \hat{\mathbf{w}}_1). \quad (3.29)$$

This is comparable to turbulence model approaches, such as the SA model. These turbulence models both i) model the closure term of the RANS equations as a function of the mean velocity (gradients) i.e. Boussinesq approximation and ii) contain mean velocity and its derivatives in the underlying turbulence transport equations, which are used to solve for the modelled (eddy-

viscosity) closure.

3.2 Data assimilation methods

Using data assimilation techniques to solve the constrained optimisation, a solution to the governing flow equations is sought, such that J , the measurement error (3.1), is minimised. Given that several formulations of the governing equations are used in this thesis (as summarised in Section 3.3), the formulation for the mean-flow equations both with and without turbulence model augmented approach will be presented in this section. However, use of other formulations (e.g. HBNS) only requires replacing the governing equations (and possibly data) that are enforced (and is summarised later).

For the turbulence model augmented formulation, the solution $\hat{\mathbf{U}}(\mathbf{x}), \hat{P}(\mathbf{x}) - \hat{\phi}(\mathbf{x}), \hat{\mathbf{f}}_s(\mathbf{x}), \hat{\tilde{\nu}}(\mathbf{x})$ to the mean-flow equations (3.7) or (3.9) is sought, which minimises the error to measurements, J . In the specific case where turbulence model augmentation is not applied, as in Equation (3.7), the eddy viscosity field $\tilde{\nu}(\mathbf{x}) = 0$. Two techniques are implemented to solve the constrained optimisation problem: a neural network approach, specifically the PINN approach from Sliwinski & Rigas (2022) and an adjoint-based variational approach, used by Foures et al. (2014) and Franceschini et al. (2020).

3.2.1 PINN

For a PINN, the solution $(\check{\mathbf{U}}, \check{P} - \check{\phi}, \check{\mathbf{f}}_s, \check{\tilde{\nu}})(\mathbf{x}; \theta)$, is defined by weights and biases, θ , which define a neural network. Figure 3.1 shows this PINN schematic. For any neural network, the optimal set of weights and biases, $\check{\theta}$, is sought by minimising some loss function C ,

$$\check{\theta} = \arg \min_{\theta} C(\mathbf{U}, P - \phi, \mathbf{f}_s, \tilde{\nu})(\mathbf{x}; \theta). \quad (3.30)$$

In the case of PINNs, the loss function, C , is used to enforce the governing equations and the

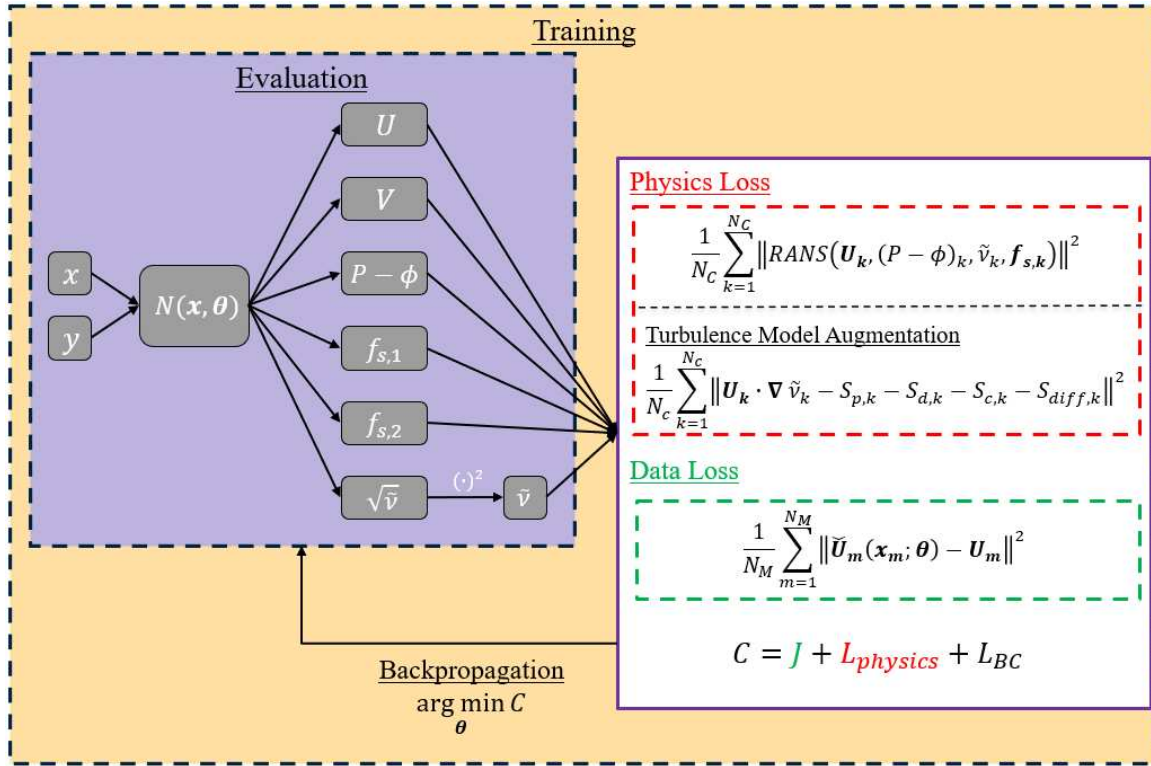


Figure 3.1: Structure of the Physics-Informed Neural Network as used in Chapter 6. A continuous mapping between spatial coordinates and flow variables is approximated by a deep neural network, denoted by N , defined by a set of weights and biases, θ . High-fidelity coarse data measurements and the (turbulence model augmented) RANS equations constrain the training of the network. k, m subscript the physics and data loss, respectively. i represents the i -th component of a vector.

boundary conditions. Whilst these two terms are sufficient to define an optimisation for the forward problem, when using data assimilation (the inverse problem) the cost function must also include the measurement error term, J . To distinguish the use of PINNs for the inverse problem (as in this research) compared to the forward problem, it will be referred to as PINN-DA hereon-in to emphasise the use of data. For the mean-flow problem described in Section 3.1.1, the loss function can be defined as

$$\begin{aligned}
 C = & \frac{\lambda^D}{N_M} J + \frac{\lambda^P}{N_C} \sum_{k=1}^{N_C} \|RANS((\mathbf{U}, P - \phi, \mathbf{f}_s, \tilde{v})(\mathbf{x}_k; \theta))\|^2 \\
 & + \frac{\lambda^P}{N_C} \sum_{k=1}^{N_C} \|(\mathbf{U} \cdot \nabla \tilde{v} - S_p - S_{diff} - S_c - S_d)(\mathbf{x}_k; \theta)\|^2 \\
 & + \frac{\lambda^B}{N_B} \sum_{b=1}^{N_B} \|BC((\mathbf{U}, P - \phi, \mathbf{f}_s, \tilde{v})(\mathbf{x}_b; \theta))\|^2 + \frac{\lambda^R}{2} \frac{1}{N_C} \sum_{k=1}^{N_C} \|\mathbf{f}_s(\mathbf{x}_k; \theta)\|^2,
 \end{aligned} \tag{3.31}$$

where the governing laws are evaluated at N_C collocation points (subscripted by k), and boundary conditions are imposed at N_B points along the boundary (subscripted by b). The data loss is subscripted by m . Here, $RANS$ defines an operator which calculates the residual of the mean-flow equations ((3.7) or (3.9)) and BC defines an operator which calculates the error to the imposed boundary conditions. To evaluate the governing laws, the gradients in the RANS equations are calculated accurately to machine-precision using automatic differentiation (leveraging the chain rule to trace the derivatives of all the constituent operations in the mapping $N(\mathbf{x}, \theta)$ from Figure 3.1). The data term, physics terms and boundary condition terms are weighted by factors λ^D , λ^P and λ^B , respectively. To select these weights, a grid-search method, adaptive weighting algorithms and spatial PDE loss weighting are trialled and are discussed in their respective chapters. In particular, the sensitivity to loss weighting is considered in Appendix B.1 and C.1.

The decomposition of the forcing into a modelled and corrective component (3.8) is not unique, as discussed in Foures et al. (2014). As a result, additional regularisation is applied to ensure uniqueness of the distribution between modelled and corrective forcing during optimisation. For the PINN-DA, an L_2 regularisation is used for its effectiveness and ease of implementation. This regularisation term (also evaluated at collocation points), weighted by λ^R , controls the magnitude of the corrective forcing field and subsequently the decomposition between modelled and corrective forcing. By increasing the penalisation of forcing magnitude, the importance of the modelled forcing component will grow and the governing equations will tend towards the RANS-SA equations (no corrective forcing). In the case without turbulence-model augmentation (effectively $\tilde{\nu}(\mathbf{x}) = 0$), $\lambda^R = 0$ and the SA equation can be removed from the loss function. The total loss C function (3.30) is minimised using backpropagation by iteratively adjusting the weights and biases based on the gradient of loss with respect to the network parameters, $\partial C / \partial \theta$.

As mentioned prior, for other formulations of PINNs (such as HBNS), one can substitute the terms in (3.31), with the governing HBNS equations ((3.21) or (3.27)) and the corresponding boundary conditions (e.g. through the $HBNS_N$ operator).

3.2.2 Variational

In the variational approach, the corrective forcing $\mathbf{f}_s$ is adjusted such that the measurement error, J , is minimised, constrained by the RANS equations (3.9). The solution to this minimisation problem is given by the stationary points of the Lagrangian

$$\begin{aligned} \mathcal{L}([\mathbf{U}, P - \phi, \tilde{\nu}, \mathbf{f}_s], [\mathbf{U}^*, P^*, \tilde{\nu}^*]) &= J - \langle P^*, \nabla \cdot \mathbf{U} \rangle \\ &\quad - \langle \mathbf{U}^*, \mathbf{U} \cdot \nabla \mathbf{U} + \nabla (P - \phi) - \nabla \cdot ((\nu + \nu_t) \nabla_s \mathbf{U}) - \mathbf{f}_s \rangle \\ &\quad - \langle \tilde{\nu}^*, \mathbf{U} \cdot \nabla \tilde{\nu} - S_p - S_{diff} - S_c - S_d \rangle. \end{aligned} \quad (3.32)$$

The inner product is defined as

$$\langle \mathbf{a}, \mathbf{b} \rangle = \int_{\Omega} \mathbf{a} \cdot \mathbf{b} d\Omega, \quad (3.33)$$

where $\mathbf{a}, \mathbf{b}$ are two spatial fields and Ω is the volume over which the field is defined. $\mathbf{U}^*, P^*, \tilde{\nu}^*$ are the Lagrange multipliers (or adjoint variables) enforcing the constraints. By setting the variation of the Lagrangian with respect to the direct flow variables $\mathbf{U}, P - \phi, \tilde{\nu}$, one can derive the adjoint equations

$$\nabla \cdot \mathbf{U}^* = 0, \quad (3.34a)$$

$$\mathbf{U}^* \cdot (\nabla \mathbf{U})^T - \mathbf{U} \cdot \nabla \mathbf{U}^* - \nabla P^* - \nabla \cdot ((\nu + \nu_t) \nabla_s \mathbf{U}^*) + \tilde{\nu}^* \nabla \tilde{\nu} + \nabla \cdot (\tilde{\nu}^* \partial_{\nabla \mathbf{U}} (S_p + S_d)) = \frac{\partial J}{\partial \mathbf{U}}, \quad (3.34b)$$

$$-\mathbf{U} \cdot \nabla \tilde{\nu}^* + (\partial_{\tilde{\nu}} \nu_t) \nabla \mathbf{U}^* : \nabla_s \mathbf{U} - \tilde{\nu}^* \partial_{\tilde{\nu}} (S_p + S_d) + \nabla \cdot (\tilde{\nu}^* \partial_{\nabla \tilde{\nu}} (S_c + S_{diff})) = 0. \quad (3.34c)$$

Furthermore, by calculating the gradient of Lagrangian, $\mathcal{L}$, with respect to changes in forcing, $\frac{\partial \mathcal{L}}{\partial \mathbf{f}_s}$, one finds $\frac{\partial \mathcal{L}}{\partial \mathbf{f}_s} = \frac{dJ}{d\mathbf{f}_s} = \mathbf{U}^*$. Specifically, this means the gradient $\frac{dJ}{d\mathbf{f}_s}$ used to set the minimisation direction for optimisation is the solution to the adjoint equations (3.34).

As with the PINN-DA-SA, regularisation is required to ensure uniqueness of the corrective forcing. Furthermore, the measurements are sparse, inducing a pointwise forcing of the adjoint momentum equations (3.34b) according to $\frac{\partial J}{\partial \mathbf{U}} = 2 \sum_{m=1}^{N_M} \delta(\mathbf{x} - \mathbf{x}_m)(\mathbf{U}(\mathbf{x}_m) - \mathbf{U}_m)$, which may lead to discontinuities in the adjoint field and in the gradient $\frac{dJ}{d\mathbf{f}_s}$, and thus ultimately in the reconstructed forcing and corresponding mean-flow solution. This can be circumvented through

H_1 -like regularisation of the gradient, which consists in getting a smoothed gradient $\frac{dJ^R}{d\mathbf{f}_s}$ from the original gradient $\frac{dJ}{d\mathbf{f}_s}$ through the inversion of the following system

$$\begin{pmatrix} \frac{1}{1+\beta}(I - \beta\Delta)\circ & \nabla\circ \\ -\nabla\cdot\circ & 0 \end{pmatrix} \begin{pmatrix} \frac{dJ^R}{d\mathbf{f}_s} \\ \pi \end{pmatrix} = \begin{pmatrix} \frac{dJ}{d\mathbf{f}_s} \\ 0 \end{pmatrix}, \quad (3.35)$$

where $\beta = l_r^2$ and l_r may be interpreted as a filter length. In the following, l_r is chosen as the spacing between measurement locations. Eq. (3.35) also involves a pseudo-pressure field π which is introduced to preserve the divergence-free character of the gradient through the regularisation procedure. A complete derivation and details of the variational approach can be found in Foures et al. (2014), Franceschini et al. (2020), Mons et al. (2022).

It has to be noted that for high Reynolds numbers turbulent cases, the variational-based data assimilation requires the use of the SA turbulence model augmentation in order to numerically solve the discretised equations, unlike the PINNs which work (albeit with different accuracy) with and without the SA model. Previously, the laminar studies in Foures et al. (2014) have not required this augmentation. To distinguish this approach it shall be referred to as variational-DA-SA.

3.2.3 Comparison: PINNs vs variational assimilation

Equations (3.31) and (3.32) highlight the similarities as well as key differences between the variational-DA-SA and PINN-DA approaches. Firstly, whilst the variational-DA-SA directly adjusts the corrective forcing, through the gradient $dJ/d\mathbf{f}_s$, the PINN-DA indirectly adjusts this forcing through adjustment of the network parameters, θ . Secondly, by the nature of PINNs, a continuous mapping of the flow solution is obtained, which can be queried at any spatial location. Additionally, the governing equations are evaluated and enforced at discrete locations within the domain (collocation points) and the spatial gradients in this mesh-free approach are determined through automatic differentiation. In the variational-DA-SA approach, solution of the direct and adjoint equations introduces mesh dependency, where the conservation laws are

enforced across the entire domain, whilst spatial gradients are calculated through discretisation (here Finite Element Method). As a result, discretisation errors are introduced, which are not present in PINNs. PINNs softly constrain the conservation laws in their continuous form whilst these are hard constraints in their discrete form for the variational-DA-SA. Lastly, the forcing regularisation method applied to the two approaches are different. The PINN-DA uses an L_2 -based regularisation unlike variational-DA-SA which uses a H_1 -based method, as discussed in Section 3.2.2. This is discussed in more detail in Chapter 6.

3.3 Summary of Equations and Data

The following table summarises the formulation of the governing equations (and data) used to constrain the data assimilation in each chapter.

Approach	Solution	Equations	Data	Chapters
DA-Baseline	$\mathbf{U}, (p - \phi), \mathbf{f}_s$	(3.7)	$\mathbf{U}$	5,6,7,8
DA-Baseline with Reynolds stresses	$\mathbf{U}, \overline{\mathbf{u}'\mathbf{u}'}$	(3.5)	$\mathbf{U}, \overline{\mathbf{u}'\mathbf{u}'}$	6,8
DA-SA	$\mathbf{U}, (p - \phi), \tilde{\nu}, \mathbf{f}_s$	(3.9)	$\mathbf{U}$	6,8
DA-HBNS	$\hat{\mathbf{u}}_{\mathbf{n}}, (p - \phi)_n, \mathbf{R}_{\mathbf{n},s}$	(3.21)	$\hat{\mathbf{u}}_{\mathbf{n}}$	7,8
DA-HBNS-SA	$\hat{\mathbf{u}}_{\mathbf{n}}, (p - \phi)_n, \tilde{\nu}, \mathbf{R}_{\mathbf{n},s}$	(3.27)	$\hat{\mathbf{u}}_{\mathbf{n}}$	8

Table 3.1: Summary of different formulations of the governing equations for mean-flow reconstruction, the required training data and the chapters, in which they are applied

One can see that when enforcing the HBNS equations, data for the resolved harmonic velocity terms must also be provided, as described in Section 3.1.2. Therefore, the measurement loss term, J , as previously described in Eq. (3.1), is adjusted as follows:

$$J = \sum_{m=1}^{N_M} \sum_{n=0}^N \left\| \hat{\mathbf{u}}_{\mathbf{n},m} - \check{\mathbf{u}}_{\mathbf{n}}(\mathbf{x}_m) \right\|^2. \quad (3.36)$$

Chapter 4

Flow Test Cases

The following chapter describes the flow cases, which are used to test the data assimilation methods proposed in this research. In order to perform data assimilation, high-fidelity data must be collected or generated either computationally or through experimental methods. The three flow cases described are a laminar circular cylinder flow (DNS, Section 4.1), a turbulent periodic hill flow (DNS, Section 4.2) and turbulent axisymmetric bluff body wake flow (experimental, Section 4.3).

4.1 2D laminar circular cylinder

The incompressible laminar circular cylinder flow is a well established test case, including the effects of number of resolved flow harmonics, as discussed by Rigas & Schmid (2022). Above $Re = u_\infty D / \nu = 47$ (where D is the cylinder diameter, u_∞ is the free-stream velocity and ν is the kinematic viscosity), the flow becomes linearly unstable, reaching a stable limit cycle with periodic vortex shedding, at the fundamental frequency ω_0 , which is dependant on the Reynolds number. This case can be solved by resolving N low-frequency harmonics using the frequency-truncated HBNS equations (see Eq. (3.16)), highlighting the importance of each resolved harmonic. Close to the critical Reynolds number (Hopf bifurcation), most energy is in the lower harmonics and a quasi-linear ($N = 1$ truncation) solution is sufficient to reach

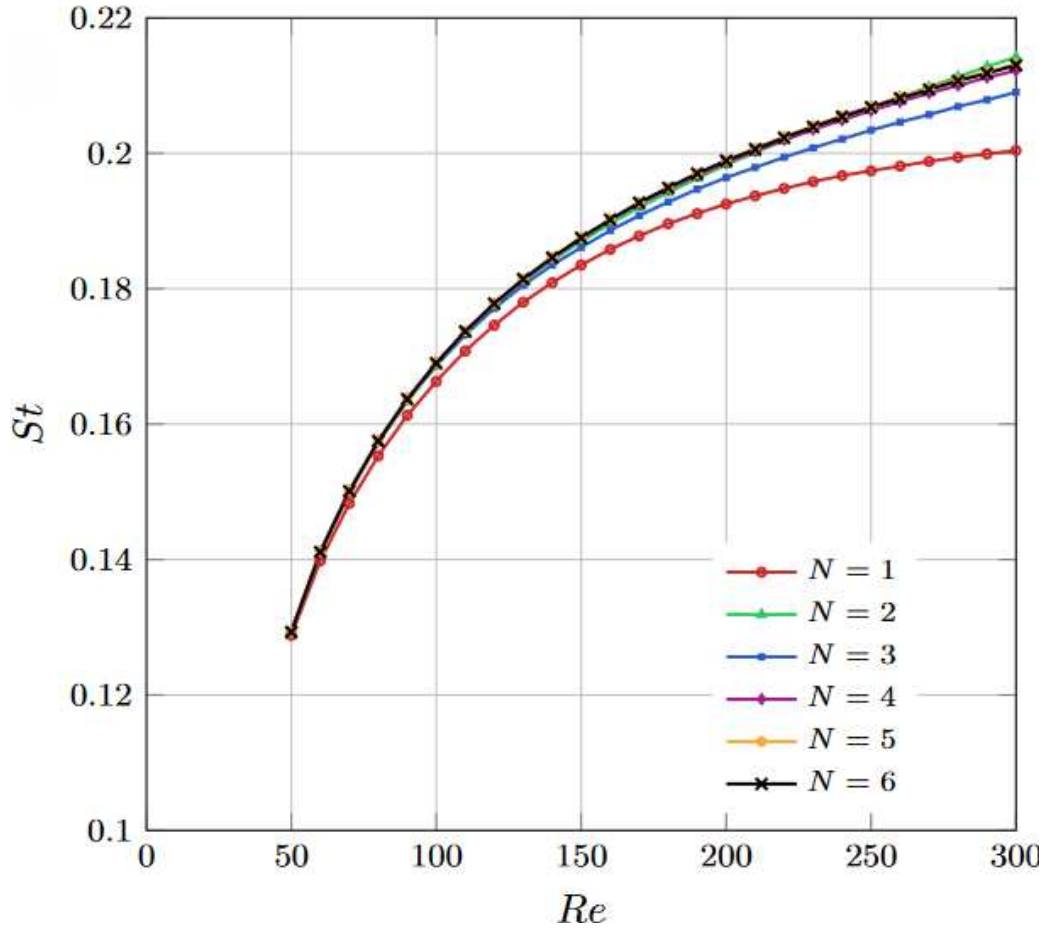


Figure 4.1: The effect of number of resolved harmonics on the converged Strouhal number of the laminar circular cylinder for different Reynolds numbers. Resolution of six harmonic scales is sufficient to achieve convergence of Strouhal number. Taken from Rigas & Schmid (2022).

an accurate and convergent solution. However, as Reynolds number increases away from the bifurcation, energy transfers to the higher harmonics and a higher order truncation of harmonics is required to converge to an accurate solution. In a convergence study, Rigas & Schmid (2022) show the influence of resolving a higher number of harmonics for a finite N number of harmonics, as shown in Figure 4.1 (from Rigas & Schmid (2022)). In the tested range of Reynolds numbers, $N = 4$ was sufficient to converge to a full DNS solution. More pertinently, as seen in Figure 4.1, at $Re = 200$, one can see the importance of resolved harmonics greater than $N = 1$, with an improvement in convergence and accuracy when truncating the HBNS equations and resolving $N = 2$ harmonics from $N = 1$. As a result, in order to evaluate the use of HBNS equations for data assimilation, the two-dimensional laminar circular cylinder flow at this Reynolds number ($= 200$) was selected.

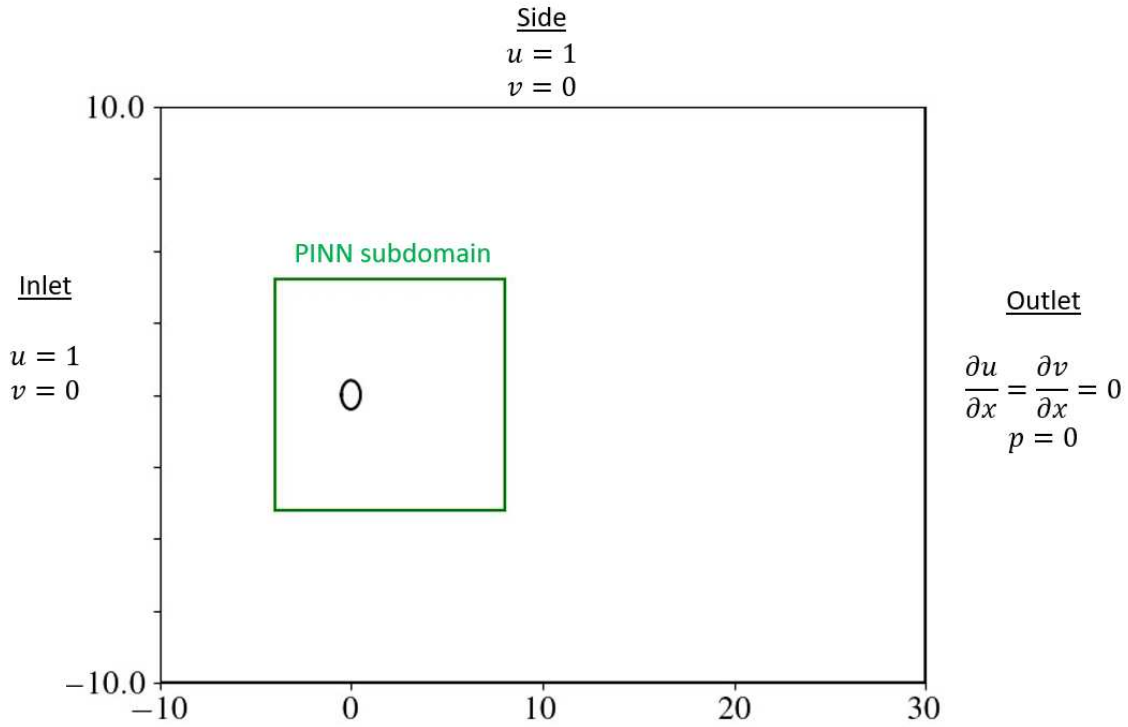


Figure 4.2: Numerical domain and boundary conditions for simulation of the laminar circular cylinder at $Re = 200$. The subdomain for flow reconstruction is also shown. Note that the cylinder body is circular and not an ellipsis as it appears, due to the axis scaling.

The remainder of this section describes how the high-fidelity simulation data is generated and the harmonic modes are calculated.

4.1.1 Numerical Simulation and set-up

Nektar++ (Cantwell et al. 2015), a spectral/hp element framework, was used to simulate the steady state solution of the cylinder flow (at $Re = 200$), which was subsequently used to extract high-fidelity measurements for data assimilation. Using a velocity correction scheme with convection advection, continuous Galerkin spatial projection and IMEX second order time integration, the unsteady Navier-Stokes equations were solved.

The 2D computational simulation domain for the laminar circular cylinder, presented in Figure 4.2, is defined as a rectangle, with $-10 \leq x \leq 30$ in the streamwise direction and $-10 \leq y \leq 10$ in the vertical direction. The cylinder, of diameter 1 is centred at $(0, 0)$ with a no-slip condition. At the inlet and sides, a fixed, uniform velocity was enforced, whilst at the outflow, a stress-

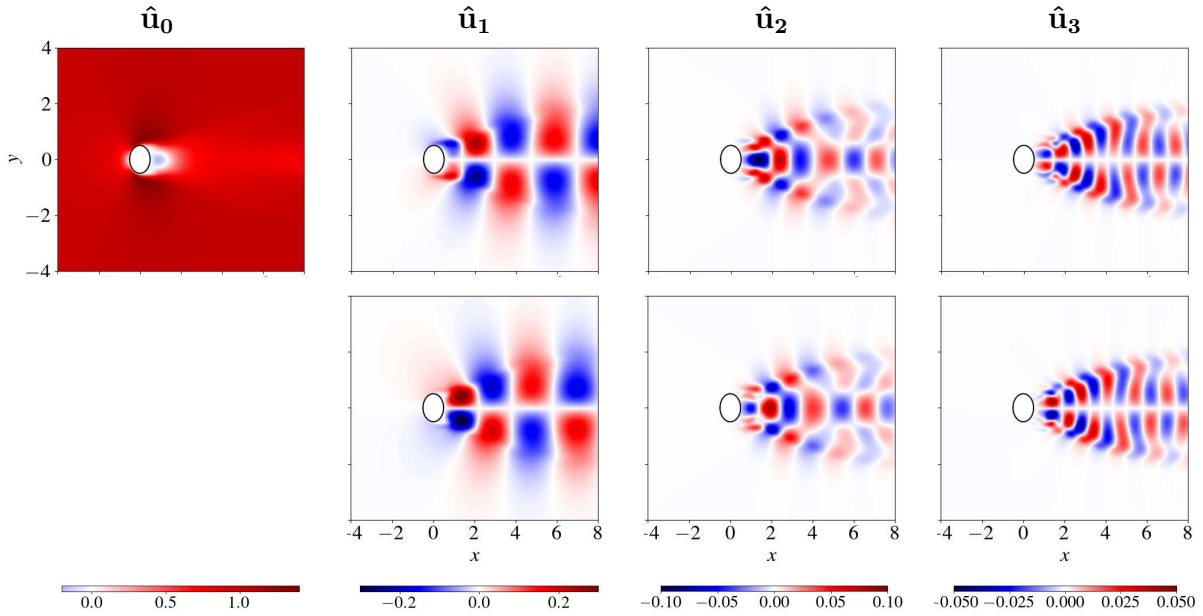


Figure 4.3: The mean mode and first three harmonic modes (extracted with DMD) of the streamwise component of $\hat{\mathbf{u}}$ for laminar circular cylinder flow at $Re = 200$. Real (top) and imaginary components (bottom).

free condition was enforced. To generate a mesh, the Nektar++ built-in mesher was used, to construct a triangular mesh with 15000 elements, and solved using a time-step of $0.025u_\infty/D$ for 12 vortex shedding cycles in the steady state limit cycle regime. The solution (Strouhal number, defined as $St = fD/u_\infty$, where f is the vortex-shedding frequency, and drag force) was validated against results from Rigas & Schmid (2022) ($St = 0.20$ compared with $St \approx 0.20$ from literature and $C_d = 0.686$ compared with $C_d \approx 0.675$ from literature - within 2%).

4.1.2 Dynamic Mode Decomposition

Data assimilation constrained by the HBNS equations requires the Fourier modes, as described in Eq. (3.10), to enforce the measurement error, J , as in Eq. (3.36). These Fourier mode shapes were extracted by applying a Dynamic Mode Decomposition (DMD) to the time-varying steady state laminar cylinder solution. DMD, developed in Schmid (2010), is a method for identifying coherent flow structures from data by determining modes which describe flow dynamics. The work by Chen et al. (2012) demonstrates the equivalency of DMD modes with discrete Fourier modes for data with zero mean. To achieve zero mean, the mean flow is subtracted from the

full unsteady flow solution. Consequently, using DMD, one can calculate a series of complex Fourier modes for each flow variable (i.e. two velocity components and pressure), as presented in Figure 4.3.

DMD, described in Schmid (2010), is an equation-free, data-based decomposition used to decompose a dynamic system $\mathbf{q}(t)$ which contains m snapshots of data into N DMD modes, ϕ , such that

$$\mathbf{q}(\mathbf{t}) \approx \sum_{k=1}^N \phi_{\mathbf{k}} e^{\omega_k t} \mathbf{a}_{\mathbf{k}}, \quad (4.1)$$

where a_k are the mode amplitudes, and ω_k are the corresponding modal frequencies. One can also see how this decomposition is comparable to a discrete Fourier series. To find the DMD modes, the data is first arranged into two data vectors, $\mathbf{Q}, \mathbf{Q}'$:

$$\mathbf{Q} = [\mathbf{q}_0 \dots \mathbf{q}_{m-1}], \mathbf{Q}' = [\mathbf{q}_1 \dots \mathbf{q}_m], \quad (4.2a)$$

$$\mathbf{Q}' = \mathbf{A}\mathbf{Q}, \quad (4.2b)$$

where $\mathbf{A}$ is an operator that captures the dynamics of $\mathbf{q}(t)$ and whose eigenvectors and eigenvalues are approximated by the DMD modes and eigenvalues

$$\mathbf{Q} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*, \quad (4.3a)$$

$$\tilde{\mathbf{A}} = \mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{U}^* \mathbf{Q}' \mathbf{V} \mathbf{\Sigma}^{-1}. \quad (4.3b)$$

By computing a reduced SVD of $\mathbf{Q}$ (Equation 4.3a), and projecting $\mathbf{A}$ onto the matrix $\mathbf{U}$ (Equation 4.3b), one can compute the smaller matrix, $\tilde{\mathbf{A}}$, which has the same eigenvalues as $\mathbf{A}$. Finally, one can compute the mode shapes, Φ of $\mathbf{A}$ by performing an eigen-decomposition of $\tilde{\mathbf{A}}$ using

$$\tilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\mathbf{\Lambda}, \quad (4.4a)$$

$$\Phi = \mathbf{Q}' \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{W}. \quad (4.4b)$$

By solving the described system, one can find the final DMD modes, Φ . The first three resultant DMD modes of streamwise velocity and the mean velocity have been presented in Figure 4.3.

4.2 Periodic Hill

The canonical turbulent periodic hill has been used in many data-driven machine learning problems, such as Brener et al. (2021). When testing data assimilation methods for turbulent mean flows, it is important to evaluate their capacity to capture common turbulent flow features. The periodic hill case at $Re = U_b H / \nu = 5600$ (where H is the hill height, U_b is the bulk velocity over the hill apex and ν the kinematic viscosity) contains separation, reattachment, recirculation as well as turbulent boundary layers. This makes the turbulent periodic hill both an interesting and useful case to test the capability of the data assimilation methods.

The rest of this section will describe the details of the flow case and the numerical setup of the simulation employed.

4.2.1 The periodic hill set-up

The high-fidelity data for the periodic hill is obtained from the DNS database¹ using the Incompact3D solver (Laizet & Li 2011) and the simulation setup is detailed in Xiao et al. (2020). The periodic hill is simulated at $Re = 5600$. Out of the several geometry configurations in the database, this work uses the canonical configuration with the domain length given by $L_x / H = 3.858\alpha + 5.142$ for which the hill stretch factor is $\alpha = 1.0$. The no-slip condition is enforced along the upper and lower walls, whilst periodicity is enforced in the streamwise direction. In order to maintain the required massflow and thus Reynolds number, body forcing is applied in the periodic streamwise direction. This is determined at each iteration, such that the Reynolds number matches the preset Reynolds number.

¹The data is available from <https://github.com/xiaoh/para-database-for-PIML>

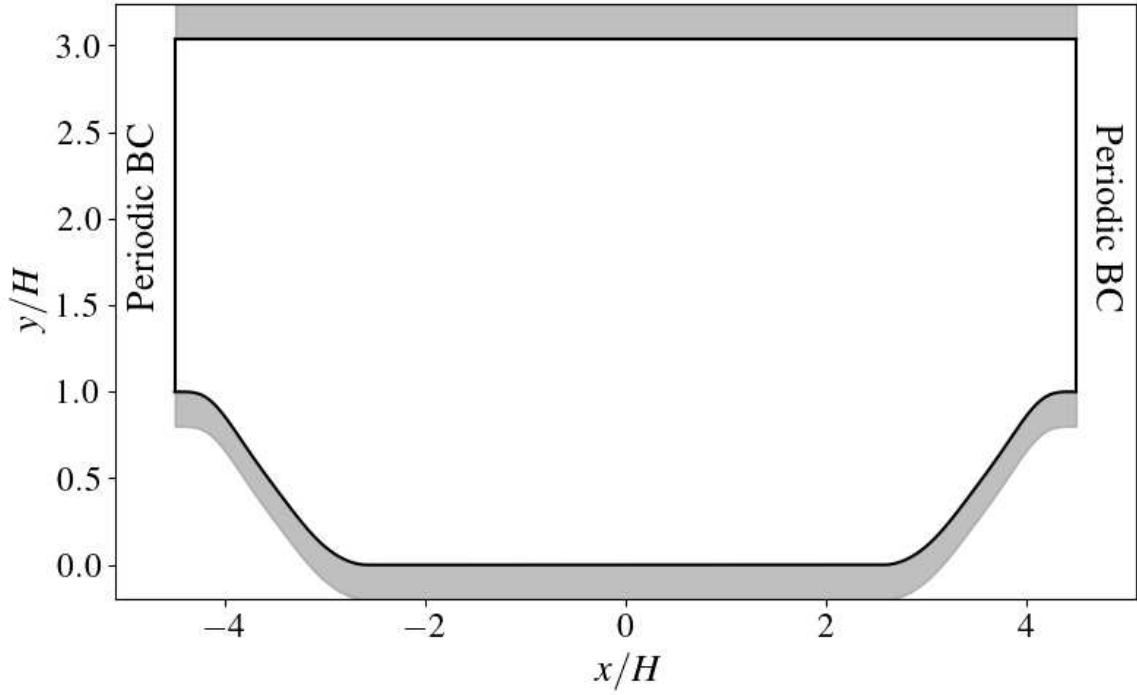


Figure 4.4: Periodic hill domain and boundary conditions simulated at $Re = 5600$.

4.2.2 Comparison with a RANS-SA simulation

The DNS mean-flow (time-averaged) solution, for streamwise velocity U is shown in Figure 4.5(a). The streamline (solution to the streamfunction) dividing the recirculation region from the rest of the flow is also shown. The low-fidelity RANS solution with the SA turbulence model is shown in Figure 4.5(b). Whilst the DNS reattachment occurs at $x/H = 0$, the RANS solver massively overpredicts the size of the recirculation region, as shown. The error field likewise, in Figure 4.5(c), shows high error across the entire domain. This highlights the limitations of using a RANS-SA simulation and demonstrates the potential advantages that can be gained from data assimilation.

4.3 Axisymmetric bluff body wake

The final case used in this research, to evaluate the use of data assimilation, is the axisymmetric, turbulent bluff body wake. Whilst the previous two cases are useful to evaluate data assimilation methods, they have certain limitations. Firstly, these flow cases have been numerically

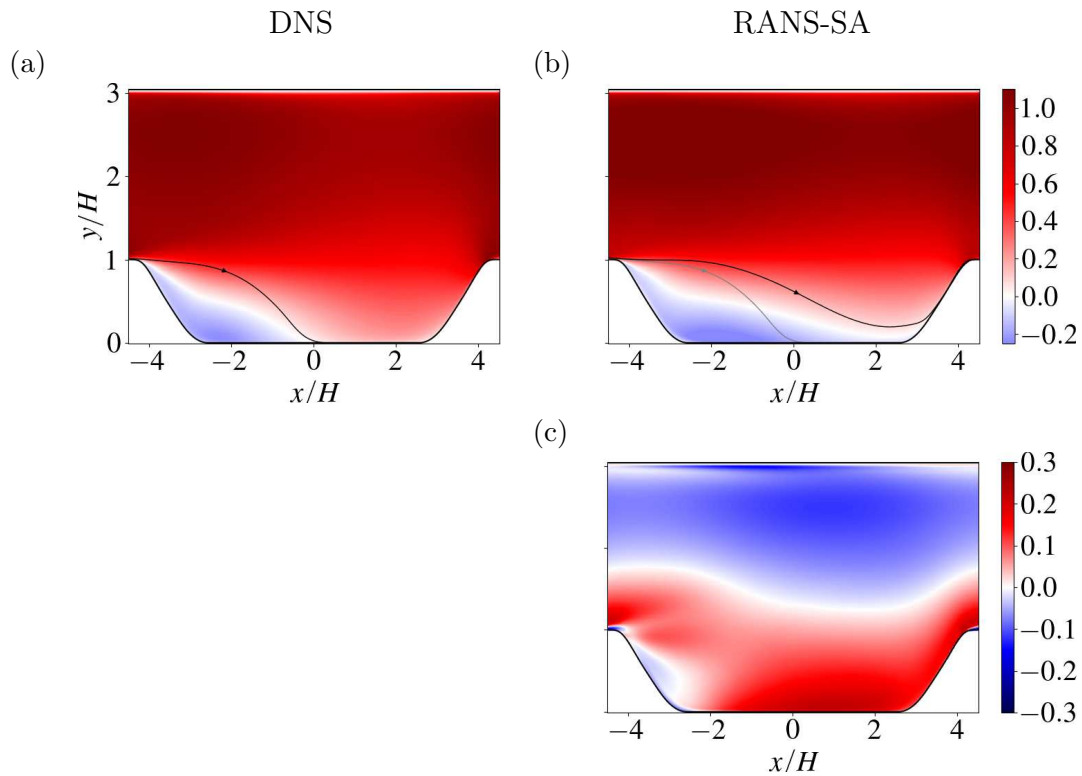


Figure 4.5: High-fidelity (DNS) solution of the flow over a periodic hill compared with the RANS solution using an SA model. The dividing streamline between the recirculation region and rest of the flow is shown. (a) Streamwise mean velocity - DNS; (b) streamwise mean velocity - RANS; (c) absolute error of streamwise velocity relative to DNS solution ($U_{err} = U_{DNS} - U_{RANS}$). For comparison, the DNS dividing streamline is shown in grey.

simulated and thus when used for data assimilation are treated as “synthetic” data. Therefore, challenges associated with the use of real experimental data are avoided. Secondly, these cases are not particularly representative of realistic engineering problems, as the Reynolds number is relatively small compared to engineering flows. The bluff body wake flow has been produced using experimental methods and is at a high Reynolds number $Re = U_{\infty}D/\nu = 1.88 \times 10^5$ (where U_{∞} is the free-stream velocity, D is the bullet diameter and ν is the kinematic viscosity), representative of real engineering flow problems. Testing data assimilation on this experimental high Reynolds number cases directly tackles the limitations of the previous cases and validates the approaches on a more complex flow case, meeting one of the key objectives from Section 2.4.

The remainder of this section will be used to describe the flow case, the experimental setup and how the data is post-processed.

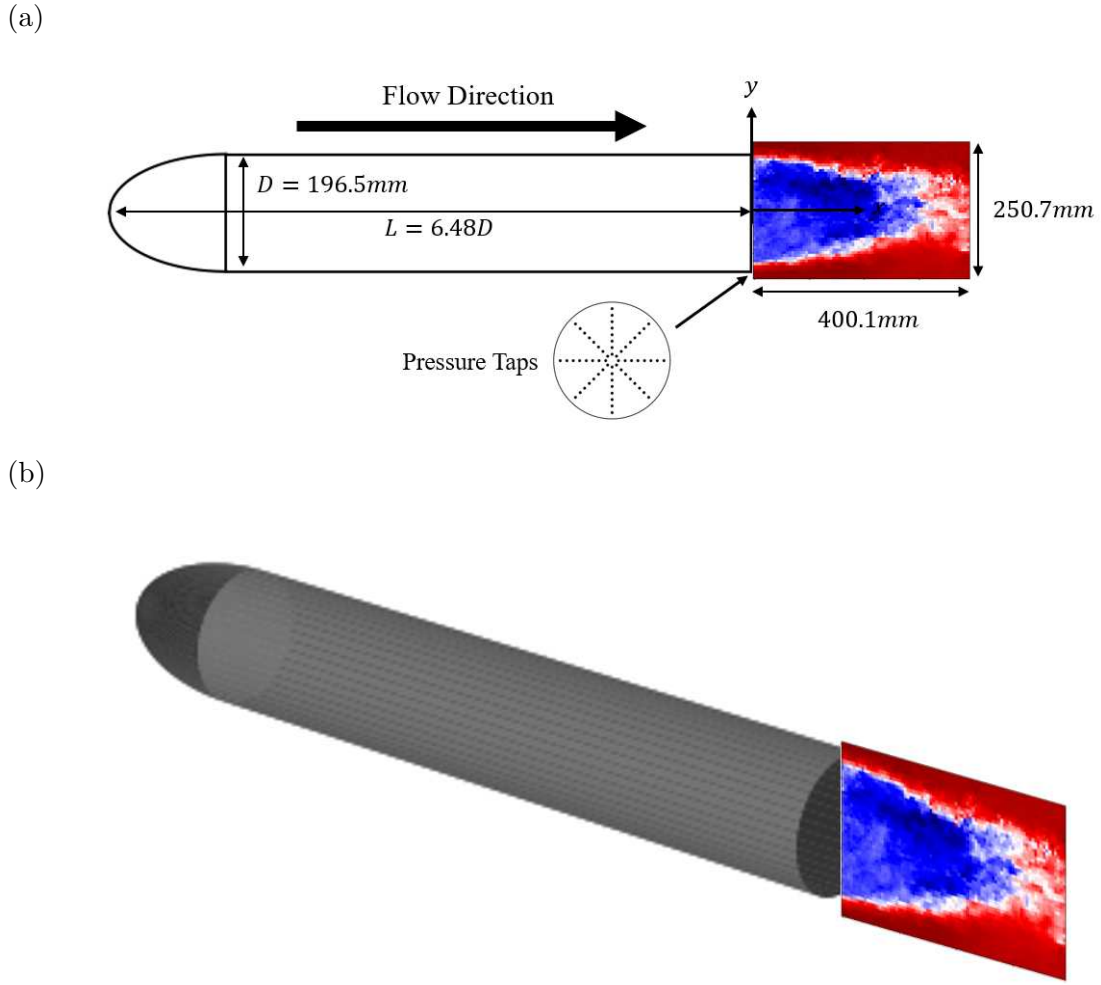


Figure 4.6: Experimental setup for axisymmetric bluff body wake as performed by Zhu (2023), including the position of the 2D2C PIV plane, aft of the body, and the pressure taps on the rear surface.

4.3.1 Experimental setup

The experiment for the axisymmetric bluff body wake was conducted in a wind tunnel of test section 1.15 m high, 1.61 m wide and 4 m long, performed by Zhu (2023). The wind tunnel has a turbulence intensity and mean velocity variation (at the wind tunnel cross section) of under 1%. To achieve a $Re = 1.88 \times 10^5$, the wind tunnel inlet velocity, U_∞ , was set at 15 m/s, based on the bluff body diameter, $D = 0.1965$ m. The bluff body shape and experimental setup is presented in Figure 4.6. The bullet-shaped axisymmetric body has a diameter, $D = 0.1965$ m and a body length to diameter ratio of $L/D = 6.48$. The nose shape is elliptical and defined using the ellipse equation. Experimental measurement data generated from this setup was used

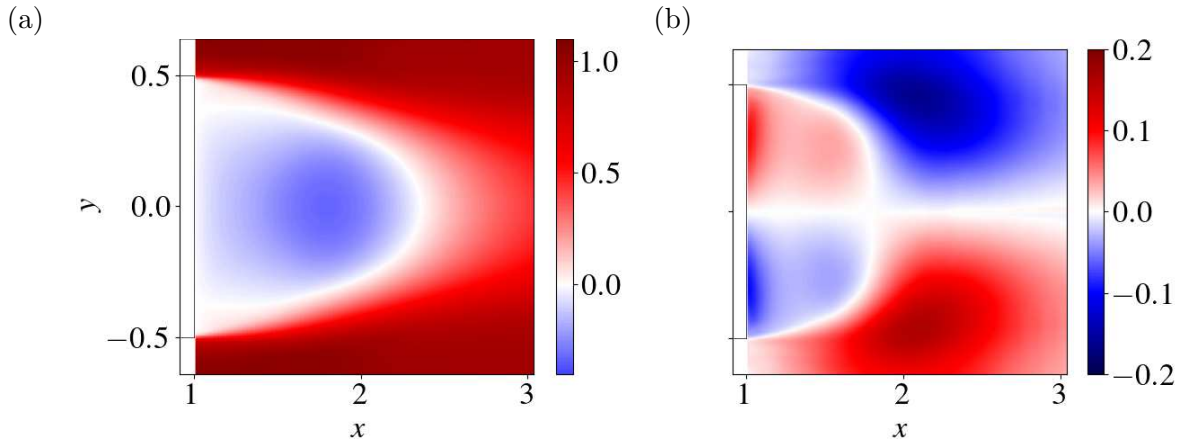


Figure 4.7: PIV mean velocity fields for the axisymmetric bluff body wake at $Re = 1.88 \times 10^5$ for (a) streamwise and (b) vertical components.

for data assimilation.

To generate the dataset, a time resolved 2D2C PIV was used, to provide velocity data for two velocity components on a single plane, as shown in Figure 4.6. The PIV field of view is 0.4001 m wide and 0.2507 m high and aligned to the vertical centreline of the bluff body, sitting immediately behind the trailing edge (there is a gap of 0.003 m). To capture the PIV fields, a Phantom VEO 640L camera with a Nikkor 135 lens was used. A dual-port Litron LDY300 laser was used to create the laser sheet for seed particle visualisation, whilst LaVision DaVis 10 software was used to post-process the raw data. The experiment was repeated thirteen times. For every repeat, data was collected, containing 5990 velocity field snapshots on a 158×100 uniform grid, sampled at 720 Hz (and therefore lasting 8.32s). The entire dataset contains 108 seconds of velocity data. This dataset was also used in Mo et al. (2024), in which a method was proposed to interpret the latent space of a non-linear auto-encoder.

Before using the data, several post-processing steps were required. All data used to train and evaluate data was non-dimensionalised by the free-stream, $U_\infty = 15$ m/s, and body diameter, $D = 0.1965$ m. This non-dimensionalised velocity field was used to determine the time-averaged mean velocity fields, shown in Figure 4.7. One can see how this test case, as captured in the PIV fields, show the bluff body wake with a large recirculation region and the shear layer that comes away from the bluff body sides. The final post-processing step applied to the data, is used to extract harmonic information about the flow and is presented below in Section 4.3.2.

4.3.2 Spectral Proper Orthogonal Decomposition

Spectral Proper Orthogonal Decomposition (SPOD) is another method used to identify coherent flow structures from data. Unlike a typical POD, SPOD, first described in Lumley (1970), identifies modes which vary in both space and time. Importantly, assuming flow is statistically stationary, the resulting SPOD modes are harmonic and describe flow structures occurring at single discrete frequencies. For this experimental bluff body wake, SPOD was used to extract the harmonic modes, which were then used to evaluate the use of data assimilation with the HBNS equations for this flow case.

The relationship between SPOD and DMD (as used for the laminar circular cylinder in Section 4.1.2) is explored in Towne et al. (2018). In Towne et al. (2018), it is shown that SPOD can be considered as optimally averaged DMD modes from an ensemble DMD problem for turbulent flows. Furthermore, SPOD modes represent structures that are dynamic as in DMD, but also capture the statistical variability of turbulent flows. Whilst DMD is suitable for the laminar circular cylinder flow, for broadband flows, including the turbulent, axisymmetric bluff body wake flow, SPOD is more appropriate (as discussed in more detail in Section 8.1.1).

To calculate SPOD modes for the bluff body wake, the algorithm based on Welch's method is used and is detailed in Schmidt & Towne (2019). Firstly, the time-resolved flow data, $\mathbf{q}(\mathbf{x}, t)$, is divided into N_b blocks, each containing the a time series of N_f data snapshots, as

$$\mathbf{Q}^{(n)} = \left[\mathbf{q}_1^{(n)}, \mathbf{q}_2^{(n)}, \dots, \mathbf{q}_{N_f}^{(n)} \right], \text{ for } n = 1, 2, \dots, N_b, \quad (4.5)$$

where $\mathbf{Q}^{(n)}$ is the data for the n -th data block, containing the N_f snapshots from this n -th data block, $\mathbf{q}_k^{(n)}$ (k is the index of the individual snapshot within the n -th block). The fast Fourier transform (FFT) is applied to each data block, to find the N_f Fourier modes, $\hat{\mathbf{q}}_k^{(n)}$, of each data block. $\hat{\mathbf{q}}_k^{(n)}$ is the k -th Fourier mode at frequency f_k for the n -th data block.

$$\hat{\mathbf{Q}}^{(n)} = FFT(\mathbf{Q}^{(n)}) = \left[\hat{\mathbf{q}}_1^{(n)}, \hat{\mathbf{q}}_2^{(n)}, \dots, \hat{\mathbf{q}}_{N_f}^{(n)} \right], \text{ for } n = 1, 2, \dots, N_b. \quad (4.6)$$

For each frequency, f_k , the matrix $\hat{\mathbf{Q}}_{f_k}$ is constructed from the k -th Fourier modes from each block as

$$\hat{\mathbf{Q}}_{f_k} = \sqrt{\kappa} \left[\hat{\mathbf{q}}_k^{(1)}, \hat{\mathbf{q}}_k^{(2)}, \dots, \hat{\mathbf{q}}_k^{(N_b)} \right], \text{ for } k = 1, 2, \dots, N_f. \quad (4.7)$$

Finally, one can solve the eigenvalue problem, for each $\hat{\mathbf{Q}}_{f_k}$

$$\hat{\mathbf{Q}}_{f_k}^* \mathbf{W} \hat{\mathbf{Q}}_{f_k} \boldsymbol{\Theta}_{f_k} = \boldsymbol{\Theta}_{f_k} \tilde{\Lambda}_{f_k}, \quad (4.8a)$$

$$\tilde{\Psi}_{f_k} = \hat{\mathbf{Q}}_{f_k} \boldsymbol{\Theta}_{f_k} \tilde{\Lambda}_{f_k}^{-\frac{1}{2}}, \quad (4.8b)$$

where $\mathbf{W}$ is a weighting matrix, $\boldsymbol{\Theta}_{f_k}$ are the eigenmodes and corresponding $\tilde{\Lambda}_{f_k}$ modal energies for k -th discrete frequency. Thereafter, the SPOD modes, $\tilde{\Psi}_{f_k}$, at this frequency can be calculated. These SPOD modes are ranked in terms of energy at the k -th frequency and are subsequently used to evaluate use of the HBNS equations with data assimilation problems for the experimental bluff body wake.

Chapter 5

Methodology for correcting spatially misaligned experimental measurements using transformation embedded PINNs

In this chapter, the data assimilation problem for flow reconstruction is considered, using partial observations, which have a systematic measurement error (or bias). In particular, spatial misalignment error is explored, where sensors (and thus measurements) are not exactly where they are assumed to be, resulting in a spatial offset of the observation data. A method for determining this systematic spatial misalignment error using Physics-Informed Neural Networks (PINNs) based on affine transformations is proposed and validated on a 2D cylinder flow.

In an experiment, sensors are used to extract information about the system. For example, in a fluid experiment, PIV can be used to measure flow velocity. However, experimental methods often have some source of error. For example, Zhou & Grauer (2023) examine the effects of particle inertia on velocity measurements (and apply a PINN approach to mitigate the effects), whilst Mo & Magri (2024) consider flow reconstruction from noisy measurement data. During laboratory experiments, there are often small errors in the sensor calibration. This introduces a corresponding systematic error or bias in the underlying measurements, such as in the nominally measured locations. Whilst small percentage differences are unnoticeable for basic flow analysis,

the measured data will not accurately obey the governing equations and boundary conditions of the system. This is particularly problematic for physics-informed data-driven methods, such as PINNs. Using biased measurements to build a model will embed this error into the trained model.

For many PINN flow reconstruction studies, such as in Sliwinski & Rigas (2022), synthetic training data is used, which is generated using high-fidelity numerical methods. As a result, the aforementioned systematic measurement error does not exist (although some studies do examine the effect of noise, as mentioned in Chapter 2). Given one of the objectives of this thesis is to apply PINNs to a realistic engineering use case, one must consider the effect of a systematic measurement error. Kelshaw & Magri (2024) explore the impact of reconstructing flow from data with spatially varying systematic measurement error (epistemic uncertainty) caused by sensor miscalibration, using PC-CNNs. In this chapter, an alternative source of systematic measurement error is investigated, spatial misalignment error. This occurs when sensors and measurements are not exactly where they are nominally assumed to be, and thus the measurements have a spatial offset. Specifically, cases where incorrect calibration of a PIV set up has led to spatial misalignment error in the measured velocity field are considered. Incorrect calibration of the PIV camera setup can be caused by factors, such as incorrect camera focus or misalignment of PIV planes. As a result, the nominal measured velocity field may be translated (from its true location), stretched or rotated. These small errors have implications on the physical accuracy (residual of the governing equations, assuming no modelling error) of the solution. To enable use of miscalibrated measurements for data assimilation, PINNs will be used to learn a transformation matrix (parameterised by affine transformations), to correct the measured velocity field.

5.1 Methodology

5.1.1 Defining the transformation matrix

In fields, such as image processing, affine transformations are commonly used to map and preprocess images. An affine transformation describes a linear matrix mapping, which can be used to transform a field. For a 2D domain, this can be written as

$$\mathbf{T} = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \beta_x \\ \alpha_{yx} & \alpha_{yy} & \beta_y \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.1)$$

This matrix can be used to transform the spatial coordinates of the measured PIV field, $\mathbf{x}_m$, to the real coordinates, $\mathbf{x}$, using the following equation:

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \beta_x \\ \alpha_{yx} & \alpha_{yy} & \beta_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_m \\ y_m \\ 1 \end{pmatrix} = \mathbf{T}\mathbf{x}_m. \quad (5.2)$$

Different transformation types, such as translation, scaling and rotation, can be independently defined. To scale a field, the following matrix can be used.

$$\mathbf{T}_s = \begin{pmatrix} S_{xx} & 0 & 0 \\ 0 & S_{yy} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (5.3)$$

where S_{xx} and S_{yy} describe stretching in the x- and y-axis respectively. Similarly, an anti-clockwise rotation, about the origin, by an angle of θ is described as

$$\mathbf{T}_r = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.4)$$

Finally, a translation by b_x and b_y in x and y directions can be defined as

$$\mathbf{T}_t = \begin{pmatrix} 1 & 0 & b_x \\ 0 & 1 & b_y \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.5)$$

Each of these matrices can be multiplied together to create a single transformation matrix, such that $\mathbf{T} = \mathbf{T}_t \mathbf{T}_r \mathbf{T}_s$. By assuming one can describe the spatial sensor misalignment using this affine transformation, $\mathbf{T}$ can be used to describe the mapping between measured and true coordinates.

When transforming the coordinate system, there is also an effect on the velocity vectors. For example, if the measured PIV velocity field is rotated, the velocity vectors will no longer align with the new transformed vector directions. This can be corrected by also applying a transformation matrix to the measured velocity field. Firstly, given that translations do not affect the velocity components (translational invariance), one can neglect the translation component of the complete transformation matrix, $\mathbf{T}$, such that

$$\mathbf{T}_v = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} \\ \alpha_{yx} & \alpha_{yy} \end{pmatrix}. \quad (5.6)$$

Thereafter, the true velocity field can be mapped to the measured PIV velocity field using the inverse of $\mathbf{T}_v$,

$$\mathbf{u}_m = \mathbf{T}_v^{-1} \mathbf{u}, \quad (5.7)$$

where $\mathbf{u}_m$ is the measured PIV velocity field and $\mathbf{u}$ is the true field.

The above formulation considers a 2D field and domain. However, the same approach can be used for a three-dimensional domain by increasing the size of the transformation matrix to a 4×4 .

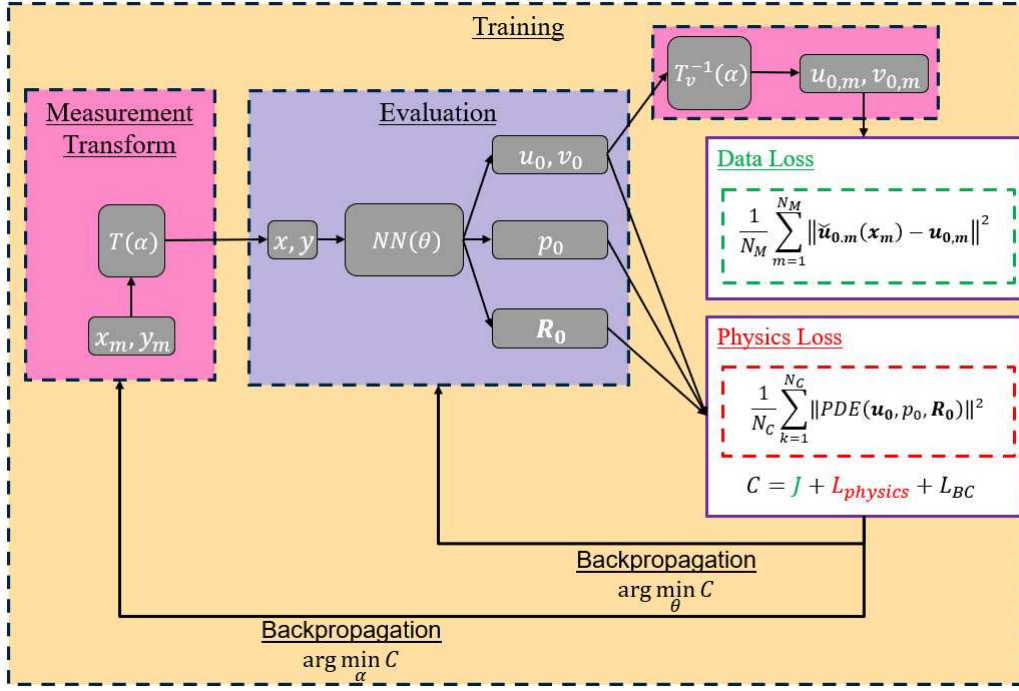


Figure 5.1: Transformation embedded PINN architecture used to learn and correct spatial measurement misalignment. Importantly, one can see the introduction of the embedded transformation block of learnable parameters α , such that they correct the measurement locations.

5.1.2 Applying and learning the transformation matrix using PINNs

To find the transformation matrix, $\mathbf{T}$, where α represents the matrix elements, the transformation matrix is embedded into the PINN architecture, as presented in Figure 5.1. The measurement error function is thus modified to incorporate the transformed measurements,

$$J = \sum_{m=1}^{N_m} \|\hat{\mathbf{u}}_{0,m} - \tilde{\mathbf{u}}_{0,m}\|^2 = \sum_{m=1}^{N_m} \|\hat{\mathbf{u}}_{0,m} - \mathbf{T}_v^{-1} \tilde{\mathbf{u}}_0(\mathbf{T}\mathbf{x}_m; \boldsymbol{\theta})\|^2. \quad (5.8)$$

Provided spatially misaligned, miscalibrated and thus biased measurements, a constrained optimisation problem can be constructed, such that a flow solution is sought, which obeys the governing equations, whilst also minimising error to measurements, accounting for some unknown miscalibration error. This enables the use of spatially misaligned data to train the PINN. Whilst typically the optimal set of neural network weights and biases, θ , are sought, the transformation parameters, α , are also included in the optimisation problem,

$$\hat{\theta}, \hat{\alpha} = \arg \min_{\theta, \alpha} C. \quad (5.9)$$

Now optimal weights and measurement transformation are sought, that minimise residual PDE and BC error, but also correct the measurements such that they better match the underlying physics and boundary conditions. To update the weights, the gradient of the cost functions must be calculated. For the model weights this is $\frac{\partial C}{\partial \theta}$ and for the transformation parameters this is $\frac{\partial C}{\partial \alpha}$. Using gradient descent, these learnable parameters can be updated as:

$$\theta_{n+1} = \theta_n + \frac{\partial C}{\partial \theta} \Delta \theta, \quad (5.10a)$$

$$\alpha_{n+1} = \alpha_n + \frac{\partial C}{\partial \alpha} \Delta \alpha. \quad (5.10b)$$

5.2 Validating use of an embedded transformation matrix on laminar circular cylinder case

In this section, the proposed embedded transformation approach is tested and validated on the 2D laminar circular cylinder (see Section 4.1) using the same PINN-DA-Baseline approach as in Sliwinski & Rigas (2022). This enables a known transformation to be applied to a high-fidelity numerical dataset, creating a “synthetic” misaligned dataset, to simulate the effect of a miscalibrated PIV camera setup, as discussed in the previous section. The training dataset was modified with predefined, known systematic errors, which can be learnt using the transformation embedded PINN-DA, such that the true transformation, $\mathbf{T}_{\text{calc}}$, and the estimated transformation, $\mathbf{T}_{\text{est}}$, could be compared.

The proposed approach was validated using the following two steps:

- Firstly, each transformation type (translation, rotation, scaling) was independently applied, to evaluate how well the transformation embedded approach was able to estimate each type of transformation.
- Secondly, the different transformations are combined at random magnitudes, to evaluate the capacity of the proposed approach to estimate more complex transformations.

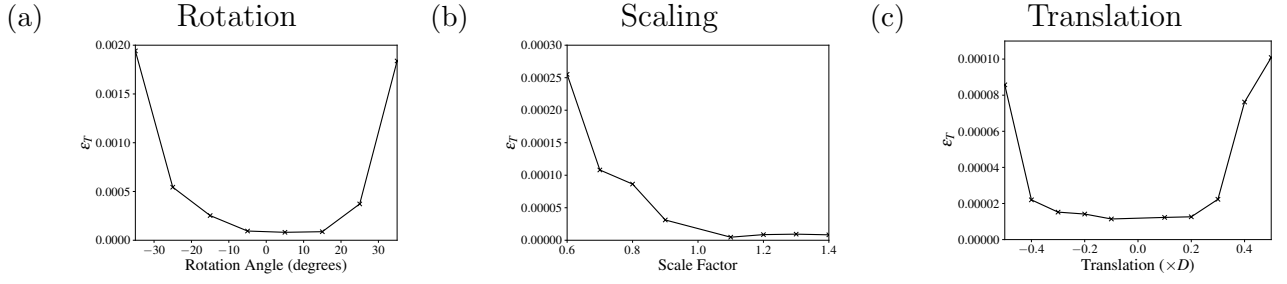


Figure 5.2: Error (ε_T) of the estimated transformation matrix, $\mathbf{T}_{\text{est}}$ with respect to the magnitude of the true transformation. Tested on a) rotation only, b) scaling only, c) translation only.

To evaluate the matrix transformation error, a Frobenius norm was calculated using

$$\varepsilon_T = \frac{\|\mathbf{T}_{\text{est}} - \mathbf{T}_{\text{calc}}\|_F}{\|\mathbf{T}_{\text{calc}}\|_F}. \quad (5.11)$$

5.2.1 Flow reconstruction from measurements with individually applied transformations (rotation, scaling, translation)

Figure 5.2 shows the accuracy of the estimated transformation matrix from data, on which the three transformation types (rotations, scaling and translations) have been applied in isolation, across a range of magnitudes. Irrespective of transformation type, use of the transformation embedded PINN-DA, enables estimation of the transformation matrix to a high degree of accuracy. Considering rotations alone, presented in Figure 5.2(a), across the range of tested angles, the transformation embedded PINN can estimate the applied measurement transformation with maximum error in the order of 10^{-3} . Applied rotations are equivalent to the effect of a (PIV) camera being slightly rotated. Furthermore, for smaller angles ($-20^\circ \leq \theta \leq 20^\circ$), the transformation matrix is estimated with error, one order of magnitude lower (10^{-4}). However, as the applied transformation angle increases and approaches the limits of the tested range $\pm 35^\circ$, the matrix estimation error begins to increase rapidly. This result demonstrates that the transformation embedded PINNs can mitigate the effect of sensor (e.g. camera) rotation from the misaligned measurement data, with error in predicted transformation growing with magnitude of applied transformation (or in a real experimental setup, with increasing misalignment).

Correcting for scaling transformation applied to the measurement data (as if a PIV camera is slightly out of focus), as presented in Figure 5.2(b), shows different behaviour for “stretched” or “squeezed” measurement data. Across the range of applied scaling transformations, the embedded transformation captures the effect to higher accuracy than rotation, suggesting the proposed approach works better for this type of sensor miscalibration. Whether data is stretched (i.e. scale factor > 1.0) or squeezed (i.e. scale factor < 1.0), the prediction error of the transformation matrix increases. However, the rate of error increase is significantly higher for data squeezing than it is for data stretching. Examination of the prediction capacity of transformation embedded PINNs with measurements that are scaled shows that the proposed approach can estimate scaling effects (equivalent to focus error of PIV camera) and that measurement stretching is better predicted than squeezing.

Finally, the effect of sensor misalignment through translation is analysed, as seen in Figure 5.2(c). This corresponds to a PIV plane sitting in a slightly different position to where it is nominally placed, and thus the measurements are slightly offset. Similar to rotations, for translation magnitudes smaller than $0.3D$, the error in estimation of the transformation matrix remains small, thereafter increasing at a faster rate, as the extremities of the tested range are approached.

Examination of Figure 5.2 indicates that different misalignment and miscalibration of sensors can be mitigated for by the transformation embedded PINN-DA approach, but that the accuracy of the learnt transformation, is dependent on the type of transformation. Rotations are captured less accurately than translations and scalings. This can be attributed to rotations having a larger effect on the resultant flow field. Whilst scaling and stretching at most affects two elements in the transformation matrix, $\mathbf{T}$, rotations affect four. Thus the learnt transformation is more complex. In summary, in isolation, the proposed transformation embedded method can combine data assimilation and enforcement of physical conditions (PDE and boundary conditions), to capture the three transformation types (rotations, scalings and translations) to a high degree of accuracy.

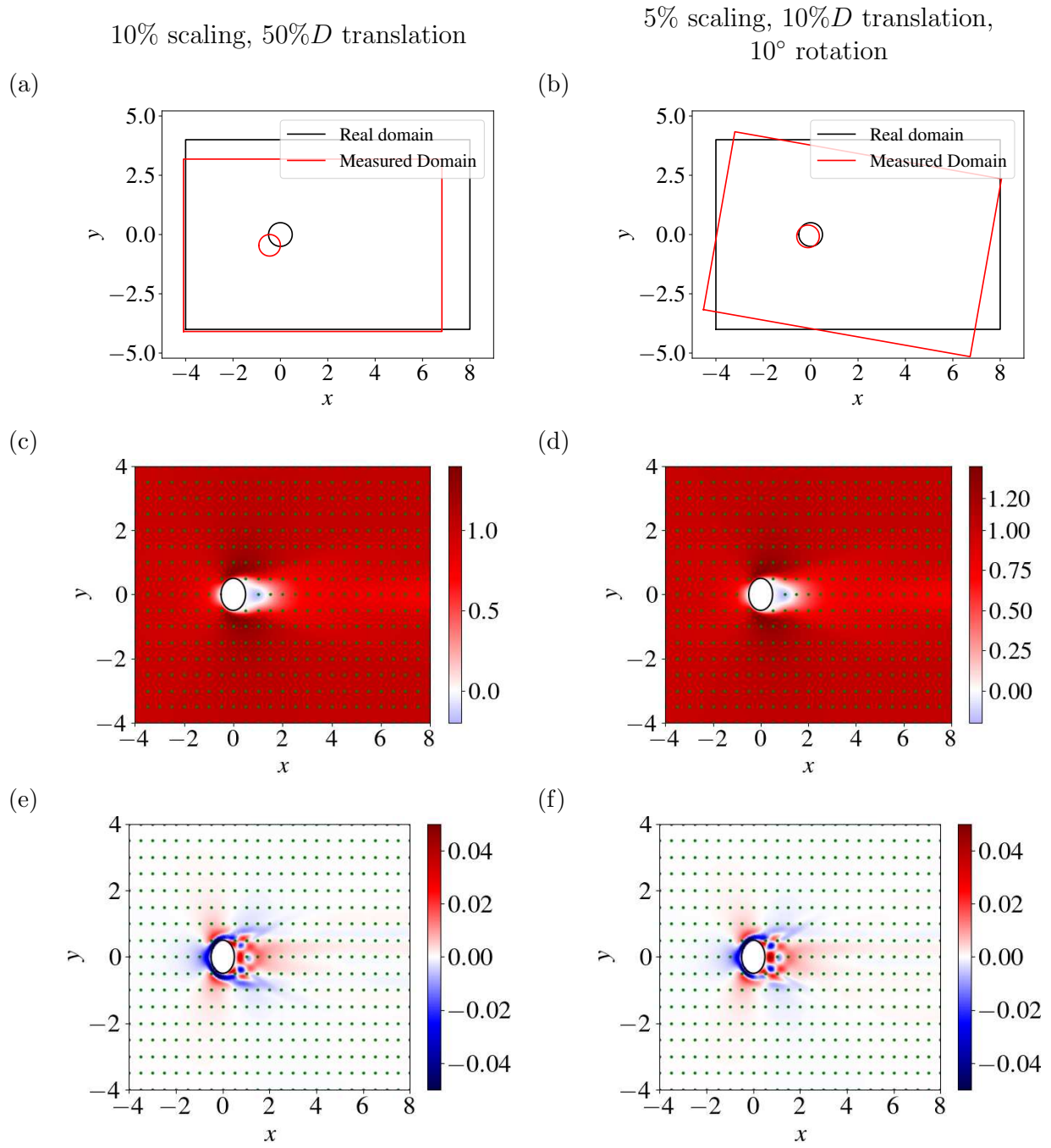


Figure 5.3: Reconstruction of mean laminar circular cylinder flow from sparse misaligned measurements (left) without applied rotation, (right) with applied rotation. (Top) the misaligned measurement domain relative to the real location of measurements, (middle) reconstructed streamwise mean velocity, (bottom) reconstruction error.

5.2.2 Flow reconstruction from measurements with multiple combined transformations applied

In this section, use of the proposed transformation embedded PINN-DAs to mitigate sensor misalignment and miscalibration are explored, investigating how a combination of different transformations can be estimated, through examination of the reconstructed flow fields and the predicted transformation matrix.

Measurements with combined translation and scaling without rotation

Firstly consider the case, where translation and scaling has been applied to the measurement data, but no rotation, as presented in Figure 5.3(a,c,e). Using the transformation embedded PINN-DA, the mean flow was reconstructed from mean velocity measurement data, on which a 50% diameter translation and 10% scaling has been applied, the effect of which is shown in Figure 5.3(a). Here, one can observe where the measurements, $\mathbf{x}$, are actually being taken from (real domain) and where the measurements, $\mathbf{x}_m$, nominally seem to be due to sensor miscalibration (measured domain). In particular, one can see how the “measured” cylinder location is now misaligned with the true location, and the domain has been shortened (which subsequently affects the measured recirculation length). Without mitigating these misalignment effects, enforcement of the measured data would conflict with the boundary condition and governing equation constraints, leading to poor flow reconstruction accuracy.

The reconstructed mean velocity field using the transformation embedded approach, presented in Figure 5.3(c), resembles the true field as the cylinder has been re-centred. The L_2 -based mean velocity reconstruction error is 1.2×10^{-5} compared with 6.6×10^{-6} for reconstructing the flow from accurate measurements (which is presented in Chapter 7). Whilst this represents an increase in reconstruction error, the maximum error is of the same magnitude (≈ 0.05). The main source of this reconstruction error, presented in Figure 5.3(e), appears around the leading edge of the cylinder, which does not occur when reconstructing from accurate, unbiased measurements. This is directly caused by cylinder wall misalignment, in particular due

to translation of measurements, which has not been fully mitigated for. The reconstructed mean velocity from sparse, misaligned measurements shows that the embedded transformation approach is able to correctly apply corrections to the data, but that there is a sensitivity around the cylinder wall, as small errors in the learnt translations result in proportionately larger reconstruction error in the high velocity gradient near wall flow.

The increase in reconstruction error (from misaligned measurements) around the cylinder leading edge can be explained through comparison of the learnt, predicted transformation matrix, $\mathbf{T}_{\text{est}}$ and the true, applied one, $\mathbf{T}_{\text{calc}}$, shown below:

$$\mathbf{T}_{\text{calc}} = \begin{pmatrix} 1.1 & 0.0 & 0.5 \\ 0.0 & 1.1 & 0.5 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{T}_{\text{est}} = \begin{pmatrix} 1.1001 & -0.0003 & 0.4843 \\ 0.0 & 1.0998 & 0.4990 \\ 0 & 0 & 1 \end{pmatrix}.$$

Comparing the predicted and true transformation matrix, one can see that the transformation embedded PINN-DA accurately estimates the applied transformation. The scaling parameter, in particular, has a low prediction error (less than 1%). The translation parameter, however, whilst still estimated well, has a higher prediction error, around 3%. It is this translation error, albeit small, which leads to the high reconstruction error around the cylinder leading edge, due to a slight offset in measured and true cylinder location. This is particularly affected by the high flow gradients in the near wall and thus a small offset has a proportionately larger effect on reconstruction accuracy. Analysis of the learnt transformation matrix shows that the embedded transformation approach is able to accurately mitigate the effects of misaligned data, enabling the accurate reconstruction of the cylinder flow, and that the main reconstruction error can be attributed to the small errors (for translation) in the estimated transformation matrix.

Measurements with combined translation, scaling and rotation

In this section, flow reconstruction from sparse misaligned measurements, on which 5% scaling, a translation of 10% diameter and rotation of 10° has been applied, as presented in Figure 5.3(b,d,f). As before, the comparison of measured and real domain (Figure 5.3(b)) shows the

effects of misalignment, with the coordinate axes and subsequent velocity vectors rotated, in addition to the effects of translation and scaling. The proposed approach has corrected the reconstructed mean velocity field, shown in Figure 5.3(d), with no apparent sign of rotation. For this case, the L_2 -based reconstruction error is 1.1×10^{-5} . Likewise, the reconstructed error field is both qualitatively and quantitatively similar to the previous case, with a similar maximum error magnitude. Furthermore, the high error region in the near wall at the leading edge also occurs here. This result shows that the proposed approach is robust to misaligned measurements with a degree of rotation and that the most sensitive area of reconstruction for the laminar cylinder flow is near the leading edge, due to offsets in the measured cylinder location.

Comparing the true and learnt transformation matrix shows consistent behaviour as the previous case, as the effects of misalignment are well captured:

$$\mathbf{T}_{\text{calc}} = \begin{pmatrix} 1.0340 & -0.1823 & 0.1 \\ 0.1823 & 1.0340 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{T}_{\text{est}} = \begin{pmatrix} 1.0344 & -0.1860 & 0.0856 \\ 0.1821 & 1.0324 & 0.1006 \\ 0 & 0 & 1 \end{pmatrix}.$$

Whilst the learn matrix is accurately predicted, one can observe, as in the previous section, a slightly larger error in translation, specifically in the x direction, which subsequently leads to a rise in reconstruction error at the leading edge, due to the high flow gradients. The learnt transformation matrix shows that, whilst the proposed framework is robust across transformation types, reconstruction accuracy is still sensitive to small measurement corrections errors, when the resultant effect leads to a mismatch in wall boundaries or areas of high flow gradient.

5.2.3 Concluding Remarks

The preceding results show that the proposed transformation embedded PINN-DA framework can be used to successfully capture the effect of measurement misalignment across a range of transformation types, which subsequently enables use of these measurements for flow reconstruction. However, small errors in predicted transformations leads to larger than expected

errors in regions with high flow gradients (such as near wall flow) indicating the potential for further research into the minimum number and optimal placement of sensors.

Finally, one can assume that if an experiment has been well set up, the sensor misalignment and miscalibration will be relatively small. The results in this work show that the method is robust across a larger magnitude of transformations, and therefore can be confidently applied to real experimental (PIV) measurement data.

Chapter 6

Comparing mean-flow reconstruction from sparse mean velocity data using variational data assimilation and turbulence model augmented PINNs

In this chapter, accurate, sparse pointwise mean velocity measurements are combined with the RANS equations (3.5) to reconstruct a turbulent mean flow using two data assimilation methods. Specifically, the gap between data assimilation using PINNs (as described in Section 3.2.1) and variational, adjoint-based methods (described in Section 3.2.2) is bridged by carrying out a direct comparison of both approaches. Furthermore, whilst augmenting the mean-flow equations with a turbulence model have been demonstrated for variational data assimilation (Franceschini et al. 2020), this implementation has yet to be demonstrated with PINNs. Therefore, the following chapter will compare the use of variational data assimilation for turbulent mean-flow reconstruction with PINNs and evaluate turbulence-model augmented PINNs.

As discussed in Chapter 2, whilst PINNs and variational data assimilation have been applied to many similar mean-flow reconstruction problems, direct comparisons between these approaches

are lacking. However, many developments in variational data assimilation, such as the aforementioned turbulence model augmented formulation discussed in Equation (3.9) could also be applied to PINNs. Therefore, a more substantial analysis and comparison between both variational and PINN data assimilation, as completed in this chapter, would bridge the gap between areas, increasing general understanding. To this aim, this chapter has two primary objectives:

- i to investigate whether an existing technique applied in variational data assimilation (turbulence model augmentation) can be successfully extended to the PINN framework and the subsequent challenges involved in implementation, with a primary focus on application to a turbulent flow;
- ii to evaluate and compare the relative performance of both data assimilation methods, when reconstructing the same mean flows in order to build an understanding on the differences between these methods, with regards to theory, implementation and performance.

Although both PINN-DA and variational-DA have the same objective, how the optimisation problem is implemented differs slightly. The use of spatial discretisation and changes in regularisation amongst other factors when using variational-DA, has a quantifiable effect on the flow reconstruction. These results will therefore aim to highlight the effect of these differences. Furthermore, whilst variational-DA requires the use of turbulence model augmentation for convergence with a turbulent flow, PINN-DA can successfully converge both with and without it. This provides an opportunity to isolate the effects of turbulence model augmentation. The turbulence model augmented formulation, discussed in Section 3.1.1, increases the physics constraints on the flow solution by modelling a component of the Reynolds forcing component, which in turn reduces the unclosed portion of the solution and subsequently improves reconstruction accuracy. Understanding how turbulence model augmentation affects mean-flow reconstruction will therefore demonstrate the effect of increasing physics constraints when performing data assimilation.

This chapter is divided into two parts. Section 6.1 contains a direct comparison of the mean-flow reconstruction of the same laminar cylinder flow from sparse, mean velocity measurements,

using variational-based methods and PINNs is completed from the results found in Foures et al. (2014) and Sliwinski & Rigas (2022). This comparison provides a benchmark for the following sections. Sections 6.2-6.4 applies the PINN-DA framework found in Sliwinski & Rigas (2022) to the canonical periodic hill turbulent mean flow at $Re = 5600$ (see 4.2). This PINN-DA approach is then extended by implementation of the turbulence model augmentation technique, proposed by Franceschini et al. (2020) for variational-DA-SA. This is followed by a comparison of the PINN and variational data assimilation.

The sections will be set out as follows. Section 6.1 will contain the aforementioned benchmark comparison of PINN-DA and variational-DA from Sliwinski & Rigas (2022) and Foures et al. (2014). Section 6.2 presents the numerical set up of the problem. This will include a description of the PINN architecture and training procedure, in addition to the numerical setup for the variational data assimilation. The model optimisation is discussed in Section 6.3 followed by the final data assimilation results in Section 6.4. Firstly, the results of the PINN mean-flow reconstruction will be presented, without use of a turbulence model. This will demonstrate the basic capabilities of PINNs and its limitations. Following this, results will be presented, showing the use of turbulence model augmentation (SA) with PINNs. These results will be compared first to the baseline PINN (without turbulence model) and then compared with results from the equivalent variational-based method. Finally, the conclusions are discussed in 6.5.

6.1 Benchmark comparison of mean-flow reconstruction of a laminar cylinder flow from existing research

Both the variational-DA (Foures et al. 2014) and PINN-DA-Baseline (Sliwinski & Rigas 2022) techniques (without the use of turbulence model augmentation) have already been successfully used to reconstruct a laminar circular cylinder mean flow, using the same formulation of the RANS equations. However, a quantitative comparison has not been completed. In this section, a comparison of the results from this existing research is presented, providing a benchmark for the turbulent cases examined in Section 6.4.

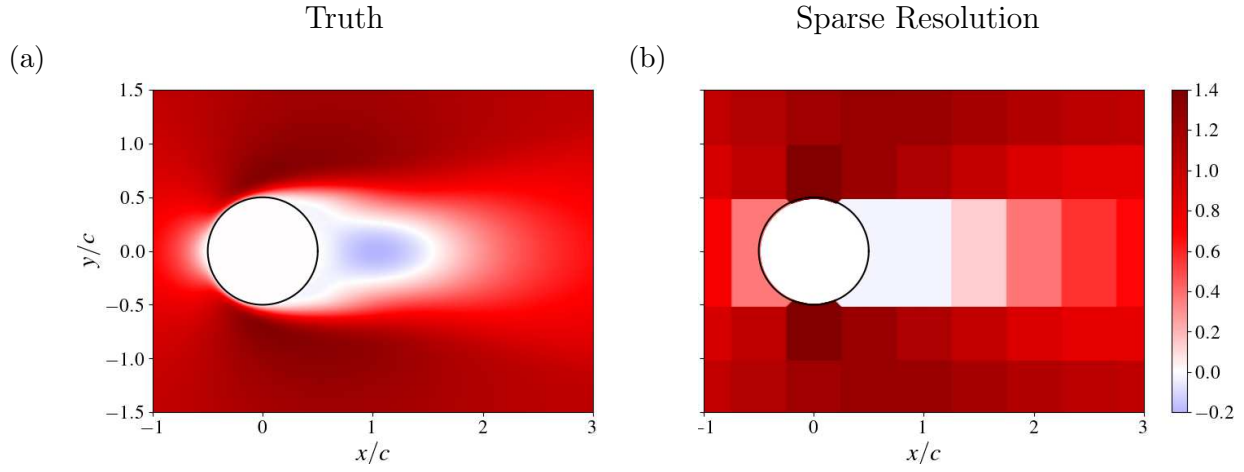


Figure 6.1: High-fidelity DNS data of U (x-component) mean velocity component for laminar circular cylinder at $Re = 150$. (left) Full, true velocity field. (right) Sparse, high-fidelity field sampled at $\Delta L = 0.5$ data resolution used to reconstruct full flow field.

A high-fidelity database was generated using direct numerical simulation of a 2D circular cylinder flow at $Re = 150$. Figure 6.1(a) shows the time averaged x-component velocity. This was then used to extract sparse, mean velocity measurements, simulating a resolution restricted velocity field. The following comparison of PINN-DA-Baseline with variational-DA-Baseline is completed using a data spacing, $\Delta L = 0.5$, on a rectangular grid, as demonstrated in Figure 6.1(b). Further details of the simulation setup and methodologies can be found in Foures et al. (2014) and Sliwinski & Rigas (2022) for the variational and PINN case respectively. The aforementioned papers also contain results for different data resolutions.

Figure 6.2 shows the reconstructed U field (a,b) and the prediction error (c,d). This highlights several trends, which are consistent across all velocity components. Firstly, the reconstructed mean velocity fields are similar to the true fields and accurately capture the expected physical features, such as the symmetrical wake and the stagnation at the front of the cylinder. Additionally, a key pattern emerged from this result. Both the variational-DA and PINN-DA approaches have almost identical reconstruction error field topology. These occur between datapoints, in the wake and recirculation regions and near the wall. Finally, as one moves away from the cylinder (in y), the error reduces as the flow becomes more uniform.

Whilst the reconstruction error fields are comparable, it is clear that the absolute error for the PINN-DA-Baseline is much smaller in magnitude. The variational-DA has a maximum

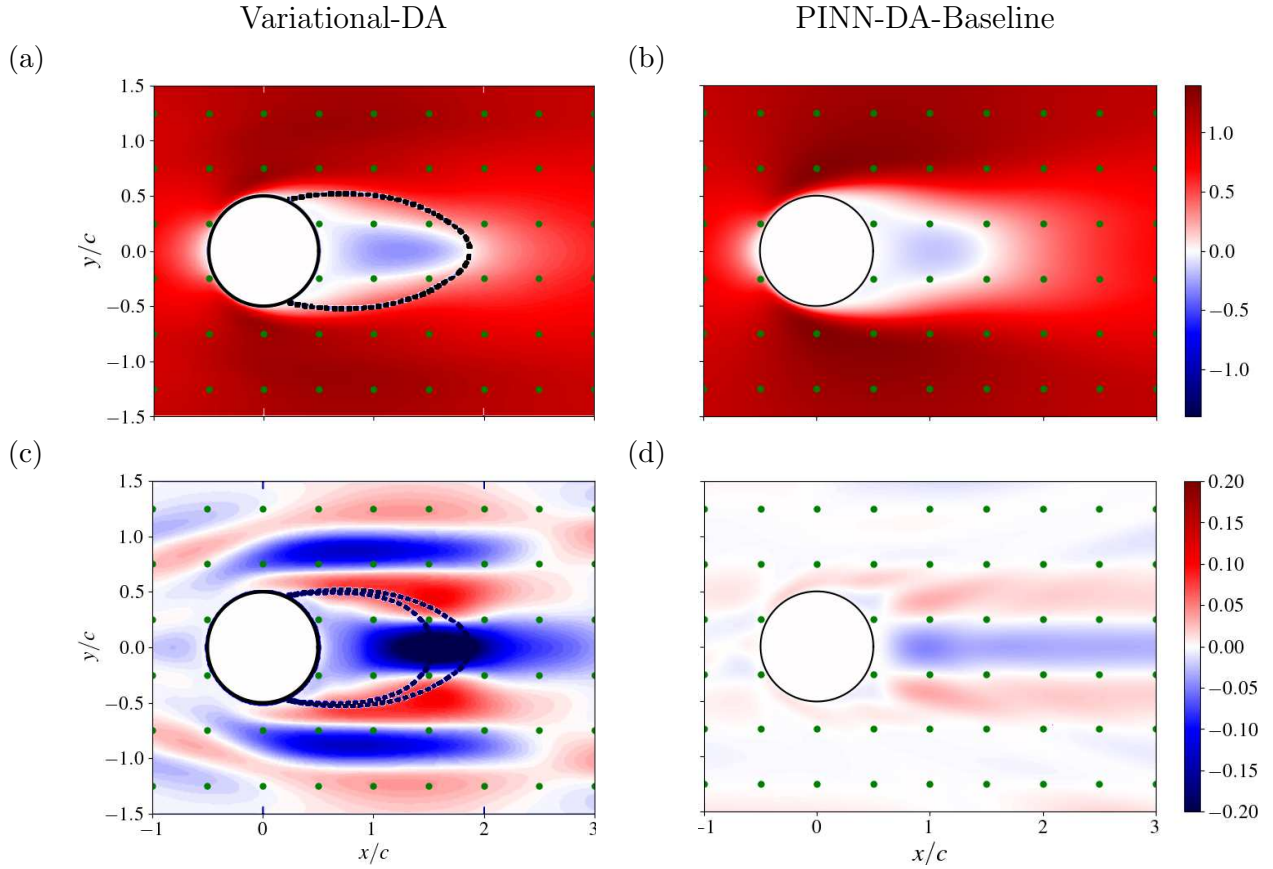


Figure 6.2: Comparison of the U (x-component) mean velocity component reconstructions for $d = 0.5$ data resolution with DNS (top). The variational-DA approach (left) and the PINN-DA-Baseline (right) figures are taken from (Foures et al. 2014) and (Sliwinski & Rigas 2022), respectively. (a,b) - U prediction. (c,d) - Absolute U error.

error magnitude of 0.3 (for absolute U error), whereas for PINN-DA it is around a third of this at 0.11. This comparison is important as it suggests, all else equal, for a laminar cylinder flow, the PINN-DA produces better results. This is most apparent when comparing the length of the wake. Whilst the variational approach predicts the wake recirculation region closes at $x = 1.85$, the PINN approach closes at $x = 1.55$ (compared with $x = 1.55$ for the DNS flow field). Furthermore, the decay in error, observed as one moves away from the body, is much more apparent in the PINN-DA results. This may indicate the fundamental nature of PINNs and their construction, allows them to better generalise these simple uniform regions.

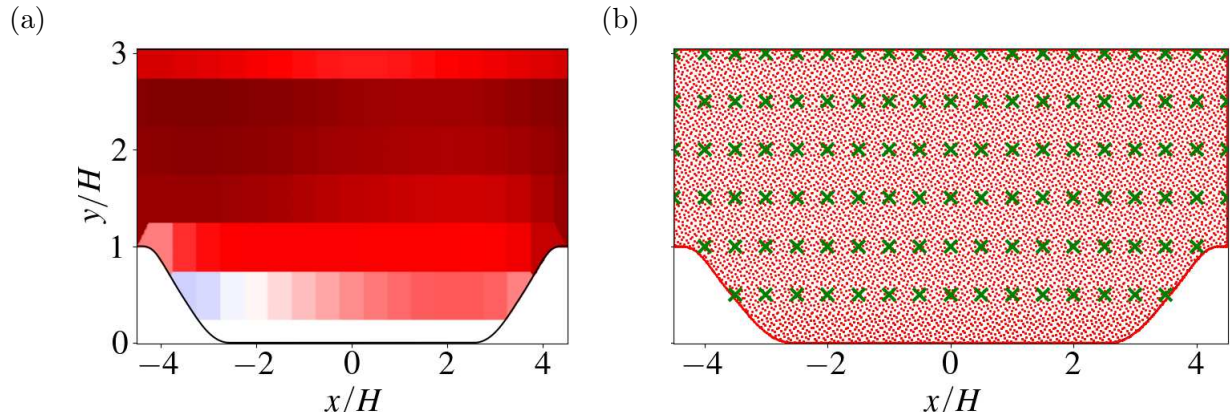


Figure 6.3: Turbulent periodic hill domain used for data assimilation, with $\Delta L = 0.5$ data resolution used to compare the variational and PINN approaches. (Left) The sparse velocity field at selected data resolution. (Right) The collocation points, where the PINN evaluates residual error (red) sampled across the domain using a Hammersley distribution and measurement locations (green crosses).

6.1.1 Final Remarks on benchmark laminar comparison

This direct comparison of both variational-DA and PINN-DA approaches can be used to benchmark the turbulent case presented in the following section. This comparison demonstrates that for laminar unsteady flow, the PINN-DA approach flow reconstruction can match and improve the variational-DA reconstructions and indicates that when applied to a turbulent flow, PINN-DA and variational-DA should be comparable. However, use of PINNs does offer certain advantages. Firstly, the simplicity of PINN code ensures one can quickly reconfigure the model. This includes changing the formulation of the PINN-DA-Baseline from the Helmholtz decomposition from Eq. (3.7) to the Reynolds stress formulation (Eq. (3.5)) with a few lines of code. By using this second formulation and additionally providing second order statistics (Reynolds stress) data, Sliwinski & Rigas (2022) were able to demonstrate how PINNs could be used to predict the pressure field (this will be discussed in greater detail in Section 6.4.2. This was not demonstrated in the variational approach in Foures et al. (2014) and highlights the flexibility offered by PINNs.

6.2 Numerical and Model Setup for turbulent case

The PINN and variational approaches described above are applied to a canonical turbulent test case, the flow over a periodic hill (described in 4.2). To analyse the effect of data resolution on mean-flow reconstruction accuracy, the high-fidelity DNS data is spatially sampled on a square grid at nine data resolutions - $0.05 \leq \Delta L \leq 1.0$, where $\Delta L = \Delta x/H = \Delta y/H$. All results and figures hereon-in will compare the results from the $\Delta L = 0.5$ resolution, unless otherwise stated. This resolution can be visualised in Figure 6.3(a) with exact measurement locations also shown in Figure 6.3(b). The choice of a sparse square sampling grid represents a velocity field as measured using methods such as Laser Doppler Velocimetry or PIV (although PIV actually takes a spatial average of velocity over the pixel, not a pointwise measurement as done here).

The rest of this section will describe the details of the numerical setup of the two methods employed.

6.2.1 PINN set-up

The PINNs are implemented using the DeepXDE Python library by Lu et al. (2021), as in Sliwinski & Rigas (2022). This toolbox simplifies the process of building PINNs (with the Raissi et al. (2019) paradigm). The code and data to run the PINN cases from this chapter are available at https://github.com/RigasLab/PINN_SA.

PINN architecture and hyper-parameter selection

All PINNs use the same architecture, as presented in Figure 6.4. The neural network was constructed as a multilayer perceptron with 7 fully connected hidden layers, each with 50 nodes. The input layer consists of two nodes, representing spatial coordinates, x/H and y/H . The output layer consists of varying number of nodes depending on the test case. The PINN initially contains five output nodes with two mean velocity components, the combined mean pressure and potential forcing field, $P - \phi$, and two solenoidal forcing nodes. When the SA turbulence model

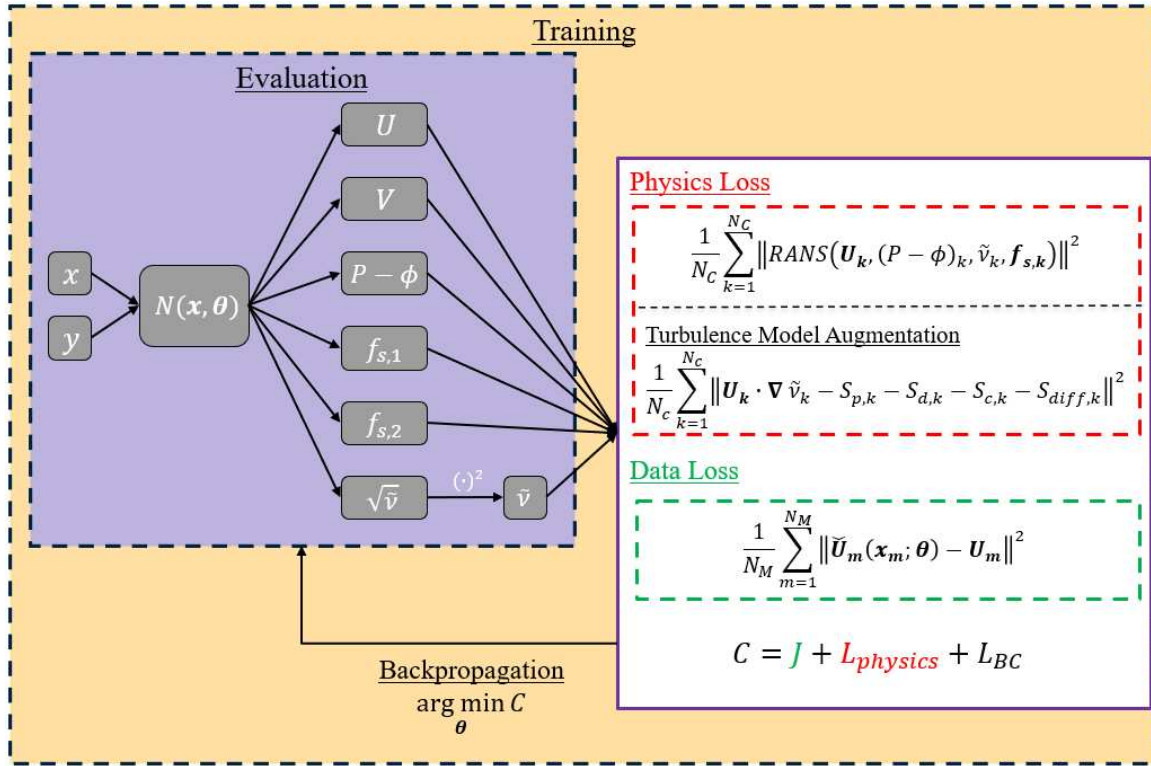


Figure 6.4: Structure of the Physics-Informed Neural Network. A continuous mapping between spatial coordinates and flow variables is approximated by a deep neural network, denoted by N , defined by a set of weights and biases, θ . High-fidelity coarse data measurements and the (turbulence model augmented) RANS equations constrain the training of the network. k, m subscript the physics and data loss, respectively. i represents the i -th component of a vector.

augmentation was used, an additional output node is required for $\tilde{v}$. Each node used a tanh activation function for its smoothness and its second order differentiability. This architecture was deemed to perform well across a range of data resolutions. A grid search method was employed to find optimal hyper-parameters (number of nodes and layers, activation function and learning rate). The effect of hyper-parameters on reconstruction accuracy are included in Appendix B.1. For PINNs, the weights and biases defining the network are randomly initialised using the Glorot uniform algorithm. This contrasts the variational-DA-SA approach, in which an “initial guess” (the RANS-SA solution) was used, around which the system is initially linearised.

For PINN optimisation, collocation points must be specified, in order to define where the governing laws are evaluated. In this work, the PINNs use 10000 collocations points which are distributed using a Hammersely distribution, such that there are sufficient points across

all regions of the domain. This number of collocation points for the PINNs was selected after the reconstruction performance showed little sensitivity to further increases of the number of points. An additional 1000 points were distributed to the walls and periodic boundaries to enforce the boundary conditions. The collocation points (and domain) are shown in Figure 6.3(b).

As in Sliwinski & Rigas (2022), a two step optimisation was used for the PINNs. Initially an ADAM optimisation is applied, followed by an L-BFGS-B phase. L-BFGS-B is a quasi-Newton method and thus its convergence depends on the initial guess. By using an ADAM optimiser first, one can reach an acceptable loss level to provide a good initial guess for the L-BFGS-B optimiser. Details on PINN convergence are discussed further in Section 6.3. The PINNs were trained on one NVIDIA RTX 6000 GPU.

Modification to PINNs for Spalart-Allmaras augmentation

Using PINNs in combination with the SA turbulence model introduces additional complexity, which require modifications to the neural network architecture and the problem formulation to achieve convergence.

Firstly, the SA transport equation contains $\frac{1}{d^2}$ terms in the production and destruction terms, where d is the true wall distance, resulting in singularities at the wall, where $d = 0$. This is not a problem for the finite element method (RANS and variational) where the residual is evaluated at a cell centre, which is not on the wall. For the PINN paradigm used, however, the governing equations are evaluated exactly at the wall boundaries and so this leads to singularities and subsequently failure to convergence. Consequently, the equation was reformulated by multiplying the SA equation with d^2 to eliminate the singularity. This new formulation can be found in Appendix A.1.2. Problems arising from singularities when used with PINNs have also been mentioned in Molnar et al. (2023).

The second change is to eliminate the effect of clipping negative values of $\tilde{\nu}$. The negative SA equation formulation used in the variational approach has distinct equations for positive and

negative $\tilde{\nu}$ (detailed in Appendix A.1.1) to ensure the transport equation is continuous across both positive and negative $\tilde{\nu}$. The final eddy viscosity, ν_t , is then clipped at negative values of $\tilde{\nu}$ and set to zero, resulting in only positive values of ν_t . Convergence of the PINNs is improved when the complexity caused by the discontinuity from clipping ν_t is removed. Thus, to enforce positive $\tilde{\nu}$, the output of the neural network is defined as $\sqrt{\tilde{\nu}}$. A transformation is then applied to the output, squaring the value, to produce a positive $\tilde{\nu}$. This is shown in Figure 3.1.

The direct effect of these modification on PDE enforcement and loss convergence is discussed in Appendix B.2.

6.2.2 Variational assimilation set-up

The solutions of the RANS-SA system (3.9) and its adjoint (3.34) are obtained through a finite-element method (FEM) spatial discretisation implemented in FreeFEM++ (Hecht 2012). Contrary to the presentation in Section 3.2.2, the adjoint model is obtained following a discrete adjoint approach, where the Jacobian matrix that is associated to the RANS-SA equations is transposed. This Jacobian matrix is also used to invert the nonlinear RANS-SA equations based on a Newton method. Piecewise-linear functions that are enriched by bubble functions are used for velocity and pseudo-turbulent viscosity variables, while piecewise-linear functions are used for pressure. The FEM is known to be unstable at high Reynolds numbers. Accordingly, both streamline-upwind Petrov-Galerkin (SUPG) (Brooks & Hughes 1982, Franceschini et al. 2020) and grad-div (Olshanskii et al. 2009) formulations are here employed to stabilise the method. A low-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) (Nocedal 1980) algorithm is used to exploit the gradient $\frac{dJ^R}{d\mathbf{f}_s}$ from Eq. (3.35) in order to perform the minimisation of the cost function J in Eq. (3.1). As mentioned above, the first-guess for the optimisation procedure is $\mathbf{f}_s = \mathbf{0}$, i.e. the baseline solution of the RANS-SA equations. The full optimisation procedure was carried out on 28 cores Intel Xeon Broadwell E5-2680v4, 2.4 GHz. The total number of degrees of freedom in the RANS and adjoint simulations is around 2×10^5 . Further details on the implementation of the variational-DA-SA can be found in (Franceschini et al. 2020, Mons et al. 2022).

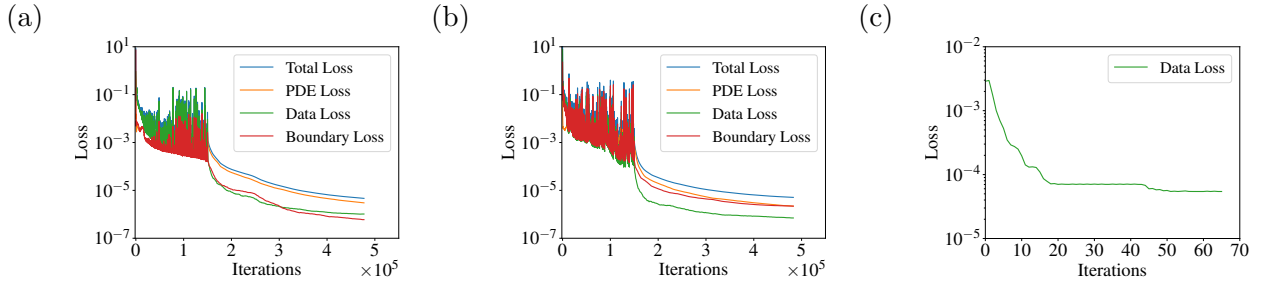


Figure 6.5: Residual loss during optimisation for (a) PINN-DA-Baseline, (b) PINN-DA-SA and (c) Variational-DA-SA. For PINN optimisation (training), there are two distinct phases: the ADAM (stochastic gradient descent based) optimisation for 150000 iterations followed by the L-BFGS-B phase.

6.3 Model Optimisation

Figure 6.5 shows the optimisation convergence of the PINN-DAs (a,b) and the variational-DA-SA approach (c). The PINN-DA loss convergence curve is decomposed into three loss components - PDE residual, data error and boundary condition error. For PINN-DA-SA the additional loss from the SA transport equation contributes also to the PDE loss. Each individual loss component appearing in Eq. (3.31) was given initial weight values λ based on the relative magnitudes of the flow variables. Thereafter, the weights were tuned using a grid-search approach to find the combination with the lowest loss, as compiled in Table 6.1. The weight sensitivities are presented in Appendix B.1. Similarly, the effect of regularisation and importance of λ^R is demonstrated in Appendix B.3.1. During the first phase of the training, ADAM optimisation is performed for 150000 iterations. The second phase applies L-BFGS-B optimisation for 300000 iterations or until the model has sufficiently converged to the floating point tolerance. Figures 6.5(a,b) shows that during optimisation (for both PINN-DA methodologies), the ADAM phase reduces loss from the high value caused by the initial randomised state to approximately 10^{-3} . During the ADAM optimisation phase, one can observe frequent jumps in the loss. This is an effect of the adaptive step size used in ADAM. The step size

Table 6.1: Loss function weights used in PINN optimisation. For PINN-DA-Baseline, the $\tilde{\nu}$ and λ^R weights can be ignored. Additionally, this regularisation weight is only valid for the $\Delta L = 0.5$ data resolution. For other resolutions, see Figure B.4.

Weight	λ^P	λ^D	$\lambda^B(U_i) - \text{Wall}$	$\lambda^B(\tilde{\nu}, f_{s,i}) - \text{Wall}$	$\lambda^B - \text{Periodic}$	λ^R
Value	1	10	2.5	10	1	0.01

is inversely proportional to the exponential average of squared (loss) gradient (as described in Kochenderfer & Wheeler (2019)). As the loss flattens out, the adaptive step size becomes large, causing the model to “jump” too far to a higher loss level. The L-BFGS-B phase is then used to fine tune the optimisation, reducing total loss to below 10^{-5} .

The convergence of the variational-DA-SA, seen in Figure 6.5(c) shows the data loss, corresponding to J/N_m , where J is defined in (3.1). This data loss is equivalent to the corresponding curves in Figures 6.5(a,b). Whilst convergent, the data loss is an order of magnitude higher for the variational-DA-SA compared with the PINN-DAs. Whilst there is a vast difference in number of iterations between the PINN and variational approaches, an iteration is not equivalent across both methods. For PINNs, a single iteration involves evaluating the PINN at all collocation and data points, calculating the cost function and back-propagating the loss to update the model weights (and thus indirectly the flow solution). For the variational approach, in each iteration the direct and adjoint equations are solved after which the forcing solution is directly updated. An average PINN iteration takes approximately 0.84 s/iteration, whilst for the variational approach this is approximately 55 s/iteration. This difference in computational cost can be attributed to computational complexity, more mesh points in the variational method (than PINN collocation points), reduced number of degrees of freedom in the neural network (updating only weights and biases instead of the actual solution across the domain) and hardware.

6.4 Results

The mean-flow reconstruction from sparse mean velocity measurements for the turbulent periodic hill case using three approaches are compared: PINN-DA-Baseline (without turbulence model), PINN-DA-SA (with SA) and variational-DA-SA (with SA). Firstly, an evaluation of the PINN-DA-Baseline approach is presented. Secondly, it is shown that significant reduction in reconstruction error is achieved by augmenting the PINN-DA-Baseline approach with the SA turbulence model (PINN-DA-SA). Thirdly, the variational-DA-SA is directly compared to

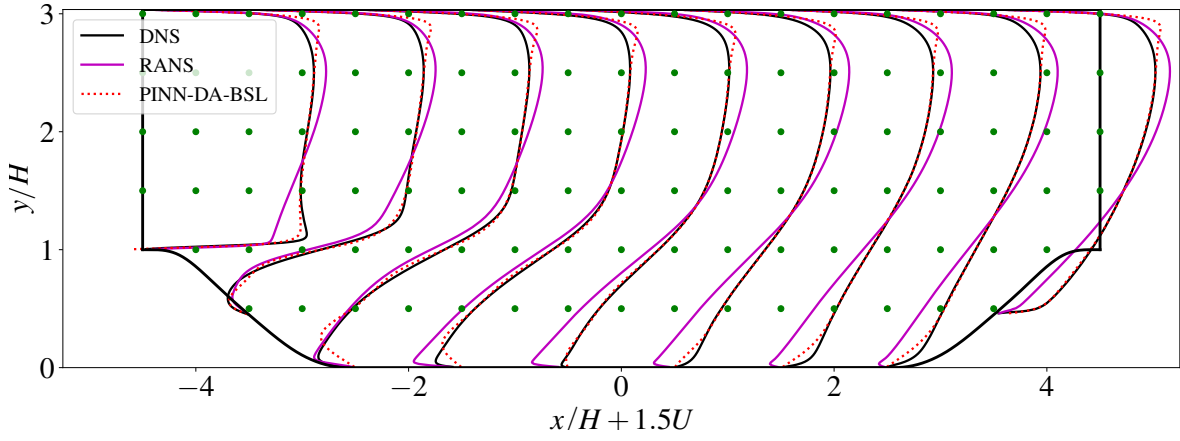


Figure 6.6: Streamwise velocity profiles for DNS, RANS-SA and PINN-DA-Baseline.

PINN-DA-SA showing the increased accuracy of the PINN-DA-SA relative to the variational-DA-SA.

To evaluate the mean-flow reconstruction accuracy, a volume weighted L_2 error will be used and is calculated as

$$\varepsilon_2 = \sqrt{\frac{1}{\Omega} \sum_{p=1}^{N_P} \delta\Omega_p \|\mathbf{U}_p - \hat{\mathbf{U}}(\mathbf{x}_p)\|^2}, \quad (6.1)$$

where Ω is the volume of the domain, $\delta\Omega_p$ is the cell volume of each point used to determine the error (as defined by the DNS grid) and N_p is the number of points across the domain. $\mathbf{U}_p$ is the high-fidelity, DNS mean velocity at $\mathbf{x}_p$ and $\hat{\mathbf{U}}(\mathbf{x}_p)$ is the corresponding reconstructed mean velocity. The variational solution is interpolated onto the DNS grid using a linear interpolation method. Additionally, the employed finite-element mesh corresponds to typical mesh sizes that are close to those in DNS.

6.4.1 PINN-DA-Baseline

Here, the PINN-DA-Baseline approach is applied to reconstruct the mean-flow velocity field when coarse mean-flow velocity data is available with spacing $\Delta L = 0.5$. This is compared against the DNS and RANS-SA velocity profiles in Figure 6.6. The data assimilation of the high-fidelity measurements has improved the prediction of the velocity field compared to RANS-SA. The total mean velocity L_2 error is 3.60×10^{-2} instead of the 8.99×10^{-2} error from RANS-SA. High accuracy of reconstruction is observed in the bulk flow domain. However, high

mean velocity reconstruction error is observed in areas with high velocity gradients, such as near the walls and at separation. This is most apparent at the upper wall boundary layer. This near wall error accounts for over 40% of the total L_2 error as the PINN-DA-Baseline fails to accurately capture high-gradient flow features. These trends are highlighted in the streamwise velocity contours in Figure 6.8(a,d). At data points, the PINN-DA-Baseline reproduces the measurements accurately. Furthermore, in spite of the higher near-wall reconstruction errors, Figure 6.8(a) shows improved reconstruction of the recirculation bubble, even with limited data in this area. This is most noticeable when comparing the PINN-DA-Baseline bubble from Figure 6.8(a) with the RANS bubble in Figure 4.5(b). The provision of sparse high-fidelity datapoints in the PINN-DA-Baseline approach has enabled better prediction of the shape of the recirculation region. However, the error field makes the limitation of the PINN-DA-Baseline approach more apparent, with the near wall errors particularly noticeable.

6.4.2 Reconstructing pressure with PINN-DA-Baseline

Instead of applying the Helmholtz decomposition (3.7) with the PINN-DA-Baseline, as in Section 6.4.1, one can alternatively enforce the base RANS equations as presented in (3.5), using both high-fidelity first order statistics (mean flow) and second order statistics (Reynolds stress) $(U_i, \overline{u'_i u'_j})$ at discrete locations to reconstruct the flow field. This approach enables reconstruction of both mean velocity and Reynolds stresses fields, but also allows the inference of the pressure field.

Using the data spacing $\Delta L = 0.5$, shows that this alternate PINN-Baseline approach produces comparable mean velocity reconstruction as the PINN-DA-Baseline, as seen in Figure 6.7(a). The L_2 error is 3.37×10^{-2} for this second approach compared with 3.60×10^{-2} for the former approach. Likewise the key trends discussed in Section 6.4.1 all still hold. However, this

Table 6.2: Weights used in PINN optimisation when both mean velocity and Reynolds stress data are provided

Weight	λ^P	$\lambda^D(U_i)$	$\lambda^D(\overline{u'_i u'_j})$	$\lambda^B(U_i) - \text{Wall}$	$\lambda^B(\overline{u'_i u'_j}) - \text{Wall}$	$\lambda^B - \text{Periodic}$
Value	1	10	100	2.5	10	1

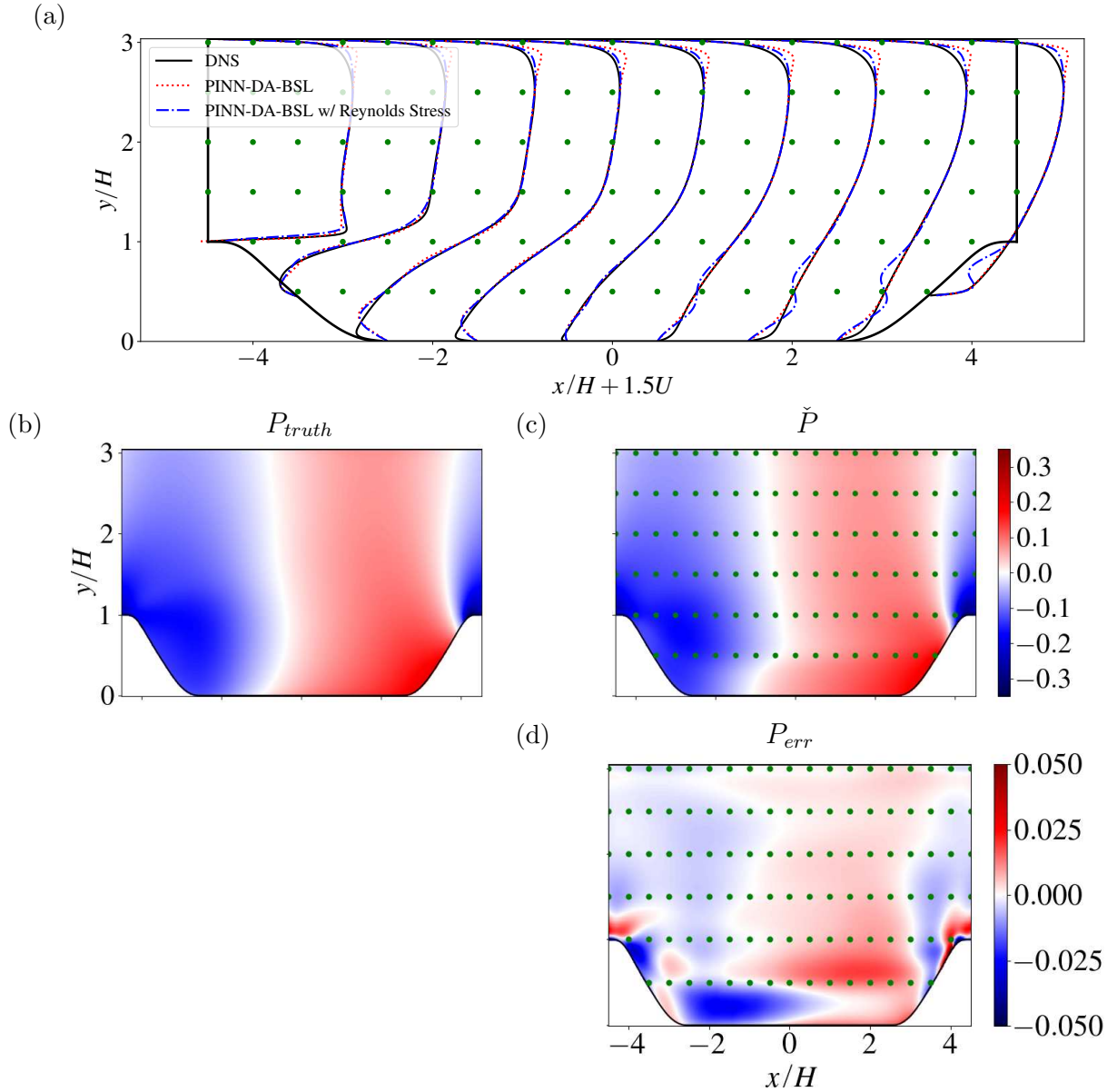


Figure 6.7: PINN-DA-Baseline mean-flow reconstruction. (a) Comparison between the DNS, PINN-DA-Baseline and PINN-DA-Baseline (with Reynolds stress data) velocity profiles for the U component. (b) True pressure field from DNS. (c) Pressure field inferred from the PINN-DA-Baseline approach from sparse mean velocity and Reynolds stress data. (d) Absolute mean-pressure error from PINN-DA-Baseline pressure reconstruction.

alternate formulation has more output variables and fewer equations when to the prior approach used in the main paper.

Although both approaches achieve similar reconstruction error, this second approach requires second order statistics (Reynolds stress measurements) along with first order ones (mean flow field). However, this formulation can be used to reconstruct the pressure field, as shown before for laminar cases (Sliwinski & Rigas 2022). This is demonstrated in the turbulent case, as seen

in Figure 6.7(b-d). This has a high potential for practical applications in order to obtain the pressure field without performing intrusive measurements. Examination of the RANS equations (3.5) shows how provision of mean velocity and Reynolds stress data at discrete points defines all terms in the equations except for the pressure term. At these measurement locations, the equations are closed and the resultant neural network can infer a pressure field, both at data points but also away from them as well. This resultant pressure field has highest reconstruction error in the regions with highest pressure gradients in the flow, matching the trend seen with mean velocity error. The pressure reconstruction error is less correlated with the location of data points, than the mean velocity. It should be highlighted that the neural network pressure output must also include an integration constant (as the RANS equations only contain derivatives of pressure). To determine this constant and recover the actual, unshifted pressure field, one needs to fix the pressure at one point. This can be done during training (by providing pressure data at a single point) or as in this case, the pressure field is shifted in post processing. Taking the pressure data at a singular point, one can find a constant, such that by adding it to the model pressure at that point, the predicted pressure exactly matches the measurement. This constant can then be applied to the full solution.

6.4.3 PINN-DA-SA

This section contains results from the augmentation of the SA model to PINNs. First, a comparison between the PINN-DA-Baseline and PINN-DA-SA results will be presented, followed by a discussion on the improved reconstruction when the SA model is used. Finally, the results of the variational data assimilation using SA (variational-DA-SA) are also presented, followed by a parametric analysis of the data resolution.

PINN-DA-SA vs PINN-DA-Baseline

Figure 6.8(a,d) shows the PINN-DA-Baseline and Figure 6.8(b,e) the PINN-DA-SA reconstruction results. Streamwise velocity and error contours relative to the DNS solution are shown,

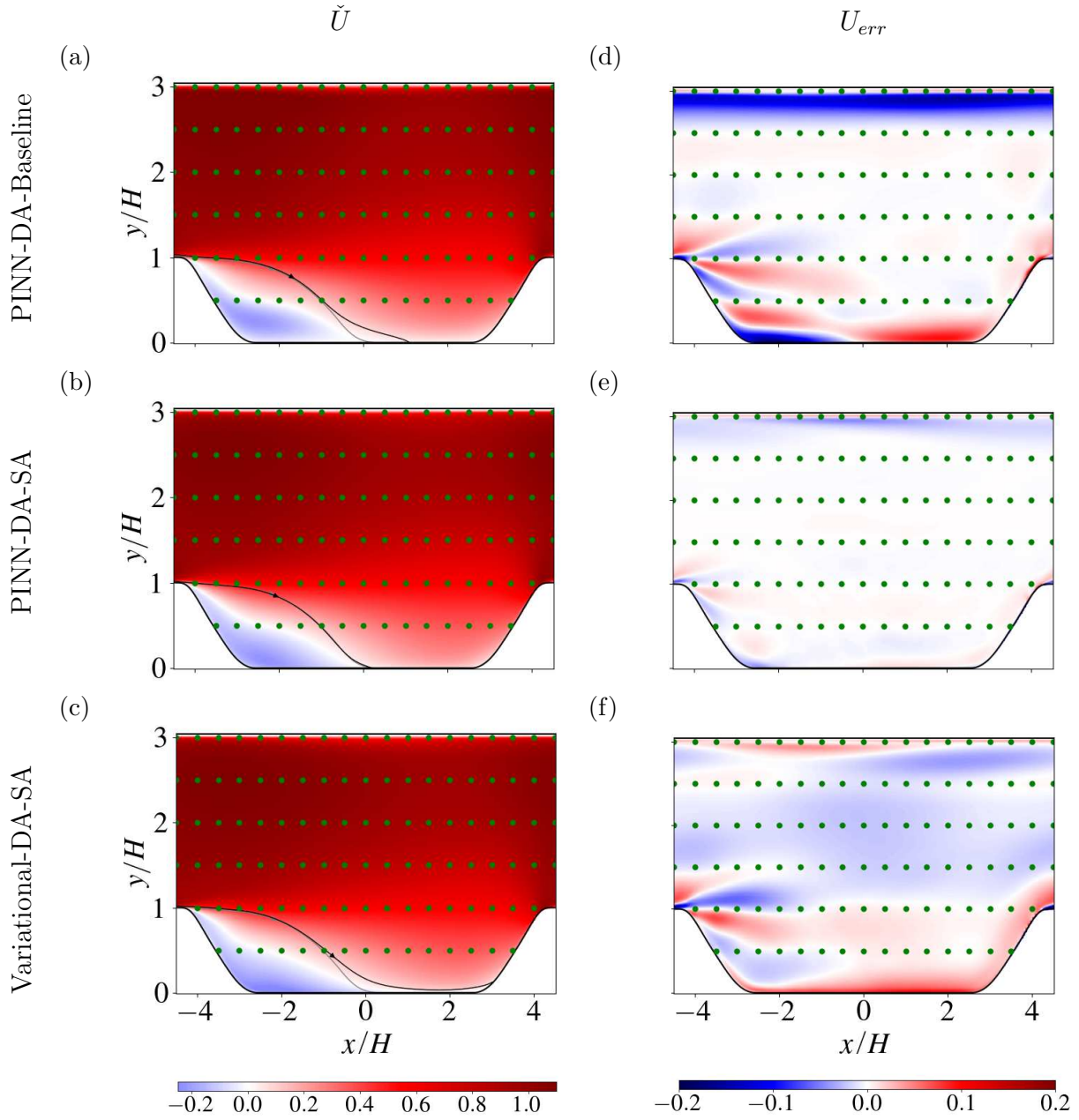


Figure 6.8: Reconstructed U velocity component for the PINN-DA-Baseline (top), PINN-DA-SA (middle), variational-DA-SA (bottom) approaches. (a,b,c) - U reconstruction. (d,e,f) - Absolute U error. PINN-DA-SA reconstructs the flow most accurately, significantly reducing the separation error seen in other approaches, whilst also reducing the near wall error seen in PINN-DA-Baseline. The dividing streamline (separating the recirculation region) is shown in (a,b,c) and compared with the DNS equivalent (in grey). Whilst the PINN-DA-SA recirculation region is indistinguishable, there is a small but noticeable difference for the PINN-DA-Baseline and variational-DA-SA solutions.

with similar trends holding for the vertical velocity component (not shown here). Whilst the error field topologies are qualitatively similar, the magnitude error is significantly reduced. For PINN-DA-SA, the recirculation bubble is almost indistinguishable from the DNS one. As a

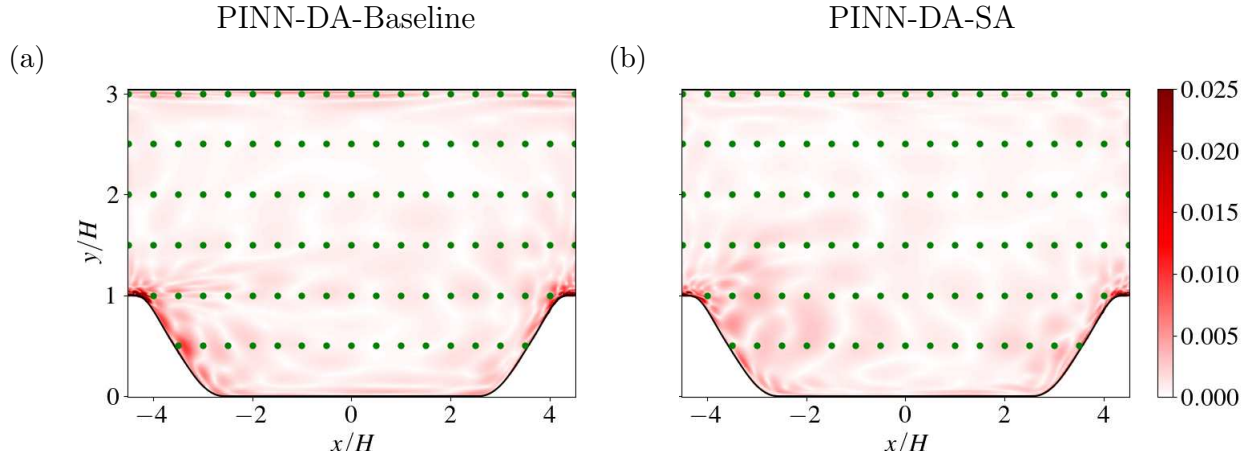


Figure 6.9: PDE residual error (x-momentum and y-momentum). (a) PINN-DA-Baseline and (b) PINN-DA-SA. The residual error between approaches are comparable despite the considerable improvements in mean velocity reconstruction error demonstrated by the PINN-DA-SA.

consequence of introducing the Spalart-Allmaras model, reconstruction error has reduced by 63% ($\varepsilon_2 = 1.35 \times 10^{-2}$). For the bulk flow region above the hills and away from the walls (between $y/H = 1.5$ and $y/H = 2.5$) and at datapoints, the PINN-DA-SA and PINN-DA-Baseline show similarly low reconstruction errors. This is expected in this region given the lack of mean shear and production of turbulent kinetic energy.

Comparing the PINN-DA-Baseline error against the PINN-DA-SA highlights the effect of augmenting the PINN with the SA turbulence model. The high mean-flow error along the upper wall has largely reduced indicating that use of the SA model has enabled a better reconstruction of the near-wall region. There are also similarly significant improvements along the lower wall both at separation but also after reattachment. The improvements of PINN-DA-SA compared to PINN-DA-Baseline corroborate well with the original design purpose and performance of the SA turbulence model, which shows improved performance for turbulent boundary layers in adverse pressure gradients.

Why PINN-DA-SA?

In order to understand the difference in performance between PINN-DA-Baseline and PINN-DA-SA, one can examine the individual terms of the cost function, C , for each approach. This consists of the measurement error term (data loss) and the enforcement of the governing

equations (PDE loss).

Firstly, one can examine the data loss component. Inspection of Figure 6.8(d) reveals that the PINN-DA-Baseline reconstructs the mean velocity measurements accurately. This is comparable to the performance of the PINN-DA-SA, with the convergence plots in Figure 6.5(a,b) showing similar converged data loss values. Away from the data points, the reconstruction error increases substantially for the PINN-DA-Baseline. This is most apparent along the upper wall where a high error region exists above the data at $y/H = 2.5$. Given that the reconstruction error is low at the data points for PINN-DA-Baseline (and also similar to PINN-DA-SA), further decrease of the error to measurements, J , does not reduce the mean velocity error away from datapoints, where most of the reconstruction error lies.

Secondly, one can consider the other component of the cost function, the residual PDE error. Figure 6.9(a) shows the PINN-DA-Baseline momentum residual error (for x- and y-momentum equations). The PDE residual error contour shows that there is a limited correlation between mean velocity reconstruction error and residual error from the conservation laws - most apparent along the upper and lower walls. A similar analysis for the PINN-DA-SA corroborates this trend (Figure 6.9(b)). The total residual PDE loss from both PINN-DA-Baseline and PINN-DA-SA (seen in Figure 6.5) are of similar magnitudes (approx. 10^{-5}) and the residual PDE fields are comparable, despite the significant differences in reconstruction accuracy between the two PINN-DA approaches. This indicates that further reduction in the residual PDE error does not lead to increase in reduction in reconstruction error and thus cannot be attributed as the cause of the difference between PINN-DA-Baseline and PINN-DA-SA.

The difference between PINN-DA-Baseline and PINN-DA-SA can be consequently attributed to the effect of the turbulence model augmentation. Due to the unclosed (underdetermined) nature of the RANS equations, there are infinite solutions to the forcing (divergence of Reynolds stress tensor) in the mean-flow equations. Providing high-fidelity data measurements “anchors” the PINN-DA solution at these discrete data points. However, away from them, there are still many candidate solutions to the mean flow, as the governing equations are still under-determined. These solutions all satisfy the governing equations and thus all have low residual PDE error.

There are consequently no terms within the optimised cost function (Eq. (3.31)) that can be minimised further, to reduce the mean velocity reconstruction error across the domain for PINN-DA-Baseline. The current optimisation for PINN-DA-Baseline is converged towards a solution that obeys the under-determined RANS equations. This is also briefly mentioned in Foures et al. (2014). The SA model adds physics constraints that limit further the infinite number of admissible solutions far from the measurement points and as a direct consequence lead to improved reconstruction.

Variational-DA-SA results and comparison with PINNs

The PINN-DA-SA can be directly compared to the variational-DA-SA, which also uses SA augmentation, since both methods use the same data and physics constraints. Figures 6.8(c,f) show the reconstructed mean streamwise velocity and the respective absolute error for the variational-DA-SA. As with all the data assimilation results presented, the key flow features (separation, recirculation bubble and reattachment) are well reconstructed. As with both PINN-DA methodologies, the variational-DA-SA reconstructs the shape of the recirculation bubble more accurately than the RANS-SA, despite the limited measurements in this region. The reconstruction error matches the features seen in the PINN-DA results as well as in (Foures et al. 2014), with high error between data points, near walls and at the initial separation.

The key feature seen in the comparison between variational-DA-SA and PINN-DA-SA is that the latter has demonstrably lower error across all regions of flow. The highest reduction in error in the PINN-DA-SA is observed in the region between hills ($y/H \leq 1.0$), specifically the recirculation and separation. Whilst all approaches have topologically similar error fields (such as the initial separation), the PINN-DA-SA demonstrates a vastly reduced error. At data points, the reconstruction error is much lower for PINN-DA approaches than the variational-DA-SA. Figure 6.5 shows that the PINN-DA's have measurement error of the order of 10^{-7} whilst the variational-DA-SA is 10^{-5} . The relative performance of the variational-DA-SA and PINN-DA-SA compared with both RANS-SA and the DNS flow can be summarised in the velocity profiles in Figure 6.10. Whilst both data assimilation methods represent a significant

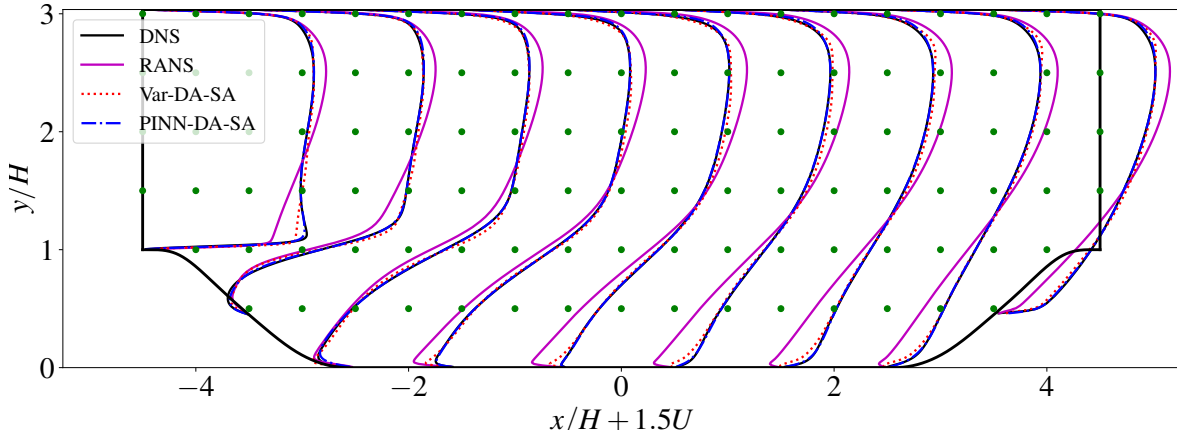


Figure 6.10: mean-flow reconstruction of streamwise velocity profiles for DNS, RANS-SA, PINN-DA-SA and variational-DA-SA.

improvement over a RANS-SA simulation, the PINN-DA-SA is almost identical to the DNS results, whilst small discrepancies are still more apparent for the variational-DA-SA.

Comparing both PINN-DA-SA and variational-DA-SA to PINN-DA-Baseline show that application of SA has been successfully used to reduce error for near-wall gradients, as a result of the additional physical description provided by the turbulence model. The specific choice of turbulence model (SA) is also a contributing factor, given the aforementioned good performance in boundary layers with adverse pressure gradients.

The comparison of PINN-DA and variational-DA is consistent with the findings from the laminar case in Section 6.1. Irrespective of use of the SA model, lower reconstruction error is observed for the PINN-DA approaches compared to the equivalent variational-DA approaches.

Forcing fields

The inferred forcing given by Eq. (3.8) consists of solenoidal and modelled eddy-viscosity (SA contribution) components. For PINN-DA-Baseline, the modelled eddy-viscosity component is zero. A further analysis of the forcing terms allows isolation of the effect of SA in the accurate reconstruction of the mean-flow field.

The PINN-DA-Baseline and PINN-DA-SA forcing fields shown in Figure 6.11 are qualitatively similar in the uniform flow region between $y/H = 1.5$ and $y/H = 2.5$, where turbulent pro-

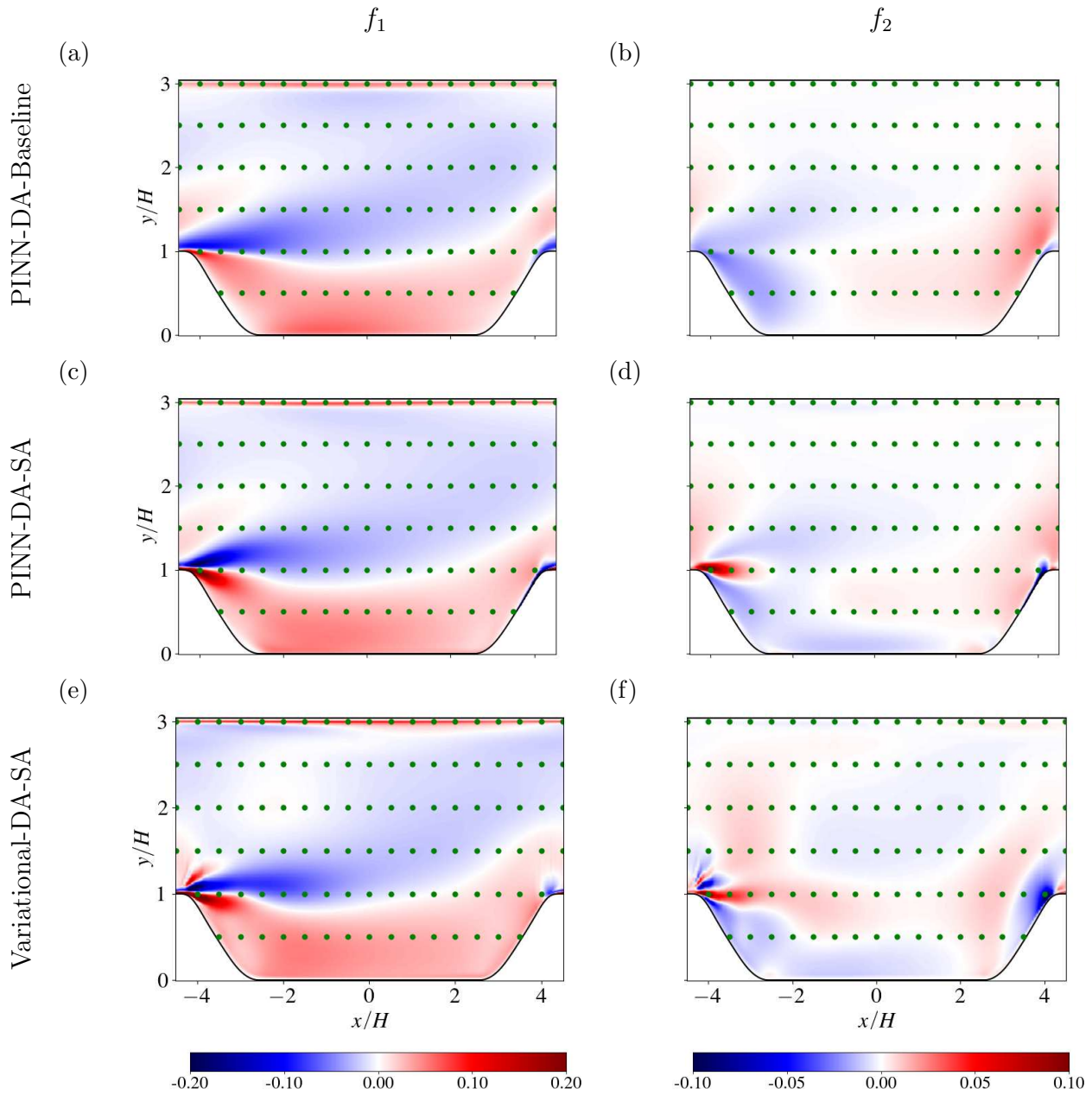


Figure 6.11: A breakdown of the Reynolds forcing (excluding the potential field) f_1 (left) and f_2 (right) from the different methodologies: $f_i = \frac{\partial(2\nu_t S_{ij})}{\partial x_j} + f_{s,i}$. For PINN-DA-Baseline, $\nu_t = 0$. These are (a,b) - PINN-DA-Baseline, (c,d) - PINN-DA-SA, (e,f) - Variational-DA-SA.

duction is low. However, the regions of the flow where PINN-DA-Baseline shows high levels of error (as seen in Figure 6.8(d)) result in larger forcing changes in the PINN-DA-SA solution. Introduction of a modelled component (via SA) has led to significant changes in the corrective forcing, most notably at the separation but also near the walls. These changes underpin the differences in reconstruction error between PINN-DA-Baseline and PINN-DA-SA. Constraining the solution with additional physical definition (via SA) has provided additional modelled

forcing in regions where data-assimilation methods with the under-determined RANS had the highest errors. Whilst the PINN-DA-Baseline had converged to a poor solution, the PINN-DA-SA has reduced the “amount” of under-determined forcing resulting in lower reconstruction errors. However, the addition of the corrective term, which is still unclosed, has allowed the data-assimilation approach to correct the SA model.

The PINN-DA-SA forcing is different to the variational-DA-SA forcing, in accordance with the differences in mean-flow reconstruction. In general, the PINN-DA-SA forcing is closer to the variational-DA-SA forcing than the PINN-DA-Baseline forcing. However, there are several differences in forcing, appearing in regions with the largest differences in reconstruction. This is most apparent in the significantly less smooth forcing around the separation in the variational-DA-SA forcing compared to PINN-DA-SA. The variational-DA-SA also exhibits a high forcing concentration in f_2 forcing along the rear hill, centred around a datapoint. This datapoint (in the $\Delta L = 0.5$ case) has shown to be particularly sensitive. Whilst the gradient regularisation procedure employed in the variational approach ensures that each update of forcing is smooth, it does not guarantee that the accumulation of updates and thus the final forcing will also be smooth. Whilst other regularisation strategies were investigated (see Appendix B.3.2), they led to poorer reconstructed mean velocity fields.

To evaluate the accuracy of the data assimilation methods, the forcing is also compared with the true DNS forcing. As the potential forcing field is absorbed into the pressure term and thus is inseparable, the curl of forcing is instead compared, to remove the contribution of the potential forcing (since $\nabla \times \nabla \phi = 0$).

$$(\nabla \times \mathbf{f})_3 = \frac{\partial f_{s,2}}{\partial x} - \frac{\partial f_{s,1}}{\partial y} + \left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \right) 2\nu_t S_{12} + \frac{\partial^2}{\partial x \partial y} (2\nu_t (S_{22} - S_{11})). \quad (6.2)$$

A direct comparison can be performed then against the curl of the DNS forcing

$$(\nabla \times \mathbf{f})_3 = \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} = \left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \right) \overline{u'v'} + \frac{\partial^2}{\partial x \partial y} (\overline{v'v'} - \overline{u'u'}). \quad (6.3)$$

The curl of the inferred forcing for all three reconstruction methods employed here is shown in

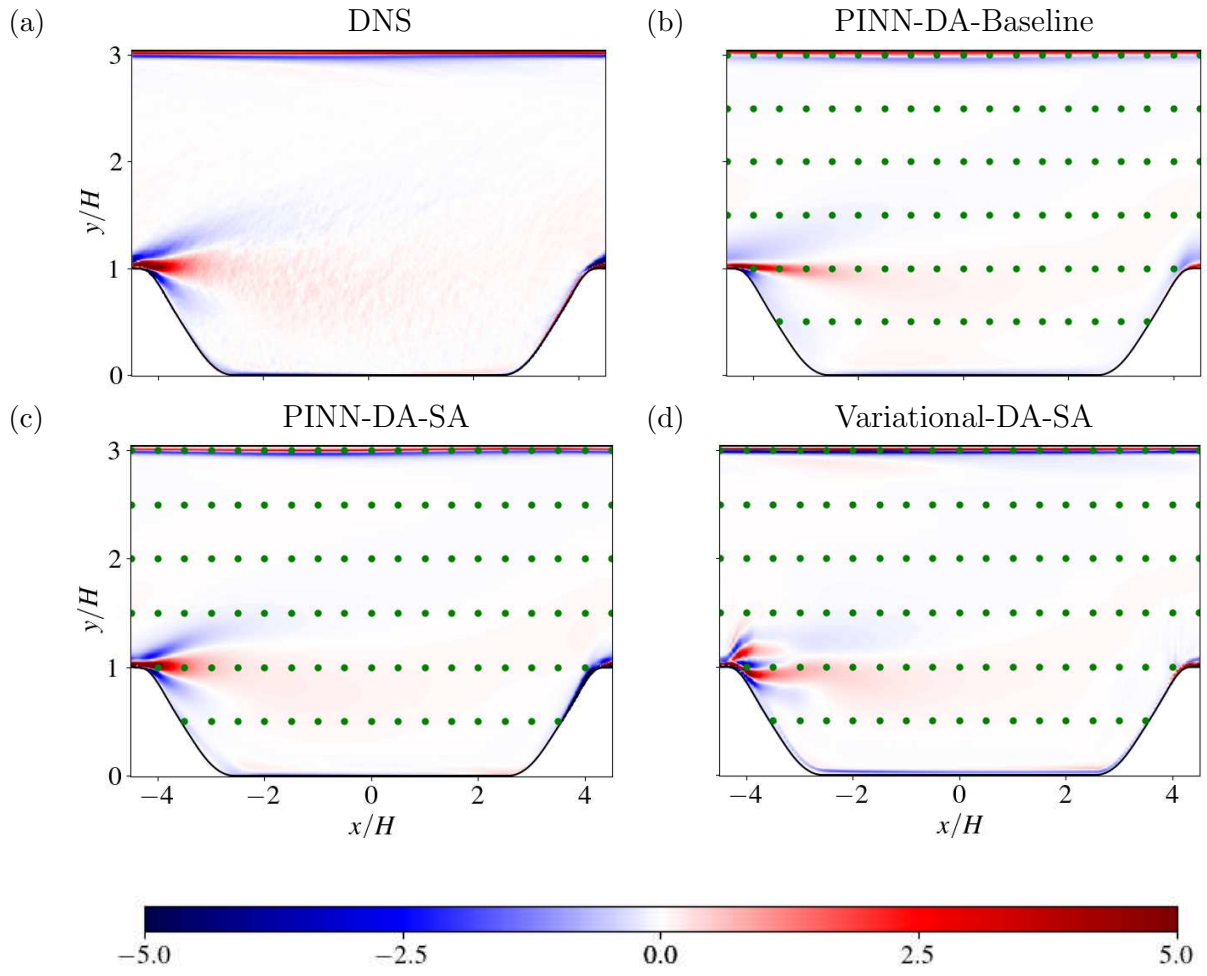


Figure 6.12: $\nabla \times \mathbf{f}$. Comparison of $\text{curl}(\mathbf{f})$ between the ground truth (DNS) and the different data assimilation techniques presented in this paper. (a) DNS, (b) PINN-DA-Baseline, (c) PINN-DA-SA, (d) Variational-DA-SA.

Figure 6.12. Best agreement with the DNS is achieved for PINN-DA-SA, which aligns with the most accurate mean-flow reconstruction. The PINN-DA-Baseline differs from DNS in several key regions - the initial separation, the recirculation, the rearward hill apex and along the walls. Differences between the variational-DA-SA and PINN-DA-SA forcing (and curl) correlate with differences in the error fields. These are most apparent at the separation region and along the upper and lower walls.

Data Assimilation with varying data resolution

Here, the dependence of the reconstruction error on the data resolution of the available high-fidelity measurements is evaluated. In Figure 6.13, variational-DA-SA, PINN-DA-Baseline

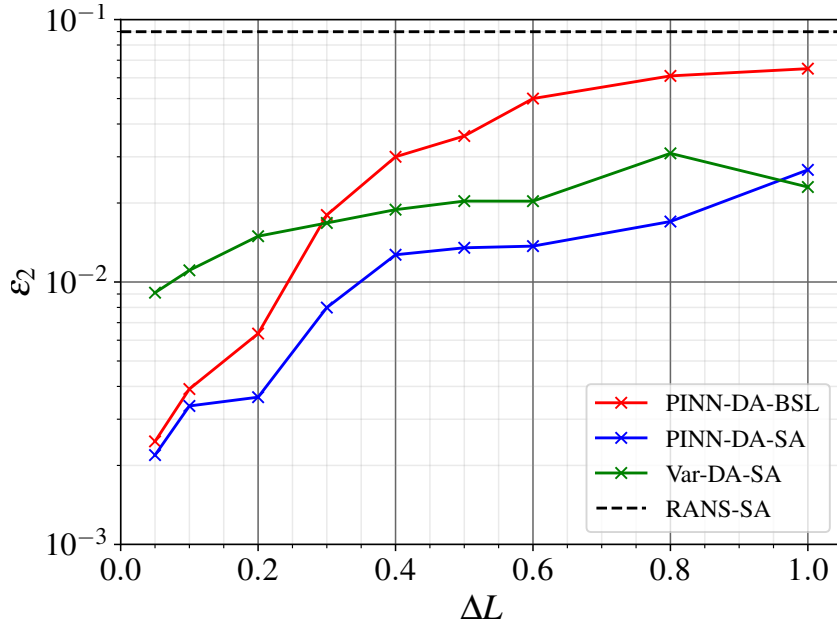


Figure 6.13: Volume-weighted L_2 mean velocity reconstruction error (as defined in Eq. (6.1)) versus data resolution for the three reconstruction methodologies (solid lines with symbols). All three approaches improve the RANS-SA predictions (dashed line).

and PINN-DA-SA approaches are compared, for a range of data-resolutions, ranging from $\Delta L = 0.05$ to $\Delta L = 1.0$. The reconstruction error increases monotonically as the (equidistant/square) spacing between data points increases. All data assimilation methods improve the RANS-SA solution, irrespective of resolution. For data spacing $\Delta L < 1$, the turbulence model augmented PINN-DA-SA method achieves the lowest reconstruction error against all the other methods. Additionally, whilst the $\Delta L = 0.8$ has not converged as well for the variational-DA-SA approach, the reconstruction error remains close to the neighbouring points.

Comparing the PINN-DA-Baseline to the PINN-DA-SA results shows the significant improvement achieved through the turbulence model augmentation. The gap between the two can mostly be attributed to the error near the walls and in the separation. Additionally, the rate of error increase with respect to data resolution is steeper for PINN-DA-Baseline. This suggests adding a modelled (closed) forcing component (i.e. PINN-DA-SA) reduces the effect coarsening data resolution has on reconstruction accuracy compared to PINN-DA-Baseline. As data resolution becomes finer, the two PINN-DA approaches converge. This is expected - as more data is provided, there are more points at which the PINN-DA-Baseline can close the solution and thus the advantage the turbulence model augmentation provides is diminished.

Comparing the PINN-DA-SA versus the variational-DA-SA solution shows that, for most of the resolution range, the PINN-DA-SA also demonstrates consistently lower error at equivalent resolutions. In fact, the PINN-DA-SA at $\Delta L = 0.8$ matches the reconstruction accuracy of the variational-DA-SA at $\Delta L = 0.3$. However, at the extreme end of coarse data ($\Delta L = 1.0$), whilst the PINN-DA-SA is still a significant improvement over the PINN-DA-Baseline (and also RANS-SA), the variational-DA-SA approach has a lower reconstruction error. This suggests that the PINNs have a greater dependence on data for good performance. As data resolution reduces (spacing increases), the optimisation problem converges to solving the direct RANS-SA problem, for which the numerical framework within the variational-DA-SA method is more robust i.e. PINNs are less effective forward solvers. Algorithmic improvements for PINNs, including adaptive weighting algorithms, can potentially provide further improvements of the PINN approach (as demonstrated by Wang, Yu & Perdikaris (2022)), which is currently under investigation. On the opposite end, for fine data resolutions, the variational-DA-SA is less accurate than the PINN-DA-Baseline.

One may attribute the difference between the reconstruction accuracy of the PINN-DA-SA and variational-DA-SA to the differences in the way the RANS equations are enforced in the data assimilation procedure. Concerning the variational-DA-SA approach, it may be first noticed that the latter solves for only the weak form of the RANS equations, following the finite-element approach. In addition, as mentioned in Section 6.2.2, stabilisation terms are added in the governing equations. Then there is the process of spatial discretisation through, among others, the choice of shape functions, which will determine the order of spatial accuracy. In contrast, PINNs enforce the RANS equations in a continuous way. Moreover, the effect of applying discrete pointwise forcing at measurement locations poses a greater challenge for the variational approach than PINNs and thus the method of regularisation has a greater impact on reconstruction performance. The PINN-DA-SA uses an L_2 -based regularisation, which was found to be sufficient to produce a smooth forcing solution. However, for the variational-DA-SA, an L_2 approach with pointwise data enforcement leads to unsmooth, pointwise forcing, thus the use of a H_1 -based method. Appendix B.3.2 contains analysis of regularisation choice. By avoiding both these sources of error, PINN-DA-SA achieves a lower reconstruction error than the

variational-DA-SA. This is evidenced by the reconstruction error for finer data resolutions. Even as data resolution increases, the aforementioned discretisation error remains. Consequently, the reconstruction error of the variational-DA-SA reduces at a much slower rate than both PINN-DAs.

These results demonstrate key advantages and disadvantages in using the PINN and variational formulation for data assimilation. These are summarised in Table 6.3.

Table 6.3: Summary table describing the advantages and disadvantages of using PINN-DA and variational-DA for mean flow reconstruction.

Methodology	Advantages	Disadvantages
PINN-DA	• Using equivalent formulations (turbulence model augmentation), higher reconstruction accuracy relative to variational-DA. • Enforces physics loss in continuous form (no discretisation error). • More robust convergence, achievable even without turbulence model augmentation. • Better convergence of data loss. • Enforcing mean flow and Reynolds stress data, enables inference of pressure. • Code simplicity.	• Requires more significant hyper-parameter tuning (network architecture and loss weights) to achieve sufficient convergence. • Successful convergence of SA equation requires greater consideration of singularity.
Variational-DA	• Better reconstruction accuracy compared to RANS-SA simulation • Unlike PINN-DA which softly constrains physics loss, variational-DA enforces strict adherence to governing equations. • For extremely sparse data, achieves better reconstruction accuracy than PINN-DA.	• Introduction of discretisation error and mesh dependence (also with use of stabilisation terms). • Greater sensitivity to regularisation methodology.

6.5 Conclusion

Three approaches to turbulent mean-flow reconstruction from sparse data measurements have been presented and compared. Firstly, it was shown that for all methods, inferring the closure of the underdetermined RANS equations or the corrective closure of the RANS-SA equations by assimilating high-fidelity sparse data improved the mean-flow reconstruction compared to RANS-SA predictions.

Secondly, the use of Spalart-Allmaras turbulence model, as used in Franceschini et al. (2020) (in the variational-DA-SA approach) was proposed for use in PINNs, particularly in the context of turbulent flows. The SA turbulence model augmented PINN (PINN-DA-SA) proved to be the most accurate flow reconstruction method followed by the equivalent approach albeit with a variational formulation. Comparing the PINN-DA-Baseline (PINN without turbulence SA model) with PINN-DA-SA showed the isolated effect of turbulence model augmentation. The PINN-DA-SA performs significantly better in regions with complex flow features, such as near walls and at separation. Injection of additional physical constraints in the form of a turbulence model provides an additional modelled component to the underdetermined Reynolds forcing, reducing the complexity of the optimisation and narrowing the solution space for the optimisation. A corrective forcing term is then tuned to correct the rest of the closure.

Thirdly, the variational-DA-SA approach, whilst an improvement over the PINN-DA-Baseline, has higher errors compared to the PINN-DA-SA, most notably off the separation. These observations for mean-flow reconstruction accuracy are also associated with equivalent error in Reynolds forcing. The difference in variational-DA-SA and PINN-DA-SA can be attributed to the effect of discretisation, which is avoided by PINNs. These observations are consistent across different data resolutions.

Given that PINNs is a growing field of research, there are potentially many avenues to improve the reconstruction capability further. The most apparent example is hyper-parameter tuning, such as loss weighting. In this study, a static weighting method was used with manual iterations to tune them. This is a laborious exercise, particularly as systems become more complex with

more boundary conditions and less uniform topologies. Adaptive weighting algorithms (van der Meer et al. 2022, Wang, Teng & Perdikaris 2021, Wang, Yu & Perdikaris 2022) or novel architectures, such as Competitive PINNs (Zeng et al. 2022), would both simplify this problem and also generalise the process to selecting weights.

Additionally, it was demonstrated that the data assimilation methods were sensitive to the provided data, such as more accurate reconstruction of the uniform, bulk flow region. It follows that investigating optimal sensor placement may reduce the volume of data required to achieve convergence.

Chapter 7

HBNS constrained PINN-DA for accurate mean-flow reconstruction and HBNS closure models

In this chapter, the use of the harmonic-balanced Navier-Stokes (HBNS) equations with PINNs, to reconstruct and super-resolve mean flow using PINNs, is explored. In the previous chapter, it was shown how augmenting the mean-flow equations with prior empirical physical information (adding a turbulence model) enabled PINNs to super-resolve and reconstruct a mean flow more accurately. However, whilst the turbulence model augmented PINN-DA formulation improves reconstruction accuracy, empirical turbulence models are based on several modelling assumptions and are calibrated for specific flow regimes. Whilst the Navier-Stokes equations have the locality property (the future flow state can be completely determined from the current flow state), due to the effects of flow averaging or filtering, the RANS or SRS equations are non-local (Girimaji 2023, Lumley 1979, Mishra et al. 2015). This means the evolution of the statistical or filtered flow variables, such as the mean flow, is dependent on the history of previous statistical flow states. Specifically, the averaging or filtering operations remove flow physics, which instead appear through non-local effects in the unclosed terms, such as the Reynolds stresses, resulting in a significant non-local spatio-temporal influence. To close

the RANS equations, traditional turbulence modelling assumes the evolution of turbulence (i.e Reynolds stress tensor) at a given point can be determined by local tensors at the same point (locality assumption). However, modelling the non-local evolution of turbulence with a local closure is fundamentally limited and may not be able to predict the true turbulent dynamics. Furthermore, the common LEVM depend on the Boussinesq approximation (the Reynolds stress tensor is proportional to the mean strain-rate tensor), which itself is not valid for all flow regimes. The effectiveness of empirical turbulence model at accurately predicting the underlying flow physics, therefore, is limited by modelling assumptions (e.g. Boussinesq approximation, locality assumption), in addition to the calibration of models with parameter tuning for specific flow regimes.

However, by alternatively using a SRS approach and enforcing frequency-truncated HBNS equations, instead of a turbulence model augmented RANS approach, to provide prior physical knowledge and accurately resolve low-frequency time scales, one can avoid the errors introduced from use of an empirical model. For example, use of the filtered (frequency-truncated) governing fluid equations, to resolve some frequency harmonics, reduces the influence of non-local spatio-temporal effects (although the filtered SRS equations still contain non-local effects, specifically on the unclosed terms). The following chapter explores a hybrid approach between constraining data assimilation with RANS (all scales modelled) and DNS (all turbulent scales resolved), by proposing an alternative, novel approach to enforcing the physics of the (fluid) system and resolving the lowest frequency turbulent time scales, through constraining PINN-DAs with the (frequency-truncated) HBNS equations (instead of the RANS equations).

Similar to other SRS approaches, by enforcing the HBNS equations with PINNs, the lowest N frequency harmonics of the flow are resolved, leaving a smaller fraction of the flow and thus turbulent scales unresolved and unclosed, albeit at the cost of enforcing additional equations and requiring more data. As before, one can use data assimilation with PINNs, to subsequently close the flow. In order to explore the use of the HBNS equations with PINNs (PINN-DA-HBNS), the following key objectives are set out as follows:

- i to investigate how resolving low-frequency harmonics by constraining PINNs with the HBNS

equations and resolved harmonic flow data (instead of an empirical turbulence model) affects its (PINN-DA-HBNS) ability to reconstruct mean flow;

- ii to investigate how a data-driven closure model for the governing HBNS equations can be developed by leveraging the (pointwise) PINN-DA inferred flow fields (resolved velocities and closure terms) found during the aforementioned flow reconstruction.

This analysis will include an investigation into the effects of resolving an increasing fraction of the lowest frequency turbulent scales, through increasing the number of resolved harmonics, N . It was shown in Section 3.1.2, that the HBNS equations, truncated to $N = 0$ harmonics is equivalent to the RANS equations. Therefore, it follows that PINN-DA-HBNS truncated to $N = 0$ harmonics is equivalent to the PINN-DA-Baseline approach from the previous chapter. This is progressed through resolving more harmonic scales, by increasing the order of truncation for the HBNS equations, enforcing additional physics constraints and also providing a greater structure to the solution, as seen with the quasi-linear approaches used in Kaiser & Oberleithner (2021), von Saldern et al. (2023). Unlike the explicit Reynolds stress approach, discussed in the previous chapter (Section 6.4.2), the PINN-DA-HBNS approach resolves a fraction of the Reynolds stresses through enforcement of the frequency-truncated HBNS equations. Understanding the sensitivity to the number of resolved harmonics demonstrates the effects of resolving more turbulent scales and providing more physics knowledge to constrain the data assimilation.

The proposed PINN-DA-HBNS is tested on the laminar circular cylinder case, due to its simplicity, understanding of its behaviour, fundamental frequency and effect of different scales, and its extensive use in other work (as discussed in Section 4.1). The PINN-DA-HBNS results are divided into two parts. Firstly, PINN-DA-HBNS is used to reconstruct a flow from sparse measurements, as in the previous chapter. This extends the previous research and shows how it can be used to tackle the same flow reconstruction problem as already studied (in Sliwinski & Rigas (2022)). Secondly, a new application of PINNs (specifically PINN-DA-HBNS) and flow reconstruction is explored. Whilst PINNs (and other data assimilation methods) have already been used to reconstruct flows, there is a fundamental limitation. Use of the trained model

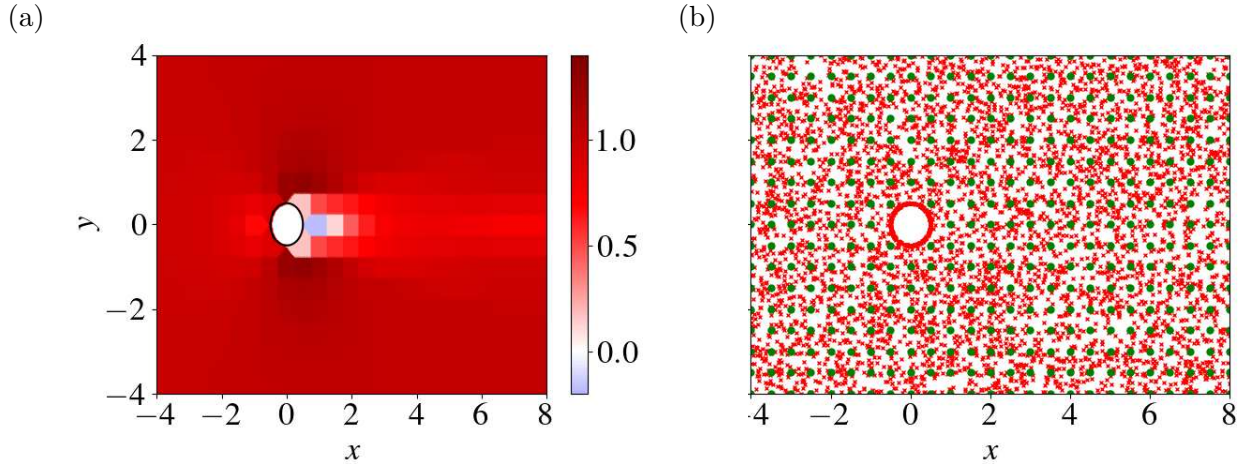


Figure 7.1: Laminar circular cylinder at $Re = 200$ with data sampling showing (a) the coarse, measured velocity field at $\Delta L = 0.5$ (b) a sample collocation point distribution.

is generally restricted to the super-resolution (and inferencing) of the reconstructed flow. In this second part, the use of PINN-DAs is expanded, by leveraging trained PINNs (forming a continuous solution operator) and their capacity to exactly determine flow gradients using automatic differentiation, to build a (pointwise, local) HBNS closure model.

The chapter is laid out as follows. Section 7.1 describes the numerical set up of the problem and implementation of PINN-DA-HBNS for flow reconstruction. The flow reconstruction results with PINN-DA-HBNS follow in Section 7.2. Thereafter, the use of PINNs to develop a closure model is presented in Section 7.3. Concluding remarks are presented in Section 7.4.

7.1 Numerical and PINN setup

7.1.1 Case and data sampling

The use of PINNs constrained by the HBNS equations (Eq. (3.21)) is tested and validated on the laminar circular cylinder at $Re = 200$. DMD was used to extract the harmonic mode shapes, as seen in Figure 4.3. The process of numerical simulation and mode extraction (with DMD) is detailed in Section 4.1. It is noted that, whilst circular cylinder flow with $Re \gtrsim 190$ exhibits three-dimensional wake instabilities, as described in Jiang & Cheng (2021), the present study focuses on validating use of the HBNS equations to reconstruct flow with PINNs, and

therefore, as with the underlying data, the flow is assumed to be two-dimensional, and three-dimensional effects neglected. The behaviour of this flow is well understood, as discussed in detail in Section 4.1. In particular, observing Figure 4.1 at $Re = 200$ highlights the effects of resolving the lowest frequency scales, with an improvement in convergence and accuracy when truncating the HBNS equations and resolving $N = 2$ harmonics instead of $N = 1$. The well-defined behaviour of the laminar circular cylinder with respect to harmonic number, forms a basis to evaluate and compare the performance of the proposed PINN-DA-HBNS approach and whether similar trends hold true for the PINNs.

The PINN-DA-HBNS was trained on a subsection of the full numerical simulation domain, as shown in Figure 4.2, defined by a box bounded at $[-4, 8] \times [-4, 4]$. The high-fidelity data used to train the model was sampled on a rectangular grid within this subdomain. To analyse the effect of data resolution, velocity measurements were sampled at several data resolutions between $0.1 \leq \Delta L \leq 1.0$, where $\Delta L = \Delta x = \Delta y$. Unless otherwise stated, all figures discussed hereon-in will compare results for the resolution $\Delta L = 0.5$ (seen in Figure 7.1(a)). Note how at this resolution, only a single data point sits in the recirculation zone. Finally, the model was trained using velocity data sampled for all N resolved low-frequency harmonics, depending on the selected order of truncation. For example, if the order of truncation was $N = 1$, measurement data was provided for $\hat{\mathbf{u}}_0, \hat{\mathbf{u}}_1$ with $\hat{\mathbf{u}}_n$, $n \geq 2$ not used to train the model (as reflected in Eq. (3.36)).

7.1.2 PINN Setup and hyperparameter optimisation

PINN-DA-HBNS was implemented using the NVIDIA Modulus package (NVIDIA 2024). This package contains several PINN architectures and provides greater control of the training and computational processes and thus was selected over DeepXDE for this problem. All PINNs trained for flow reconstruction in this section employ the architecture presented in Figure 7.2. In this architecture, the PINN-DA-HBNS is composed of three independent neural networks (multilayer-perceptrons), which individually predict a) the velocity modes, $\hat{\mathbf{u}}_n$, b) the composite pressure and potential forcing modes, $(\hat{p} - \phi)_n$, and c) the solenoidal closure terms, $\mathbf{R}_{n,s}$. Each

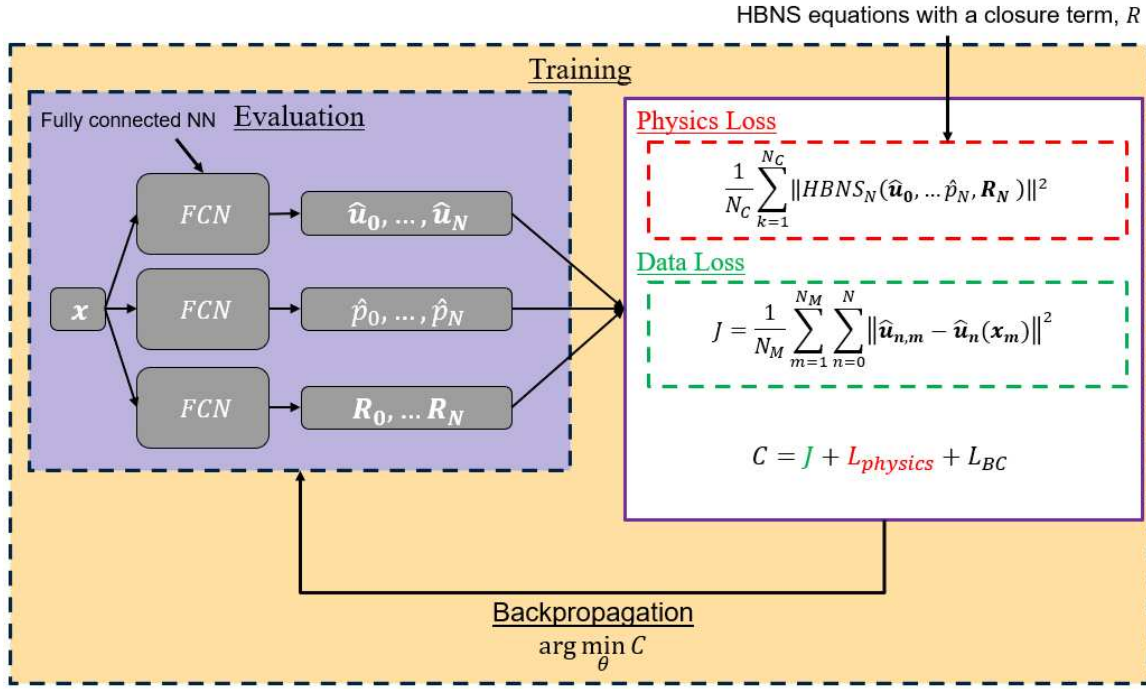


Figure 7.2: PINN-DA-HBNS architecture used in this chapter. The velocity fields, pressure fields and closure fields are separated into 3 separate sub-models. These become coupled through enforcement of the governing equations.

sub-network has 9 layers and 200 nodes/layer, irrespective of number of resolved harmonics, N . The input layer consists of two nodes, representing the spatial coordinates, x and y , which are fed into each of the sub-networks. The number of output nodes vary depending on the number of harmonics. Furthermore, as the equations and harmonic modes are complex, both the neural network flow variable outputs and the enforced governing equations are split into real and imaginary components. This approach is different to the ModalPINN approach in Raynaud et al. (2022), as the HBNS equations are directly enforced, instead of enforcing the Navier-Stokes (with the flow variables calculated as a sum of the truncated Fourier series). Furthermore, in this chapter, the unresolved high-frequency harmonics are not neglected, but accounted for via the closure term $\mathbf{R}_n$. The PINNs use the non-linear swish activation function, in addition to the adaptive activation function method described in Jagtap, Kawaguchi & Karniadakis (2020). Further detail on the selection of the neural network architecture, hyper-parameter selection and use of adaptive activation functions is included in Appendix C.1.

During training, the collocation points, at which the governing equations are evaluated, are generated using a resampling algorithm. After every iteration, 5000 collocation points are

randomly sampled using a random uniform distribution. By resampling collocation points every iteration, one can use fewer collocation points, reducing overall computational costs, whilst still ensuring sufficient coverage across the flow domain over the entire training cycle. To enforce the boundary conditions, an additional 640 points are randomly sampled along the cylinder wall, using the same resampling algorithm. An example collocation point distribution from a single resample is presented in Figure 7.1(b). The residual PDE loss was weighted using a spatial weight distribution function, as described in Sharma et al. (2023). Using the signed distance function to weight the residual loss ($\lambda^P(\mathbf{x}_c)$), collocation points nearer to the wall boundary are weighted slightly lower than away from the wall, preventing high flow gradients near the wall disproportionately affecting the loss function and subsequent optimisation. The choice of collocation point sampling, residual PDE and loss weighting is discussed further in Appendix C.1.

The PINN-DA-HBNS models were trained using the ADAM optimisation algorithm with an exponential learning rate decay. The initial learning rate was set at 1×10^{-3} . The PINNs were trained on one NVIDIA RTX 6000 GPU. Further details on the convergence of PINN-DA-HBNS are found in Section 7.2.2.

7.2 Comparison of flow reconstruction with PINN-DA-HBNS

In this section, use of PINN-DA constrained by the HBNS equations (Eq. (3.21)), to reconstruct 2D circular cylinder flow at $Re = 200$ from sparse velocity measurements is validated. Firstly, a brief summary of the problem formulation is presented. Secondly, the optimisation of PINN-DA-HBNS is discussed. Thirdly, an overview of the PINN-DA-HBNS approach across different numbers of resolved harmonics, N , is presented, considering skin-friction coefficient and velocity reconstruction accuracy across different data resolutions. Finally, the flow reconstructions are examined in greater detail, to provide further insight into the effect of resolving more frequency scales.

7.2.1 Problem formulation

In this section, PINN-DA is used to solve the constrained optimisation, where a solution, $(\hat{\mathbf{u}}_{\mathbf{n}}, (\hat{p} - \phi)_{\mathbf{n}}, \mathbf{R}_{\mathbf{n},\mathbf{s}})(\mathbf{x}, \theta)$, to the frequency-truncated HBNS equations (Eq. (3.21)) is sought, such that J , the measurement error (3.36), is minimised. Furthermore, to evaluate the effect of resolving differing numbers of resolved low-frequency harmonics (ranging from $N = 0$ to $N = 2$), the Fourier series expansion (in time) and enforced governing equations, HBNS $_N$ (3.21), are truncated to N resolved harmonics. The Fourier series expansions, for the $N = 0, 1, 2$ resolved harmonics are

$$N = 0 : \quad \mathbf{w} = \hat{\mathbf{w}}_0, \quad (7.1a)$$

$$N = 1 : \quad \mathbf{w} = \hat{\mathbf{w}}_0 + \hat{\mathbf{w}}_1 e^{i\omega t} + \overline{\hat{\mathbf{w}}_1} e^{-i\omega t}, \quad (7.1b)$$

$$N = 2 : \quad \mathbf{w} = \hat{\mathbf{w}}_0 + \hat{\mathbf{w}}_1 e^{i\omega t} + \overline{\hat{\mathbf{w}}_1} e^{-i\omega t} + \hat{\mathbf{w}}_2 e^{2i\omega t} + \overline{\hat{\mathbf{w}}_2} e^{-2i\omega t}. \quad (7.1c)$$

In the following section, the PINN-DA-HBNS reconstruction approach is referred to by the corresponding order of truncation (of the Fourier series/HBNS equations) i.e. $N = 0, 1, 2$. For example, using the expansion in Eq. (7.1a), which is equivalent to the PINN-DA-Baseline formulation, is referred to as the $N = 0$ approach. Likewise, the quasi-linear expansion, from Eq. (7.1b), is also referred to as the $N = 1$ approach, whilst resolution of two harmonics, as in Eq. (7.1c), is also referred to as the $N = 2$ approach.

To demonstrate the effect of frequency truncation on HBNS $_N$, the momentum equations, acting on the mean flow ($\hat{\mathbf{w}}_0$), have been presented for differing N

$$N = 0 : \quad \hat{\mathbf{u}}_0 \cdot \nabla \hat{\mathbf{u}}_0 + \nabla(\hat{p} - \phi)_0 - \nu \Delta \hat{\mathbf{u}}_0 + \mathbf{R}_{0,\mathbf{s}} = 0, \quad (7.2a)$$

$$N = 1 : \quad \hat{\mathbf{u}}_0 \cdot \nabla \hat{\mathbf{u}}_0 + \nabla(\hat{p} - \phi)_0 - \nu \Delta \hat{\mathbf{u}}_0 + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1} + \overline{\hat{\mathbf{u}}_1} \cdot \nabla \hat{\mathbf{u}}_1 + \mathbf{R}_{0,\mathbf{s}} = 0, \quad (7.2b)$$

$$N = 2 : \quad \hat{\mathbf{u}}_0 \cdot \nabla \hat{\mathbf{u}}_0 + \nabla(\hat{p} - \phi)_0 - \nu \Delta \hat{\mathbf{u}}_0 + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1} + \overline{\hat{\mathbf{u}}_1} \cdot \nabla \hat{\mathbf{u}}_1 \\ + \hat{\mathbf{u}}_2 \cdot \nabla \overline{\hat{\mathbf{u}}_2} + \overline{\hat{\mathbf{u}}_2} \cdot \nabla \hat{\mathbf{u}}_2 + \mathbf{R}_{0,\mathbf{s}} = 0. \quad (7.2c)$$

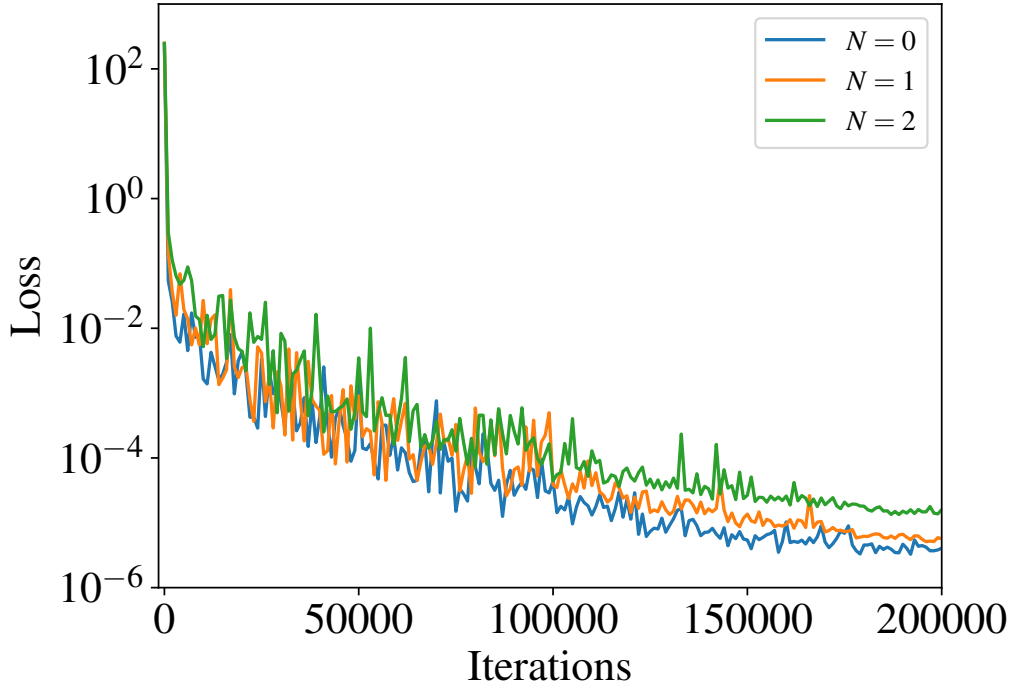


Figure 7.3: Convergence of loss during optimisation of PINN-DA-HBNS for differing number of resolved harmonics.

These equations show how resolving more flow harmonics reduces the effect of the unclosed, unresolved forcing, $\mathbf{R}_{0,s}$. This is similar to the PINN-DA-SA formulation (Eq. (3.9)) used in Chapter 6), in which most of the closure is modelled (with eddy viscosity) using a turbulence model and a small corrective forcing component is found (using data assimilation). However, in this chapter, instead of modelling a forcing component with an empirical turbulence transport equation, additional physics knowledge is added to the data assimilation, without introducing associated modelling errors (from the turbulence model), by resolving low-frequency harmonics using the frequency-truncated HBNS equations.

The following sections explore how PINN-DA-HBNS, constrained by different number of resolved harmonics N (as detailed in Eq. (7.2)), affect mean-flow reconstruction accuracy.

7.2.2 Optimisation

The convergence of PINN-DA-HBNS training loss for varying N resolved harmonics, using data sampled at $\Delta L = 0.5$, is presented in Figure 7.3. The models were trained for 200,000 iterations,

reducing loss from the randomised initial state to a low value. Whilst all three cases achieve similar levels of final loss convergence, as the number of resolved harmonics increase, the total residual loss increases slightly for each added resolved harmonic, as seen when comparing the three cases. This can be attributed to the number of terms in the loss function. Specifically, as the number of resolved harmonics increase, the number of equations, data and boundary loss terms increase and thus the total loss inherently includes more terms. One can also observe spikiness during optimisation of the PINN-DAs. This is a result of random resampling of collocation points at every iteration and is equivalent to the effect seen with use of mini-batches.

7.2.3 Reconstruction error versus data resolution

To evaluate the reconstruction accuracy for differing number of resolved harmonics, an L_2 -based error, ε_n , was calculated for each n -th velocity harmonic,

$$\varepsilon_n = \frac{1}{N_P} \sum_{p=0}^{N_P} \|\hat{\mathbf{u}}_{n,p} - \check{\mathbf{u}}_{n,p}(\mathbf{x}_p)\|^2, \quad (7.3)$$

where N_P is the number of points in the domain, $\check{\mathbf{u}}_{n,p}$ is the n -th PINN-reconstructed velocity mode at the p -th point, $\mathbf{x}_p$ and $\hat{\mathbf{u}}_{n,p}$ is corresponding high-fidelity velocity mode. Also note that for $n \geq 1$, $\hat{\mathbf{u}}_{n,p}$ is complex. Subsequently, ε_0 describes the error of the $n = 0$ velocity harmonic (mean-flow) reconstruction, whilst ε_1 shows the reconstruction error of the first, fundamental velocity harmonic.

The effect of data resolution on reconstruction accuracy, tested at different number of resolved harmonics, N , is presented in Figure 7.4. Data resolutions ranging from $\Delta L = 0.1$ to $\Delta L = 1.0$, for harmonics $N = 0, 1, 2$, have been examined. As spacing between measurements increases, the reconstruction error increases monotonically, irrespective of the number of resolved harmonics. Further inspection of the mean-flow ($n = 0$) reconstruction error in Figure 7.4(a) highlights the effects of data resolution and resolving more low-frequency scales (more resolved harmonics). At the finest data resolution, the effect of resolving more scales is marginal. However, as the data resolution coarsens, a divergence between between resolving $N = 0$ (i.e. Eq.

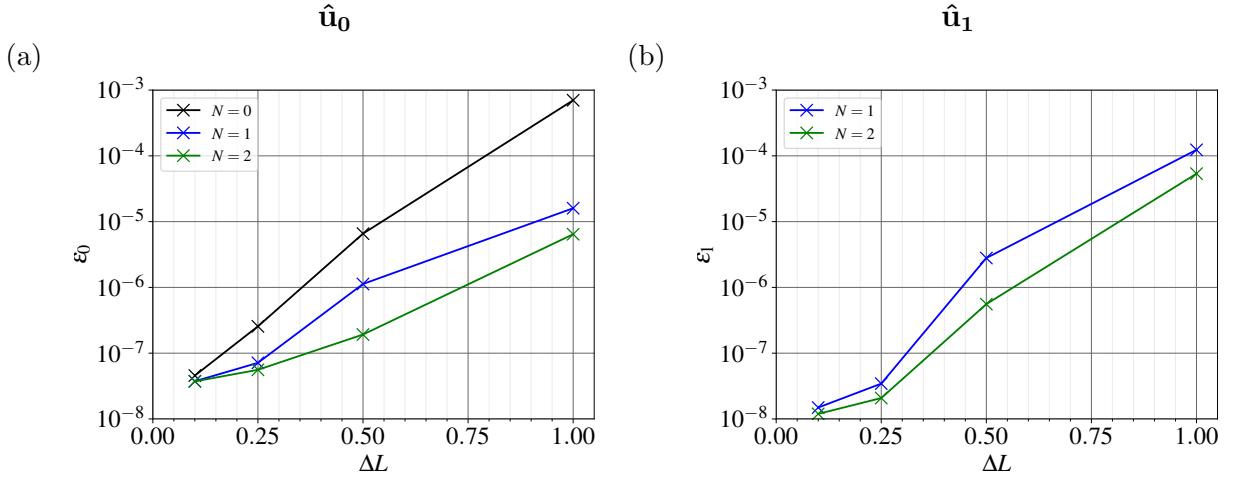


Figure 7.4: Velocity reconstruction accuracy versus data resolution for differing number of resolved harmonics, N . (a) $n = 0$ harmonic (mean) velocity reconstruction error (b) $n = 1$ harmonic velocity reconstruction error.

(7.1a)) and resolving more harmonics is observed. The reconstruction error grows at the highest rate with respect to coarsening data for $N = 0$. Likewise, whilst the difference between a quasi-linear ($N = 1$ - Eq. (7.1b)) approach and the $N = 2$ case (Eq. (7.1c)) is small at $\Delta L = 0.25$, coarsening the data resolution, increases the differences in reconstruction accuracy between different numbers of resolved harmonics. This result shows i) the dependence of PINN-DA-HBNS reconstruction accuracy on the resolution of more low-frequency scales, with $N = 2$ resolved harmonic case having the highest reconstruction accuracy, followed by resolving $N = 1$ harmonics and then $N = 0$ (equivalent to PINN-DA with RANS) case, ii) the improvement resulting from resolving more low-frequency scales increases as data resolution decreases.

One can attribute the observed improvements in reconstruction accuracy of the mean flow directly to the growing number of resolved low-frequency scales of the flow. This leads to a subsequent decrease in magnitude and importance of the unresolved, high-frequency scales and resultant closure term, $\mathbf{R}_s$, decreases in magnitude (and importance). This explains the improvement from $N = 0$ to the quasi-linear ($N = 1$) approach and the $N = 2$ case thereafter. This is observed for both the mean-flow reconstruction (as discussed prior) but also in the reconstruction of the first harmonic, as shown in Figure 7.4(b). By resolving the $n = 2$ higher frequency harmonic, the $n = 1$ harmonic is reconstructed more accurately, compared with the quasi-linear approach. This in turn propagates to the mean velocity mode, which is thus

also reconstructed more accurately. As more higher frequency harmonic scales are accurately resolved, using the corresponding frequency truncation of the governing (HBNS) equations and resolved velocity measurements, the lower frequency harmonics are more accurately captured.

The improvement in reconstruction of the mean flow as the order of frequency truncation increases is comparable to the effect of spatial discretisation (and truncation) when comparing the adjoint-based variational-DA and PINN-DA in Chapter 6. In Chapter 6, it was shown that, by avoiding the spatial discretisation error when enforcing the governing equations, PINN-DA was able to reconstruct mean flow more accurately than the equivalent variational-DA. Likewise, in this section, increasing the number of resolved harmonics, N , and thus truncating fewer harmonics reduces the temporal truncation error. This shows that reduction in discretisation or truncation error (both spatial or temporal) consistently reduces reconstruction error, as seen in both Chapter 6 and this chapter.

For the rest of this section, the figures and results presented consider reconstructions using the $\Delta L = 0.5$ data resolution.

7.2.4 Skin friction coefficient

Another metric to evaluate flow reconstruction accuracy is the comparison of the reconstructed skin-friction coefficient, C_f , with the high-fidelity data. This is calculated as

$$C_f = \frac{\tau_w}{\frac{1}{2}\rho U_\infty^2} = \frac{2}{Re} \frac{\partial \hat{u}_{0,\theta}}{\partial r}, \quad (7.4)$$

where τ_w is the wall shear stress, $\hat{u}_{0,\theta}$ is the mean velocity component tangential to the cylinder wall and $\frac{\partial}{\partial r}$ is the flow gradient in the wall normal (radial) direction.

Figure 7.5 compares the reconstructed skin-friction coefficient for different number of resolved harmonics, with the high-fidelity, DNS C_f . Irrespective of the number of resolved harmonics, N , the peak skin-friction coefficient (approx. 0.33) is accurately estimated, with the initial C_f profile for PINN-DA reconstructions indistinguishable from the high-fidelity DNS profile.

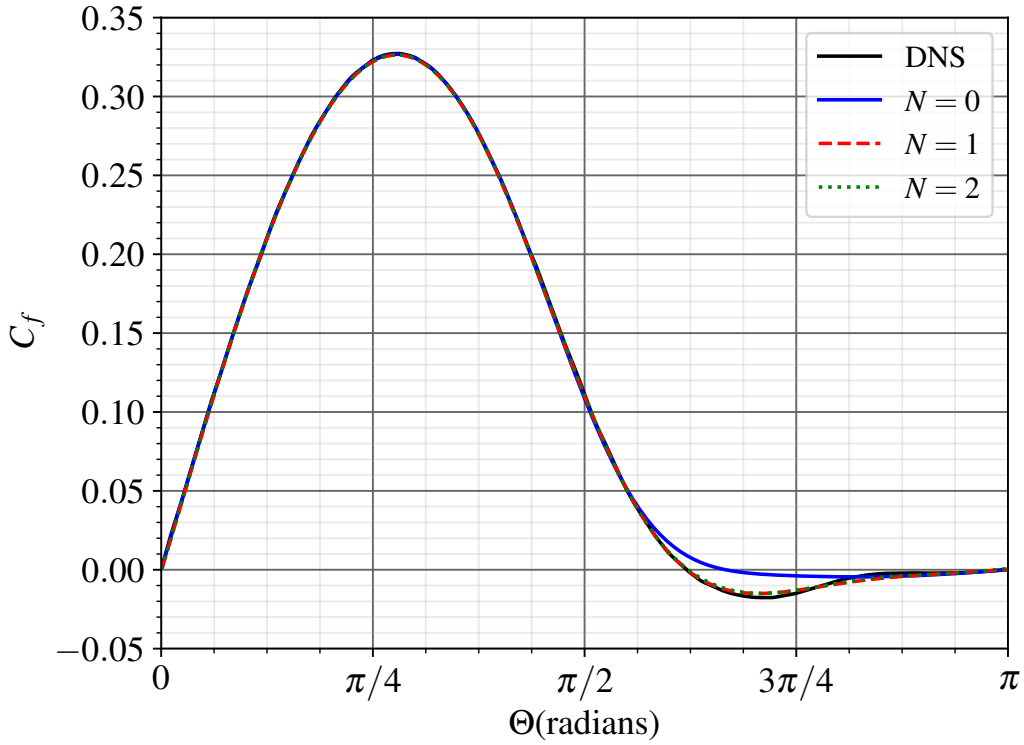


Figure 7.5: Skin-friction coefficient profile for DNS and PINN reconstructions for varying number of harmonics using data resolution $\Delta L = 0.5$.

However, around the flow separation, differences between the number of resolved harmonics (with PINN-DA-HBNS) appear. Using a quasi-linear approach, the prediction of the skin-friction coefficient in the region around the separation, $\Theta \geq \pi/2$, is improved relative to the $N = 0$ case. Here, the profile of the reconstructed skin-friction coefficient converges to the DNS data. However, unlike the velocity field reconstructions (described in Section 7.2.3), where reconstruction accuracy improves when resolving $N = 2$ harmonic scales compared with $N = 1$, the skin-friction coefficient is convergent after resolving two harmonics and increasing N further is not expected to produce significantly greater improvements in skin-friction coefficient. This result demonstrates the effectiveness of resolving additional low-frequency harmonic scales ($N \geq 1$) as skin-friction coefficient accuracy increases.

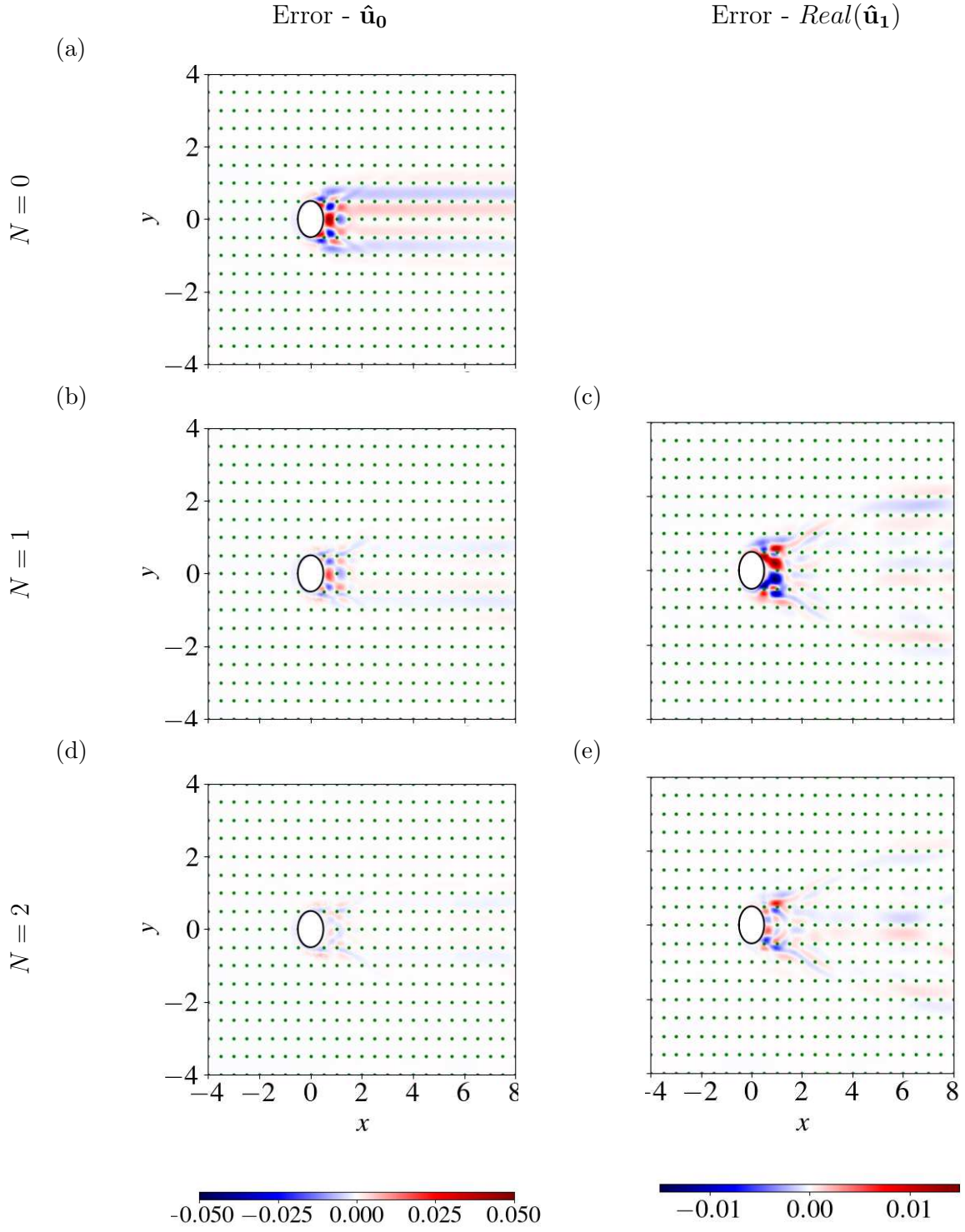


Figure 7.6: Flow reconstruction error for streamwise component of (a,b,d) mean flow ($\hat{\mathbf{u}}_0$) and (c,e) first velocity harmonic ($\hat{\mathbf{u}}_1$) for different number of resolved low-frequency harmonics, $N = 0, 1, 2$.

7.2.5 Reconstruction of the mean flow using a RANS ($N = 0$), a quasi-linear ($N = 1$) and higher order ($N = 2$) approach.

In this section, the reduction in reconstruction error seen in Figure 7.4, from resolving low-frequency harmonics is investigated, describing how the use of the (frequency-truncated) HBNS equations, to resolve some low-frequency scales, bridges the gap between RANS and DNS, resulting in improved PINN-DA-HBNS reconstruction accuracy. Whilst only the reconstruction error fields are presented here, the reconstructed fields have been included in Appendix C.2.

Firstly, the reconstruction when using $N = 0$ resolved harmonics is evaluated i.e. RANS formulation (see Eq. (7.2a)), presented in Figure 7.6(a). As demonstrated in Section 3.1.2, resolving $N = 0$ harmonics is effectively the same as performing data-assimilation using the underlying RANS equations. In this case ($N = 0$), only mean velocity data (i.e. $\hat{\mathbf{u}}_0$) is provided to reconstruct the flow. Whilst only the streamwise velocity component is shown, the same trends are also found in the vertical velocity component. When resolving only $N = 0$ harmonics, the total mean velocity reconstruction error is $\varepsilon_0 = 6.6 \times 10^{-6}$, as seen in Figure 7.4(a), whilst the maximum error observed in this reconstruction is 0.048. As expected, these reconstruction results are consistent with previous studies (Sliwinski & Rigas 2022). Firstly, high reconstruction error is predominantly observed in areas with high velocity gradients. In particular, the highest error magnitude is found in and around the near wake, within the recirculation region. There is also noticeable error in the near wall region, where the boundary layer and separation occur. Specifically, error in the latter of these (separation) explains the corresponding error in skin friction coefficient, seen in Figure 7.5. Further downstream, in the far wake, the reconstruction error, whilst still present, begins to diminish. Furthermore, in the regions with uniform flow or with small velocity gradients (above and below the cylinder), the reconstruction is accurate with minimal error. Lastly, as seen in many PINN-DA reconstructions, specifically at locations where data measurements are provided, the mean flow is accurately reconstructed as error is almost zero (and data loss is convergent). As one moves away from the measurement locations, the error begins to rise again. This is most apparent in the wake region. This result shows that use of a RANS constrained PINN-DA formulation ($N = 0$) using sparse mean velocity

measurements can reconstruct the flow well, albeit with noticeable error in specific regions.

Secondly, the use of a quasi-linear ($N = 1$) PINN-DA-HBNS approach (see Eq. (7.2b)) is examined, as presented in Figure 7.6(b). Topologically, the reconstruction error field is qualitatively similar to the $N = 0$ case (Figure 7.6(a)). In the uniform flow region, the quasi-linear case is indistinguishable from the $N = 0$ case. The already small magnitude of error combined with the limited unsteadiness in this region, means that the resolution of the lowest (vortex-shedding) frequency scale provides little new information and thus benefit here. Similarly, the principal reconstruction errors occur in the same regions as before (i.e. the wake and near wall flow field). However, the quasi-linear PINN-DA-HBNS ($N = 1$) has significantly reduced error magnitude compared to RANS formulation ($N = 0$). For the quasi-linear PINN-DA-HBNS, the total mean velocity reconstruction error reduces to $\varepsilon_0 = 1.1 \times 10^{-6}$ (a six-fold reduction in total error), and the maximum observed error magnitude is 0.019 (a reduction of 60%). These differences are most clearly observed in the near wake. The mean-flow reconstruction around the separation is also significantly improved, explaining the improvement in skin-friction coefficient seen in Figure 7.5. There is also a noticeable reduction in error in the far wake. This result shows that despite increasing the complexity of the PINN-DA-HBNS optimisation (greater number of flow variables and governing equations), resolving the first low, vortex-shedding frequency harmonic, improves the capacity of PINN-DA-HBNS to reconstruct the mean flow accurately over the RANS ($N = 0$) formulation.

Thirdly, the effect of resolving $N = 2$ harmonic scales (see Eq. (7.2c)) on mean-flow reconstruction is presented in Figure 7.6(d). As with all prior cases, the reconstruction error field is qualitatively similar to all other reconstructions of the laminar circular cylinder. However, there is another, albeit smaller, reduction of reconstruction error, relative to the quasi-linear method. The reconstruction error when resolving $N = 2$ harmonics has been reduced to $\varepsilon_0 = 1.9 \times 10^{-7}$ (a reduction of 83% compared to $N = 1$ and 97% compared to $N = 0$). The largest improvement is again observed in the recirculation. Whilst the near-wake has exhibited the highest magnitude of error for both prior cases ($N = 0, 1$), this has been further reduced, as seen in Figure 7.6(d). The maximum error magnitude has been reduced to 0.011 (a reduction of 42% compared to $N = 1$ and 77% compared to $N = 0$), which is comparable to the error in the far

wake. Similar improvements are also observed in the near wall and separation regions, whilst error at measurement locations and in the uniform flow region is typically small. The reducing improvement from resolving $N > 1$ harmonic scales is consistent with both the skin-friction coefficient prediction in Figure 7.5 as well as the findings in Rigas & Schmid (2022), where using a quasi-linear model substantially improved mean drag prediction, but resolving another low-frequency scale had a much smaller effect. Figure 7.6(a,b,d) shows that whilst resolving additional low-frequency scales does increase mean velocity reconstruction accuracy, there are diminishing returns, as seen when resolving the second harmonic.

Reconstruction of fundamental velocity harmonic

Examination of the reconstruction of the fundamental harmonic velocity field, as presented in Figure 7.6(c,e) (showing reconstruction error of the first harmonic streamwise velocity field) explains i) the improvements in mean velocity reconstruction through increases in the number of resolved harmonics and ii) the diminishing returns in reconstruction accuracy when resolving the second harmonics. Reconstruction of the first (fundamental) velocity harmonic when using the quasi-linear approach, as seen in Figures 7.6(c), is accurate across most of the domain. However, in the area corresponding to the recirculation region, high reconstruction error is observed in the first resolved harmonic velocity field, with the highest error magnitude approaching 10% of the true field ($\hat{\mathbf{u}}_1$). For the quasi-linear case, the high reconstruction error for first harmonic velocity correlates with the region that also has the highest error in the mean velocity reconstruction (as seen in Figure 7.6(b)).

A similar comparison of the reconstructed first harmonic, but when resolving $N = 2$ harmonics, presented in Figures 7.6(e), demonstrates how resolving an additional harmonic frequency scale results in an improvement in mean velocity reconstruction (as observed in Figure 7.6(d)). By resolving another harmonic ($N = 2$), the reconstruction error of the first harmonic velocity is significantly reduced in the near wake, where quasi-linear showed a large error. The reconstruction error of this field is $\varepsilon_1 = 5.6 \times 10^{-7}$ (for $N = 2$) compared to $\varepsilon_1 = 2.8 \times 10^{-6}$ for the quasi-linear approach, a reduction of 80%. Likewise, the maximum reconstruction error of the

reconstructed first harmonic is approximately half that of the quasi-linear approach (0.008 from 0.015 when resolving $N = 1$ harmonics). Resolving the second harmonic reduces the reconstruction error of the first harmonic velocity field relative to the quasi-linear PINN-DA-HBNS. Comparing the effect of resolving two low-frequency harmonics with just one shows how the reconstruction error of the first harmonic velocity is reduced, and importantly in the region where this reduction predominantly occurs, this directly corroborates with the equivalent reduction in mean velocity reconstruction error.

The improvement in mean-flow reconstruction accuracy from resolving the fundamental harmonic mode can also be partially attributed to the resolution of coherent flow structures (i.e. vortex shedding mode). As discussed in Chapter 2, coherent flow structures are characterised by flow physics which is largely influenced by non-local spatio-temporal correlations. As a result, they have a stronger effect on the non-local closure terms. By using an $N \geq 1$ approach for data assimilation of the laminar circular cylinder, the fundamental vortex shedding mode is resolved, reducing the influence of the non-local effects compared with the $N = 0$ approach, subsequently improving mean-flow reconstruction accuracy.

Residual forcing

Whilst previously, the effect of resolving different numbers of low-frequency harmonics on resolved velocity reconstruction was discussed, in this section the effect on the inferred residual (solenoidal) forcing terms (here $\mathbf{R}_{0,s}, \mathbf{R}_{1,s}$), accounting for the unresolved high-frequency flow harmonics are discussed. $\mathbf{R}_{1,s}$ is presented in Figure 7.7. By resolving more low-frequency harmonics, the number and subsequent effect of the truncated, unresolved frequency scales is reduced, with a subsequent reduction in the magnitude of the residual forcing. The unresolved residual forcing acting on the $n = 1$, x-momentum equations has a maximum magnitude of approximately 0.15, when using the quasi-linear approach. However, after another low-frequency harmonic is resolved ($N = 2$), this is reduced by over three times (maximum magnitude of approximately 0.04). Resolving more frequency scales both directly captures a greater portion of the flow behaviour (low-frequency), whilst also simultaneously diminishing the contribution

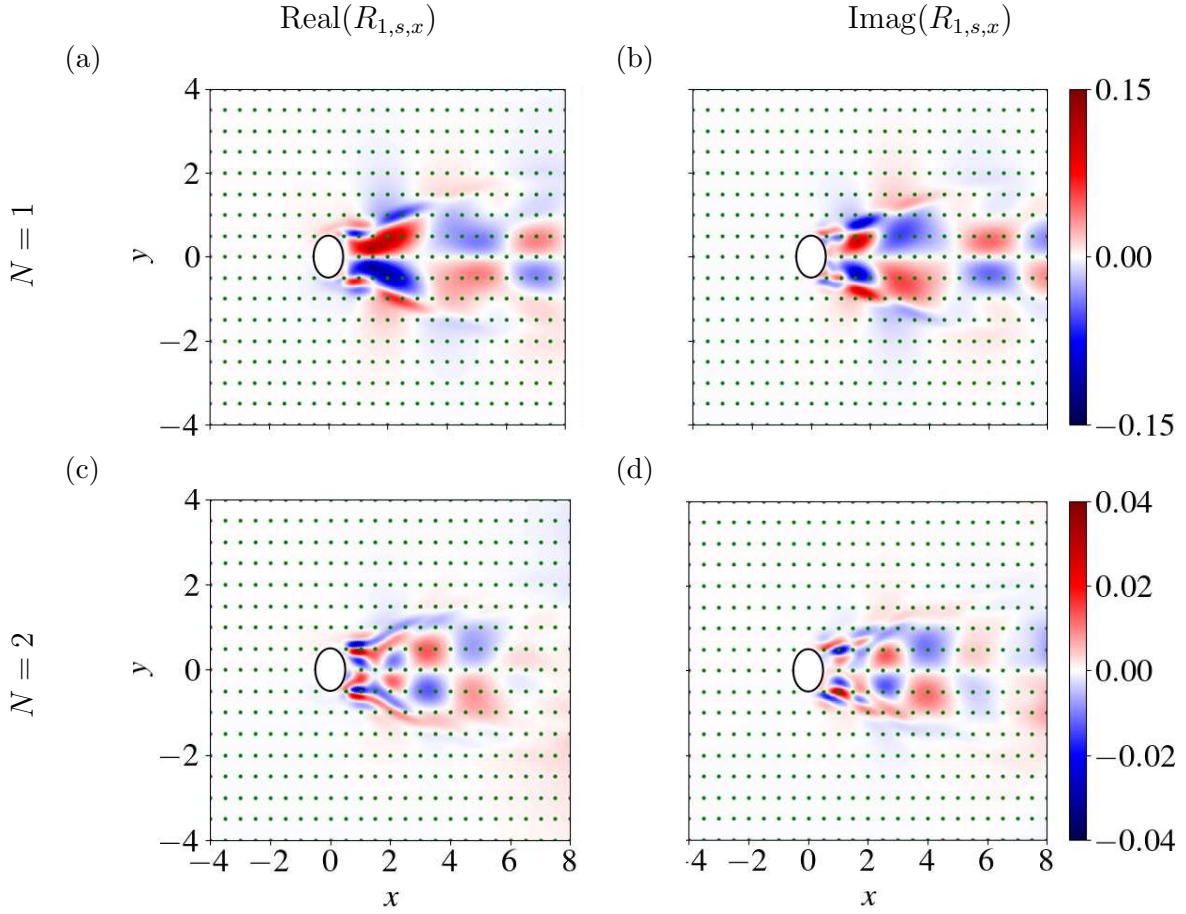


Figure 7.7: Residual forcing on $n = 1$ x-momentum equations when resolving (a,b) $N = 1$ harmonics and (c,d) $N = 2$ harmonics

of the under-determined forcing terms, unresolved scales and non-Markovian effects on the full flow solution, as seen in Eq. (7.2). This can also be seen by analysing the equations for residual unresolved forcing (Eqs. (3.17) and (3.18)). Whilst increasing the number of harmonics that are resolved does increase the complexity of the optimisation (more equations, flow variables and more neural network outputs), provided sufficient convergence of the PINN-DA-HBNS is achieved (for the more complex, higher harmonic formulation), the resolution of these scales must also improve the reconstruction accuracy, as demonstrated in Figure 7.4.

Finally, one can compare the reconstructed, residual forcing in the $n = 0$ (mean-flow) momentum equations, $\mathbf{R}_0$ with the equivalent high-fidelity DNS forcing, to evaluate the accuracy of PINN-DA-HBNS approach. The DNS forcing, $\mathbf{R}_0$ was calculated using Eq. (3.18), truncated to $N = 10$ resolved harmonics (well above the number of required harmonics found in Rigas & Schmid (2022)). However, use of the Helmholtz decomposition means direct comparison with

the DNS forcing is not possible, due to the removal of the potential forcing component. To eliminate the effect of this potential forcing component, in order to compare the DNS forcing with the inferred, reconstructed solenoidal forcing, $\mathbf{R}_{0,s}$, the curl of forcing is calculated instead as in Equation (7.5):

$$\nabla \times \mathbf{R}_0 = \nabla \times (\nabla \phi_0 + \mathbf{R}_{0,s}) = \nabla \times \nabla \phi_0 + \nabla \times \mathbf{R}_{0,s} = \nabla \times \mathbf{R}_{0,s}. \quad (7.5)$$

The comparison of curl of forcing, presented in Figure 7.8, highlights the qualitative similarity between the DNS and PINN-DA-HBNS inferred forcing. The similarity between the curl of the reconstructed and high-fidelity DNS residual forcing indicates that the PINN-DA-HBNS is correctly capturing the effects of the unresolved, high-frequency harmonic scales. Despite the qualitative similarities between high-fidelity and reconstructed curl of forcing, there are some discrepancies particularly in the near wake, correlating with the region of highest velocity reconstruction error. Whilst these discrepancies initially appear more apparent when more harmonics are resolved (such as in Figure 7.8(e,f)), the aforementioned reducing contribution from unresolved scales and therefore residual forcing component means the impact on mean-flow reconstruction accuracy of error in these forcing components is also reduced, as the resolved scales become increasingly dominant. The most significant reduction in (curl of) forcing magnitude is observed when switching between resolving $N = 0$ harmonics to resolving $N = 1$ harmonics (the magnitude of curl for the quasi-linear approach is less than 20% of the $N = 0$ case). Furthermore, there is a smaller, albeit still noticeable, reduction from resolving $N = 1$ harmonics to $N = 2$ harmonics. This corroborates with the trends seen in the flow reconstruction accuracy discussed prior (as in Figure 7.4(a)). The largest improvements in reconstruction accuracy, were seen when a quasi-linear approach was used (instead of resolving $N = 0$ harmonics), which correlates with the largest reduction in the unresolved forcing component. Subsequently, smaller improvements in reconstruction accuracy and smaller reductions in the effect of unresolved frequency scales are observed as N is increased again. The curl of forcing, therefore, demonstrates both the accuracy of PINN-DA-HBNS in capturing the effect of the unresolved high-frequency scales with the residual forcing as well as describing the di-

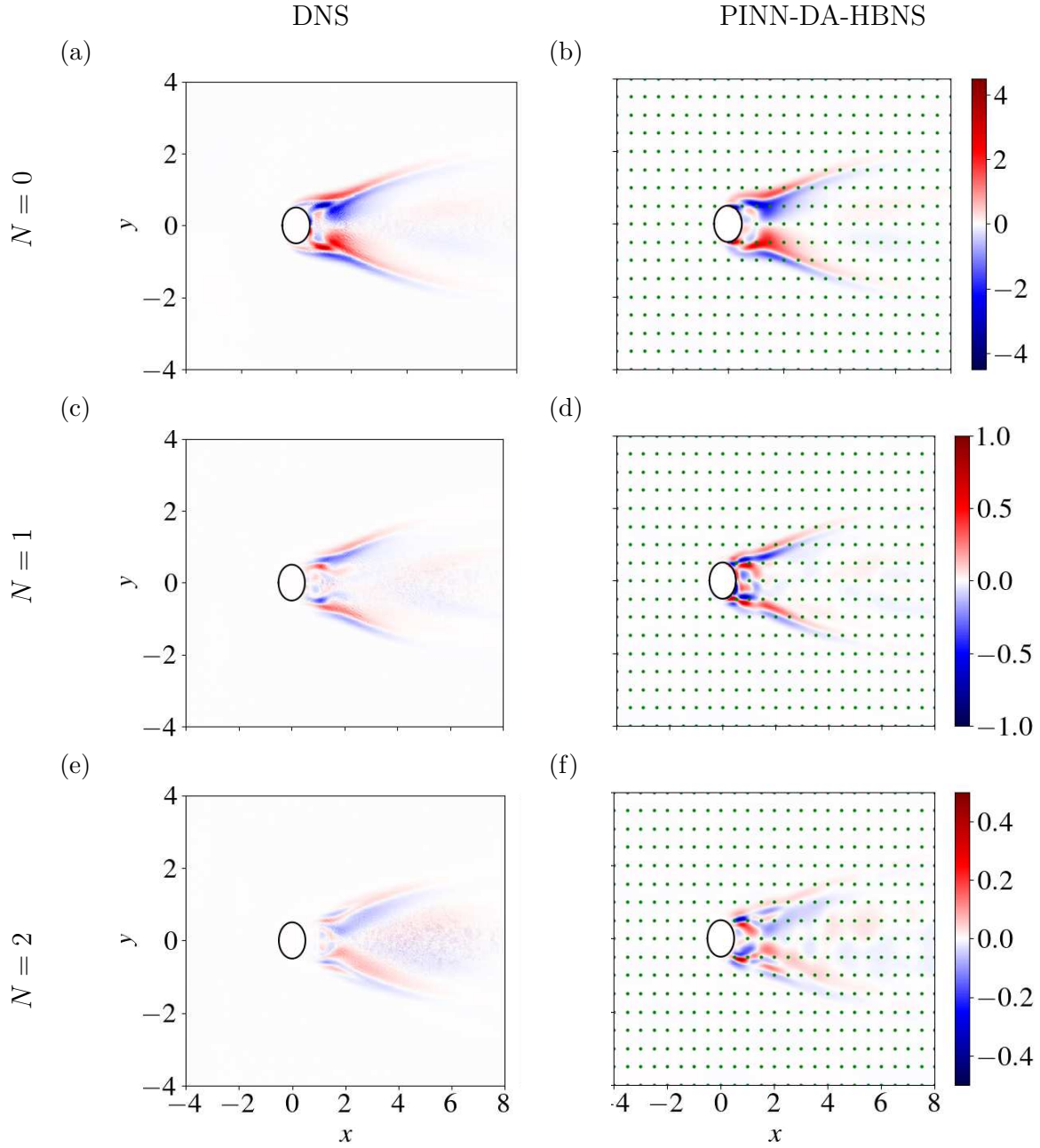


Figure 7.8: Curl of forcing, $\nabla \times \mathbf{R}_0$ for (a,c,e) the high-fidelity DNS data and (b,d,f) the PINN-DA-HBNS reconstruction. This compares (a,b) $N = 0$, (c,d) $N = 1$, (e,f) $N = 2$ resolved harmonics.

minishing improvements in mean-flow reconstruction accuracy from resolving more harmonic scales.

7.3 Expanding flow reconstruction with PINN-DA to build HBNS closure models

In this section, the use of PINN-DA-HBNS to generate datasets, which can subsequently be used to build HBNS closure models, is proposed. Whilst the previous section demonstrates how PINN-DA-HBNS can be used to super-resolve a flow field from coarse velocity measurements and infer the residual, closure field, it is limited in its capacity to only infer fields specific to the individual trained case. However, by reconstructing the flow field using PINN-DA, the volume of data and information about the flow is increased and can be used for applications, such as building data-driven closure models (using methods such as the FIML approach (Holland et al. 2019, Singh & Duraisamy 2016) amongst others). Performing data assimilation with PINNs to reconstruct several flow cases, therefore, enables the generation of a dataset from sparser (experimental) high-fidelity data measurements, which can subsequently be used to build (HBNS) closure models.

The schematic, presented in Figure 7.9, shows the proposed framework for using PINN-DA-HBNS to develop HBNS closure models. Reconstructing flows with PINN-DA-HBNS for several cases, is used to form a dataset using the inferred, super-resolved fields (solenoidal closure terms $\mathbf{R}_{n,s}$ and velocity fields). Thereafter, using pointwise, local input and output features (as in traditional LEVM turbulence models), an HBNS closure model is trained (built using a traditional neural network), as presented in Figure 7.9, to map the residual closure term at a single point from local tensors (e.g. resolved velocity) at the same point (i.e. a closure model based on the Markovian assumption).

The remainder of this section, demonstrates how datasets generated from sparse high-fidelity measurements using the PINN-DA-HBNS methodology (for the laminar circular cylinder at different Reynolds numbers), presented in Section 7.2, are used to build data-driven HBNS closure models. Sections 7.3.1 and 7.3.2 describe the the setup and optimisation of the data-driven closure model. This is followed by a the results of the data-driven HBNS closure model, discussing the accuracy of the trained model on an unseen Reynolds number and the effects of

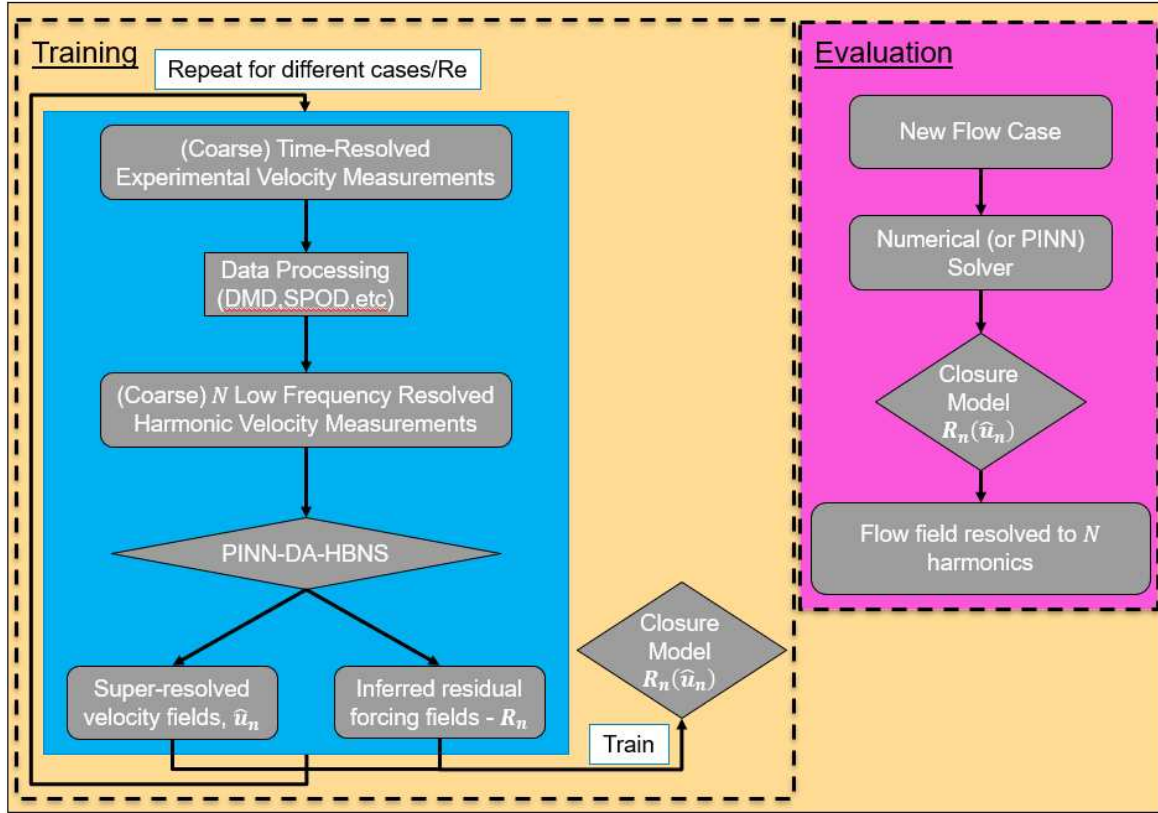


Figure 7.9: Schematic to show how sparse high-fidelity data can be leveraged using PINN-DA to produce a dataset, which can then subsequently be used to construct closure models.

resolving a low-frequency harmonic, and use of velocity (Section 7.3.3) and velocity gradient (Section 7.3.4) model input features.

7.3.1 Case setup and neural network configuration

To build the HBNS closure model, the training dataset was generated from flow fields, reconstructed with PINN-DA-HBNS, as described in Section 7.2. To this aim, PINN-DAs were trained for the laminar circular cylinder case across several different Reynolds numbers ($Re = 150, 200, 300$) and with differing numbers of resolved harmonics ($N = 0, 1, 2$) using the same approach and architecture as in the previous section. This data was split by Reynolds number into a training dataset ($Re = 150, 200$) and a testing dataset ($Re = 300$), which was used to evaluate the closure model and its capacity to extrapolate. However, due to the small magnitude of residual forcing when resolving $N = 2$, the uncertainty and noise of the predictions are of a similar order to the forcing itself, thus the approach less successful. These have

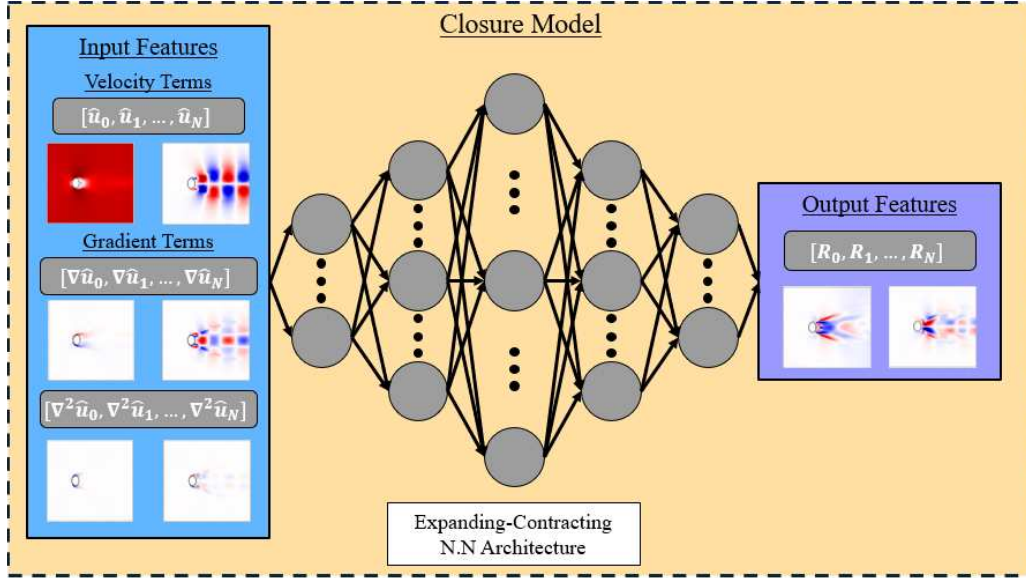


Figure 7.10: Schematic of neural network closure model to find $\mathbf{R}_n = G(\hat{\mathbf{u}}_1)$. The neural network represents a pointwise mapping, in which the low-frequency resolved velocities (and gradients) at a single (spatial) point are used to determine the low-frequency residual closures at the equivalent location.

been included in Appendix C.3.

A data-driven HBNS closure model, $\mathbf{R}_{n,s}(\hat{\mathbf{u}}_1)$, as described in Eq. (3.29), is built, mapping pointwise resolved velocity (and velocity gradient) input features at a single spatial point to the residual closure term at the equivalent location. Firstly, an optimisation problem is defined, such that the error between the closure model predicted ($\check{\mathbf{R}}_{n,s}$) and true, PINN-inferred (solenoidal) closure term, $\mathbf{R}_{n,s}$, is minimised:

$$C = \frac{1}{N_P} \sum_{p=0}^{N_P} \sum_{n=0}^N \|\check{\mathbf{R}}_{n,s}(\hat{\mathbf{u}}_1^p) - \mathbf{R}_{n,s}^p\|^2, \quad (7.6)$$

where, N_P is the number of datapoints that the closure model is trained on. The optimal weights and biases, $\check{\theta}$, defining the HBNS closure model, $\check{\mathbf{R}}_{n,s}(\hat{\mathbf{u}}_1; \theta)$, are sought such that this cost function (7.6), is minimised:

$$\check{\theta} = \arg \min_{\theta} C(\check{\mathbf{R}}_{n,s}(\hat{\mathbf{u}}_1; \theta)). \quad (7.7)$$

Using a DNN, as seen in Figure 7.10, a data-driven closure model was built, mapping low-

frequency resolved velocity harmonics, $\hat{\mathbf{u}}_1$, at a single location to the corresponding residual terms, $\mathbf{R}_n(\hat{\mathbf{u}}_1)$, accounting for high-frequency, unresolved scales, as detailed in Section 3.1.2. This architecture, built using Pytorch packages, is based on the approach in (Rigas & Schmid 2022) using a decoder-encoder (widening-narrowing) neural network architecture. This architecture choice is motivated by use of the expanding decoder segment to map the low-frequency resolved scales to the high-frequency unresolved scales, whilst the contracting encoder section maps to the low-frequency residual terms (as demonstrated in Rigas & Schmid (2022)). Point-wise, resolved velocity information is provided as input features into the neural network closure model, whilst the residual closure terms are outputs. In order to investigate the importance of gradient information to train HBNS closure models (as discussed in Section 3.1.2), use of first- and second-order resolved velocity gradients, in addition to resolved velocity fields, were also investigated as input features. As mentioned in the previous paragraph, the data used to train (and test) the HBNS closure model are sampled from the pre-trained PINN-DA-HBNS, allowing fast and accurate sampling of velocity fields and their derivatives (calculated exactly, through use of automatic differentiation).

Finally, the neural network was trained using the ADAM optimisation algorithm for 10000 iterations and a learning rate of 10^{-3} . This was completed on one NVIDIA RTX 6000 GPU.

The following sections demonstrate how the proposed data-driven HBNS closure model framework is used to map the low-frequency residual term at a given point as a function of the low-frequency resolved velocity harmonics at the same point, evaluating the effects of using resolved velocity and resolved velocity gradient input features and resolving $N = 0, 1$ low-frequency harmonics. For clarity, this has been summarised in Table 7.1.

7.3.2 Optimisation

The convergence of the closure model during training, presented in Figure 7.11, shows the effect on training loss of both choice of model input features (i.e. with or without resolved velocity gradient information) and the number of resolved harmonics. One can observe that irrespective of the number of resolved harmonics, introducing resolved velocity gradient information as a

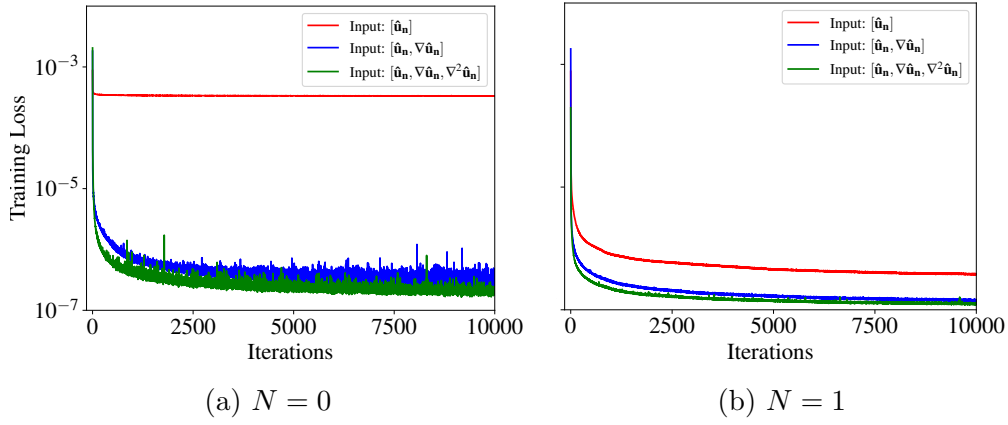


Figure 7.11: Convergence of training loss during optimisation of closure model for differing number of resolved harmonics (a) $N = 0$ (b) $N = 1$.

closure model input, improves the predictions of the closure model on the training dataset. Additionally using first-order resolved velocity gradients as an input feature reduces training loss relative to closure models using only velocity, whilst also using second-order velocity derivatives results in a further, albeit smaller, reduction in training loss. However, the degree of improvement is not the same across different number of resolved harmonics. The $N = 0$ closure model demonstrates a reduction in training loss of 99% when including first-order velocity derivatives and further 30% reduction when including second-order derivatives. This is in contrast, with 62% and 12% for the $N = 1$ closure model. This shows that for training the closure model, i) introduction of velocity gradients as input features consistently improves training performance, although the scale of improvement reduces for models with more resolved harmonics and ii) inclusion of second-order derivatives has the lowest training loss followed by using first-order gradients and finally using only velocity.

Input Features	Gradient inputs features?	Resolved harmonics, N	$n = 0$ mom. eq.	Prediction figure
$\hat{\mathbf{u}}_0$	No	0	(7.2a)	7.12(c)
$\hat{\mathbf{u}}_0, \hat{\mathbf{u}}_1$	No	1	(7.2b)	7.12(d)
$\hat{\mathbf{u}}_0, \nabla \hat{\mathbf{u}}_0$	Yes	0	(7.2a)	7.12(e)
$\hat{\mathbf{u}}_0, \hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_0, \nabla \hat{\mathbf{u}}_1$	Yes	1	(7.2b)	7.12(f)
$\hat{\mathbf{u}}_0, \nabla \hat{\mathbf{u}}_0, \nabla^2 \hat{\mathbf{u}}_0$	Yes	0	(7.2a)	7.12(g)
$\hat{\mathbf{u}}_0, \hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_0, \nabla \hat{\mathbf{u}}_1, \nabla^2 \hat{\mathbf{u}}_0, \nabla^2 \hat{\mathbf{u}}_1$	Yes	1	(7.2b)	7.12(h)

Table 7.1: Summary of different input features used to build data-driven HBNS closure model, $\mathbf{R}_{n,s} = G(\bullet)$, mapping low-frequency resolved harmonics to low-frequency residual terms.

Figure 7.11 also highlights how the number of resolved harmonics affects the closure model convergence. When closure models are built using only velocity inputs, resolving $N = 1$ harmonics enables the closure model to reduce training loss from 3.3×10^{-4} (for $N = 0$ closure model) to 3.8×10^{-7} . This effect is diminished when velocity gradient information is included in the input features, as training loss reduces from 3.0×10^{-7} to 1.4×10^{-7} . This result shows that the number of resolved harmonics used with the closure model is important, as a reduction in training loss is observed when shifting from resolving $N = 0$ to $N = 1$ harmonics, but that this effect is smaller for closure models, which use resolved velocity gradient input features.

7.3.3 Closure model using only velocity field input features

Rigas & Schmid (2022) successfully built a data-driven closure model for the HBNS equations with $N = 1$ resolved harmonics using only resolved velocity as an input feature. However, suitability of this approach was not demonstrated for $N = 0$ (or $N > 1$). Therefore, in this section, the closure model approach used in Rigas & Schmid (2022) is examined, to investigate the effect of the number of resolved harmonics and evaluate its suitability and limitations (with additional results in Appendix C.3 for $N = 2$).

Figure 7.12, showing the HBNS closure model prediction of $R_{0s,x}$ for the unseen test case ($Re = 300$), demonstrates a key limitation of building the closure model with only velocity input features. The $\mathbf{R}_0$ prediction for the $N = 0$ closure model, which is presented in Figure 7.12(a,c) and is effectively a RANS closure model, shows how using only velocity, the $N = 0$ closure model fails to accurately capture the residual term, particularly in the vicinity of the cylinder body, where large prediction artefacts are observed. This is consistent with the excessively high training loss observed in Figure 7.11(a). There is some improvement away from the cylinder, although discrepancies are still observed. This demonstrates that building an $N = 0$ (RANS) closure model using velocity input features is not possible.

However, examining the R_0 prediction for the $N = 1$ closure model, where both mean and first harmonic velocities are provided as HBNS closure model input features (seen in Figure 7.12(b,d)), demonstrates contrasting behaviour (to the $N = 0$ closure model). As in Rigas

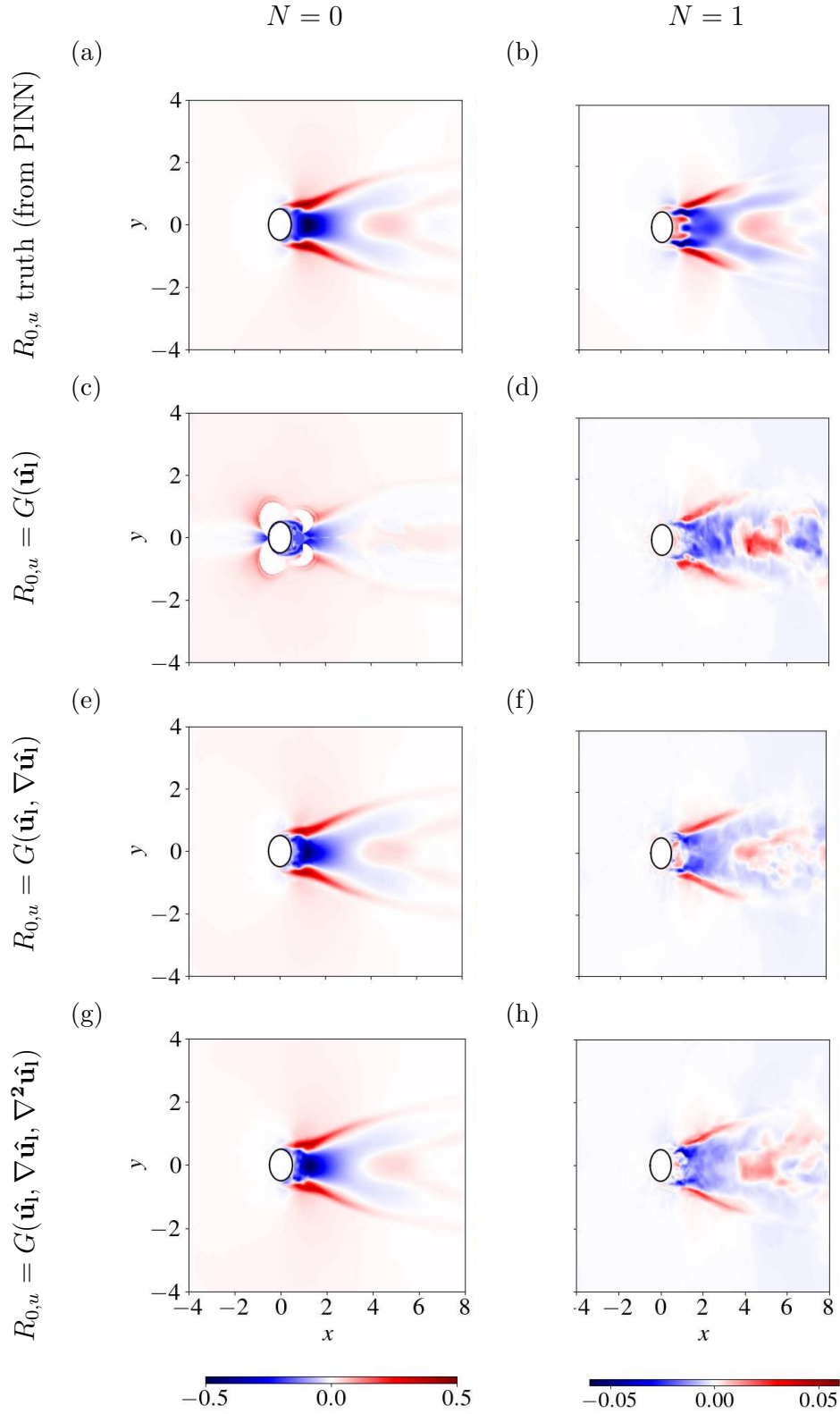


Figure 7.12: True (PINN-DA-HBNS inferred) (a,b) and closure model predictions for $R_{0,u}$ field when trained with (c,d) velocity variables ($\mathbf{R}_n = G(\hat{\mathbf{u}}_1)$), (e,f) velocity gradients ($\mathbf{R}_n = G(\hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_1)$), (g,h) velocity laplacian ($\mathbf{R}_n = G(\hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_1, \nabla^2 \hat{\mathbf{u}}_1)$). This is compared for resolved harmonics (left) $N = 0$, (right) $N = 1$. Predictions are for $Re = 300$ after training on $Re = 150, 200$.

& Schmid (2022), the $N = 1$ closure model with only velocity data can predict the closure terms with reasonable accuracy. Although the largest discrepancies appear immediately aft of the cylinder body, correlating with the recirculation region, the error is much smaller than the $N = 0$ closure model, which is also consistent with the convergence seen in Figure 7.11(b). These results also show that the $N = 1$ HBNS closure model is generalisable, as it has successfully extrapolated to a higher unseen Reynolds number. These findings, provide a key insight into the importance of harmonic information, demonstrating that whilst a RANS closure model ($N = 0$) can not be built only with velocity information, introducing resolved harmonic velocities ($N \geq 1$) as model inputs is sufficient to accurately build an HBNS closure model.

One can consider the difference between the $N = 0$ and $N = 1$ closure model, using velocity input features in terms of the non-local dynamics of the closure term. As previously mentioned, due to flow averaging (or filtering) the HBNS equations (including the RANS equations), are non-local. Specifically, flow physics is filtered out, which then appears a non-local effects in the unclosed, residual term. In particular, coherent flow structures (e.g. vortex shedding modes) have significant non-local spatio-temporal effects on the closure terms (as discussed in Section). However, as in traditional LEVM, the data-driven HBNS closure model used in this section, assumes that the closure is local. By resolving coherent structures (e.g. vortex shedding mode), one can reduce the influence of non-local effects on the closure term (Girimaji 2023), improving the quality of the locality assumption. Therefore, resolution of the fundamental ($N \geq 1$) flow harmonic and therefore also a coherent structure (vortex shedding mode) sufficiently diminishes the non-local effects, such that one is able to build a data-driven closure model using only resolved velocity input features (unlike $N = 0$ closure model).

7.3.4 Closure models using resolved velocity gradient input features

Whilst inspection of the residual closure terms of the HBNS equations (Eqs. (3.17) and (3.18) and repeated here in Eq. (7.8)) shows the importance of the resolved harmonic velocities (as demonstrated in Section 7.3.3), Eq. (7.8) (and also Eq. (3.15)) also indicates the importance

of resolved velocity gradient information.

$$\underbrace{in\omega\mathbf{M}\hat{\mathbf{w}}_1 + \mathcal{L}_P[\hat{\mathbf{w}}_1] + \mathcal{P}_l(\mathcal{N}_P[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1])}_{\text{HBNS}_N(\hat{\mathbf{w}}_1)} + \underbrace{\mathcal{P}_l(\mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h] + \mathcal{N}_P[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h] + \mathcal{N}_P[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_1])}_{\mathbf{R}_n \equiv F(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h)} = \mathbf{0}, \quad (7.8a)$$

$$\mathbf{R}_n = \sum_{\substack{n=n_1+n_2 \\ n_1 \leq n_2 \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) - \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}). \quad (7.8b)$$

Specifically, one can observe how $\nabla \hat{\mathbf{u}}_1$ appears directly in the formulation of the residual term through the non-linear operator i.e. $\mathcal{N}_P[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h]$. Furthermore, projection of HBNS equations onto a high-frequency subspace (Eq. (3.15b)) demonstrate how the unresolved high-frequency harmonics, $\hat{\mathbf{u}}_h$, can be expressed in terms of the resolved low-frequency scales (and its derivatives), $[\hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_1, \nabla^2 \hat{\mathbf{u}}_1]$. Therefore, it follows that provision of resolved velocity gradients as closure model input features should improve capacity to capture the closure terms, $\mathbf{R}_n$ and thus improve closure model accuracy (as seen in Eq. (3.29)). This is also similar to the Boussinesq approximation, which assumes the evolution of turbulence is proportional to the mean strain-rate tensor, and LEVM, which use turbulence transport equations that include (mean) velocity gradients, to determine eddy viscosity. In the following section, the effect of using (first and second-order) resolved velocity gradients, found through automatic differentiation of the pre-trained PINN-DA-HBNS, is evaluated.

The effect of including resolved velocity gradients as HBNS closure model input features on prediction accuracy of R_0 are presented in Figure 7.12(e-h). Firstly, one can observe a significant improvement in prediction of R_0 for the $N = 0$ (RANS) closure model, when first-order velocity gradients are included as input features. Unlike when only velocity is used, as in Figure 7.12(c), the $N = 0$ closure model, presented in Figure 7.12(e), is now able to accurately capture the residual closure term, explaining the large reduction in training loss seen in Figure 7.11(a). This is consistent with existing RANS turbulence models (both data-driven and traditional), which use gradient terms. For closure models with a higher number of resolved harmonics ($N \geq 1$), the addition of first-order gradients has a much smaller effect, although improvements in the prediction of R_0 from the closure model are still observed, mostly in the near wake region. One can thus conclude from Figure 7.12(e,f) that including first-order velocity gradients, when

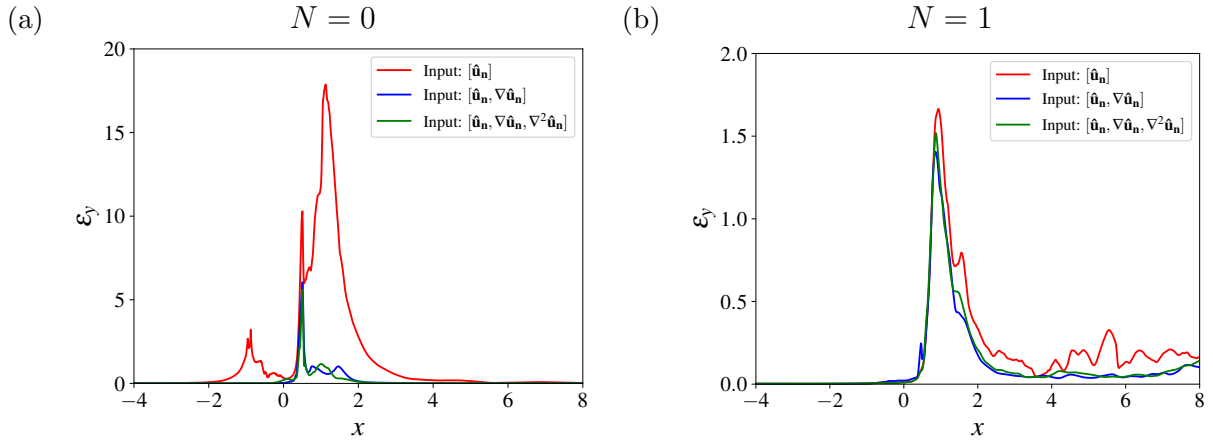


Figure 7.13: Streamwise error distribution for predicted residual closure terms $\mathbf{R}_n$, for HBNS closure model resolved with different numbers of low-frequency harmonics. (a) $N = 0$ (b) $N = 1$. This is evaluated for the testing case, $Re = 300$.

building a closure model, improves accuracy compared to using only velocity (and for the $N = 0$ closure model is necessary).

In contrast, additionally using second-order velocity gradients as input features does not appear to have a significantly noticeable effect, irrespective of number of resolved harmonics. This can be partially attributed to effects of optimisation. As the number of input features increase by adding second order derivatives, the optimisation problem becomes more complex, leading to some overfitting. Likewise, the magnitude of $\nabla^2 \hat{\mathbf{u}}_n$ can be many orders higher $\hat{\mathbf{u}}_n$ or $\nabla \hat{\mathbf{u}}_n$ also affecting optimisation. This correlates with the convergence of the training loss, in Figure 7.11, showing that the small reduction in training loss from introducing second-order velocity gradients as input features, leads to insignificant changes in the prediction of the residual closure terms, and is not an improvement on using only first-order derivatives closure model input features.

To evaluate the spatial dependence of the error in the predicted closure terms, correlation with specific flow features, and the effect of the number of resolved harmonics and input features, the following integrated error was calculated,

$$\varepsilon_y = \frac{1}{L_y} \int^y \sum_{n=0}^N \|\check{\mathbf{R}}_{n,s} - \mathbf{R}_{n,s}\|^2 dy. \quad (7.9)$$

Evaluation of the HBNS closure model prediction error with respect to streamwise location, presented in Figure 7.13, shows that for all cases there is a large increase in error occurring aft of the cylinder body (i.e $0.5 \geq x \leq 2.0$), correlating with the most complex region of the flow, where the initial separation, recirculation and largest flow gradients occur. This is also the region where the highest PINN-DA-HBNS reconstruction error occurs, indicating that error in the underlying PINN-DA-HBNS dataset has possibly also propagated to the HBNS closure model. Thereafter, further downstream, the prediction error reduces to a small value as the flow unsteadiness and fluctuations begin to decay. Examination of this aft-body peak across the different number of resolved harmonics show that the magnitude of this error decreases when using the $N = 1$ HBNS closure model. This reduction in peak error emphasises the importance of resolving more high-frequency harmonic scales.

One can also observe that introduction of resolved velocity gradients does have a positive effect, with reductions in both the maximum error, in addition to a reduction in prediction error further downstream. This is most apparent for the $N = 0$ closure model, but is still observed for closure models using $N \geq 1$. This figure also stresses how the addition of second-order gradient information has almost identical performance to closure models with the addition of first-order gradients only. These trends, associated with the effects of adding velocity gradient input features, are consistent with those seen prior. Analysis of the streamwise error quantitatively shows that inclusion of resolved velocity gradient input features, improves the performance of the HBNS closure models, but that this is secondary to the improvements seen from building a closure model with more resolved harmonics ($N = 1$).

Figures 7.12 and 7.13 all show that use of more resolved scales and resolved velocity input features, reduce the HBNS closure model prediction error in the near wake, highlighting their physical importance in more complex flow areas. These regions pose a greater challenge (including with flow reconstruction) due to higher flow gradients and more complex interactions. Provision of more flow information (velocity gradient input features and/or resolving more harmonics), specifically with input features that are represented in the underlying formulation of the residual closure terms (Eq. (7.8)) and allow the solving of the high-frequency scales (as in Eq. (3.15b)), enable the HBNS closure model to capture these interactions better, resulting in

a more accurate closure model, especially in these (near-wake) areas.

7.4 Conclusion

Use of PINNs constrained by the Harmonic-Balanced Navier-Stokes equations to reconstruct a laminar circular cylinder flow from sparse velocity measurements was presented, after which the reconstructed fields were used to build a HBNS closure model. Use of the HBNS equations allow the gap between RANS (all turbulent scales modelled) and DNS (all turbulent scales resolved) to be bridged by resolving the lowest frequency scales and using high-fidelity data to infer the effects of the high-frequency unresolved scales. Furthermore, accurate resolution of low-frequency harmonics enable the addition of prior physics knowledge to constrain the data assimilation, without introducing modelling errors from augmentation of empirical turbulence models (as in Chapter 6).

Firstly, it was shown that using the HBNS equations and resolving at least one harmonic ($N \geq 1$) improved the ability of PINN-DA-HBNS to reconstruct flow more accurately, both in terms of skin friction coefficient and mean-flow reconstruction error, than the PINN-DA-Baseline approach, where $N = 0$ harmonics are resolved. In particular, regions of higher gradients (and complexity) were improved. However, as harmonics beyond $N = 1$ are resolved, diminishing returns for reconstruction accuracy are observed, as a smaller portion of the flow is left unresolved and thus inferred by the PINN-DA-HBNS (although this is likely intrinsic to the laminar cylinder flow).

Secondly, the application of PINN-DAs was expanded beyond use for flow reconstruction, as a framework for developing (neural network based) HBNS closure models from sparse high-fidelity velocity measurements using PINN-DA-HBNS was proposed. Furthermore, the work from Rigas & Schmid (2022) was expanded, testing the physical effects of resolved velocity input features and number of resolved harmonics. The resultant model was able to generalise to a higher unseen Reynolds number. Whilst provision of velocity fields was sufficient to build a closure model for $N = 1$ resolved harmonics (due to the reduced influence of non-Markovian

effects), first-order resolved velocity gradients were requisite to build a closure model for $N = 0$, whilst also improving the accuracy of $N = 1$ closure model. Provision of second-order velocity gradients to construct a HBNS closure model were less successful. However, there are two primary considerations for using this approach. Firstly, the accuracy of the closure model is intrinsically linked to the accuracy of the underlying PINN-DA-HBNS reconstruction used to train the model. Secondly, the closure model presented here only captures the divergence-free closure term and thus does not include the potential component. Use of these HBNS closure models may pave the way for wider use of the HBNS for computation of fluids as the effects of unresolved, high-frequency harmonics can be efficiently captured without their expensive computation.

7.4.1 Future Work

Whilst these results validate constraining data assimilation with the HBNS equations for flow reconstruction, application to more complex and turbulent cases is necessary. These cases have energy distributed across a broadband spectrum of harmonics and, unlike the laminar cylinder flow, a quasi-linear model is unlikely to be sufficient.

Similarly, a review of data-driven turbulence models, such as in Section 2.3, suggests that enforcement of physical constraints on the training of the closure model, use of tensor basis, better selection of input features (as in Grabe et al. (2023)), feature normalisation or alternative modelling approaches (e.g. FIML) may improve optimisation of the closure model, increasing its capacity whilst training using a wider range of flow cases and regimes is necessary to maximise generalisability. Furthermore, the use of a non-local closure model approach (Han et al. 2023, Mishra et al. 2015, Parish & Duraisamy 2017), may lead to further improvements in closure model accuracy and generalisability. Finally, testing the HBNS closure model with a computational flow solver is required, to establish whether simulating the truncated HBNS equations can be a viable solution to improving the accuracy of RANS.

Chapter 8

Evaluating PINN-DA formulations to reconstruct a three-dimensional, turbulent wake mean flow from 2D2C sparse experimental data

In this chapter, the use of data assimilation with PINNs is expanded to a practical engineering use case through application on a three-dimensional, turbulent, experimental flow field. Both Chapter 6 and Chapter 7 showed how improved inference of the turbulent scales, through introducing turbulence modelling or resolving the low-frequency scales, improved the capacity of PINNs to reconstruct flow fields. However, these methods were trialled on synthetic flow fields, generated using high-fidelity methods such as DNS, and thus require validation on a realistic engineering case. In the following chapter, this is directly addressed by applying the PINN-DA formulations tested previously to perform data assimilation on a real experimental flow dataset.

The canonical flows that have been used for most research into data assimilation with PINNs (including those presented in previous chapters) are not fully representative of flows that are used in real industrial applications. Using experimental data for flows, such as the axisymmetric

bluff body wake described in Section 4.3 (and used in Mo et al. (2024) to evaluate non-linear auto-encoders), presents a new set of challenges. These flows are often three-dimensional and at a higher Reynolds number (than canonical flows), whilst use of experimental data has its own uncertainty. To evaluate the use of data assimilation with PINNs on a practical engineering case, the following questions are explored:

- i How well do the PINN-DA approaches, which have successfully been applied in the previous chapters, perform at reconstructing the mean flow for a high Reynolds number, three-dimensional, turbulent flow?
- ii What challenges arise when applying PINN-DA to experimental measurements and how can they be solved?

In order to investigate these questions, data assimilation with PINNs will be used to reconstruct the mean flow field for an axisymmetric bluff body wake at $Re \approx 2 \times 10^5$ from sparse time-resolved 2D2C PIV velocity measurements sampled on a single plane. This differs from works, such as Steinfurth & Weiss (2024), who use a high resolution, 3D3C PIV field. An extensive comparison between the three previously discussed PINN formulations (PINN-DA-Baseline, SA and HBNS) in addition to a new PINN-DA-HBNS-SA formulation will be completed, to evaluate if they can be successfully applied to this more complex case and whether the observations from Chapters 6 and 7 are consistent. Primarily, the accuracy of the mean-flow reconstruction from sparse measurements will be considered. Additionally, there will be a brief examination of the reconstruction of the flow field as the PINN-DA extrapolates beyond the single PIV plane. Finally, in order to demonstrate a potential industrial application of PINN-DAs, they will be used to reconstruct the pressure field in the flow domain, where pressure measurements are not available.

The results in this chapter have been divided into two parts. Firstly, by applying PINN-DA to the raw experimental PIV data using an identical approach to previous chapter, it is demonstrated how uncertainty of experimental data, specifically data which has small spatial sensor misalignment errors, affects flow reconstruction performance. In the second part, the

flow reconstruction is repeated using the embedded transformation methodology, proposed and validated in Chapter 5, in order to mitigate the aforementioned effects of spatially misaligned measurements.

The following chapter proceeds as follows. Section 8.1 describes the data sampling and pre-processing for the axisymmetric bluff body wake case, followed by details of the architecture and setup for the PINNs. This is followed by the results in Section 8.2. This contains a discussion on the use of spatially misaligned PIV data for flow reconstruction and the improvements from using the embedded transformation methodology in Section 8.2.1 and an analysis of the flow reconstruction performance across four different formulations of PINNs (PINN-DA-Baseline, PINN-DA-SA, PINN-DA-HBNS, PINN-DA-HBNS-SA). Finally, a conclusion is presented in Section 8.3.

8.1 Numerical setup

In this chapter, PINNs are used to reconstruct an axisymmetric, turbulent bluff body wake from experimental PIV data, in order to test the capacity of PINN-DAs on a realistic engineering problem. This turbulent flow ($Re = 1.88 \times 10^5$), described in Section 4.3, contains a range of flow features (such as boundary layers, separations and recirculations), such that the capacity of PINN-DAs can be fully explored. In this section, how the experimental data is processed and sampling will be described, before detailing the PINN architecture and set up.

8.1.1 Selection of harmonic modes for HBNS constrained data assimilation

As in Chapter 7, use of the PINN-DA-HBNS formulation on the experimental flow requires extraction and sampling of a low-frequency harmonic velocity modes, to train the PINN, as shown in the previous chapter. Whilst use of DMD is suitable for the laminar circular cylinder, this is not the case for the turbulent experimental data. The laminar circular cylinder

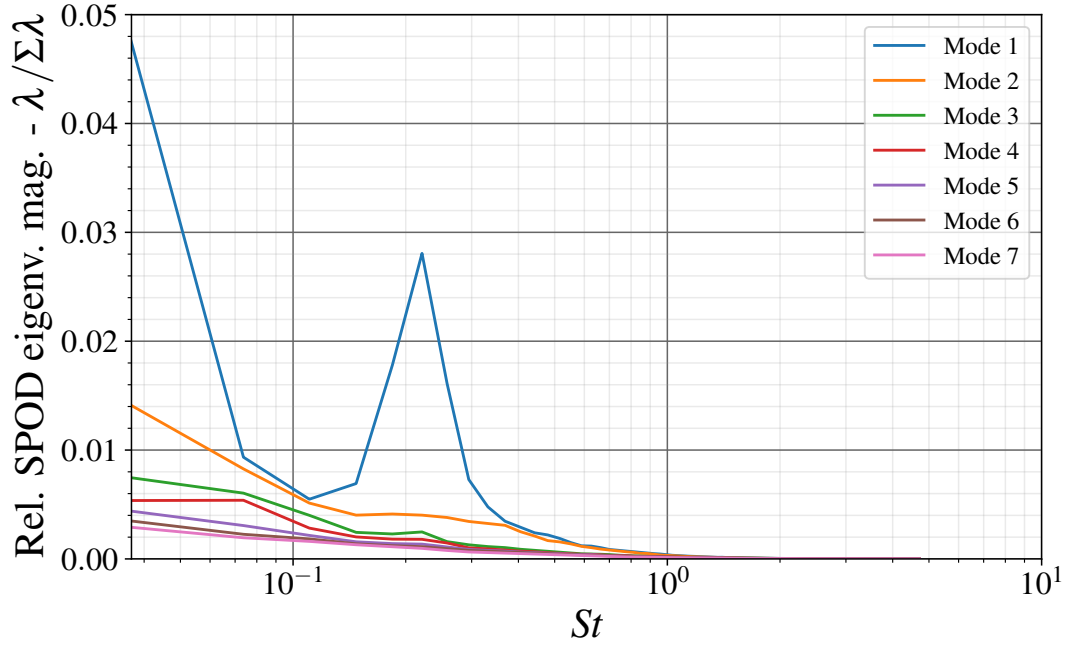


Figure 8.1: Relative magnitude of SPOD mode eigenvalues, performed on time resolved PIV data for the axisymmetric bluff body wake.

reaches a periodic limit-cycle, where energy is focused around some fundamental frequency, ω_0 , corresponding to the fundamental mode of the Fourier series expansion (Eq. (3.10)). For the experimental data, energy is instead distributed around a broadband spectrum of frequencies, as discussed in Franceschini et al. (2021). Therefore, the SPOD method, described in Section 4.3.2, was used to “discretise” the flow data across frequencies and thereafter extract harmonic modes, $\hat{\mathbf{u}}_1$.

Given use of SPOD distributes the energy and modes across a broadband spectrum for the experimental data, one must select the first discrete frequency (and thus first harmonic mode), which will be resolved using the HBNS formulation, that maximises the resolution of the low-frequency turbulent scales. This contrasts the laminar circular cylinder case, where the modes of the periodic limit-cycle correspond to a Fourier series expansion with the fundamental frequency equivalent to the vortex-shedding frequency. The relative eigenvalue magnitude of each SPOD mode for the bluff body wake PIV data are presented in Figure 8.1, showing their relative amplitudes and energies. Analysis of this figure shows there are two frequencies, at which a spike in the modal amplitude and energy occurs, and therefore represent a larger fraction of the low-frequency turbulent scales. Firstly, one can see the spike in magnitude occurring at

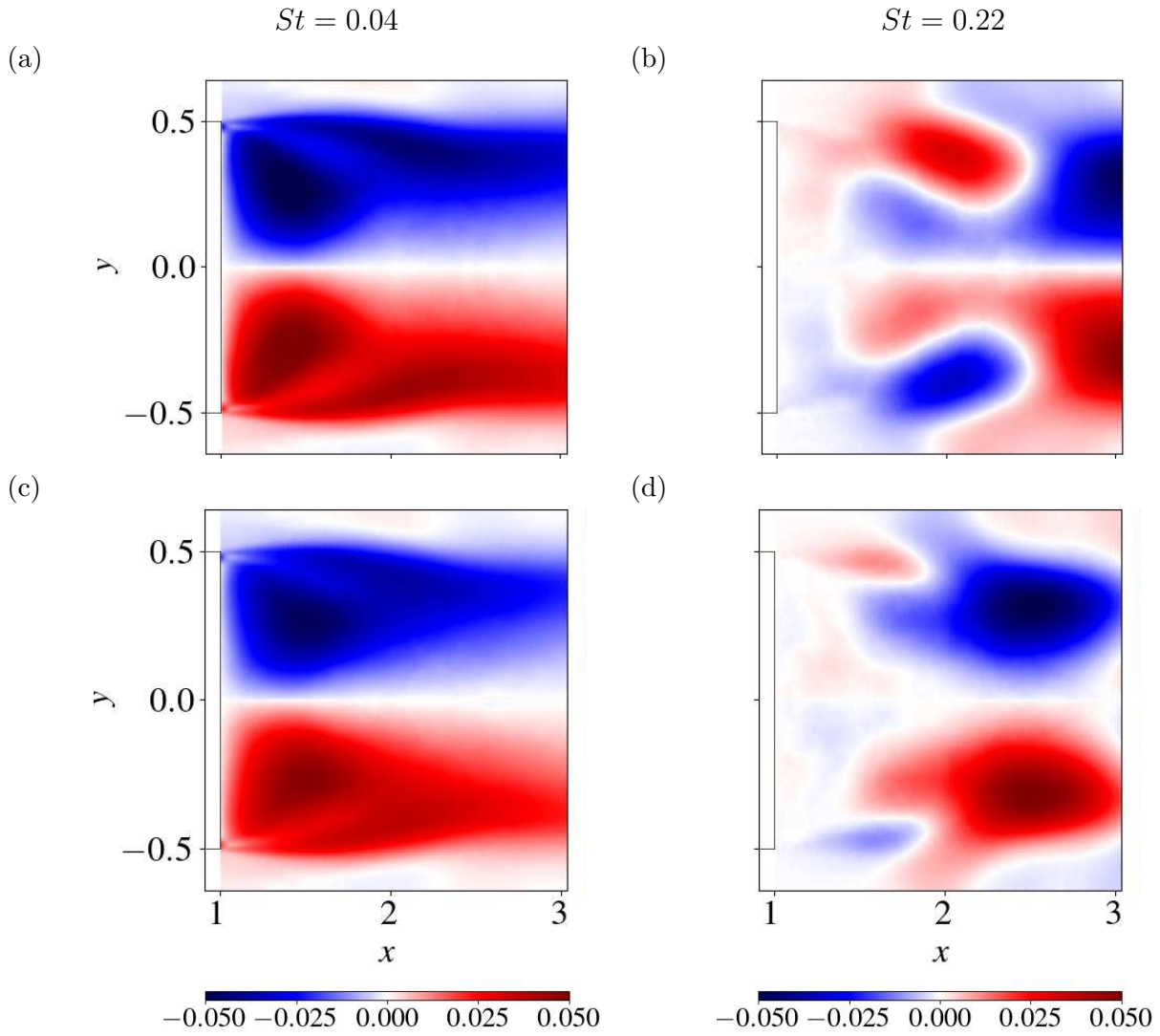


Figure 8.2: Streamwise SPOD velocity modes (real (top) and imaginary (bottom) components), extracted from experimental bluff body wake PIV data, representing two choices of first discrete resolved frequency, for flow reconstruction using PINN-DA-HBNS. (a,c) Low-frequency mode at $St = 0.04$. (b,d) Vortex-shedding frequency $St = 0.22$.

$St = 0.22$, correlating with the vortex-shedding frequency. Selection of this frequency would thus replicate the approach seen in the previous chapter (on the laminar circular cylinder). Secondly, one can observe that at the low frequencies ($St = 0.04$), the SPOD mode has energy with higher relative magnitude than the vortex shedding mode. Resolving these frequencies ($St = 0.04, 0.22$) and SPOD modes with the highest magnitude eigenvalues (and thus energy) using PINN-DA-HBNS (and PINN-DA-HBNS-SA), ensures that the largest amount of low-frequency turbulent scales from a single mode are captured, maximising the effect of resolving scales ($\hat{\mathbf{u}}_1$) on the solution and flow reconstruction.

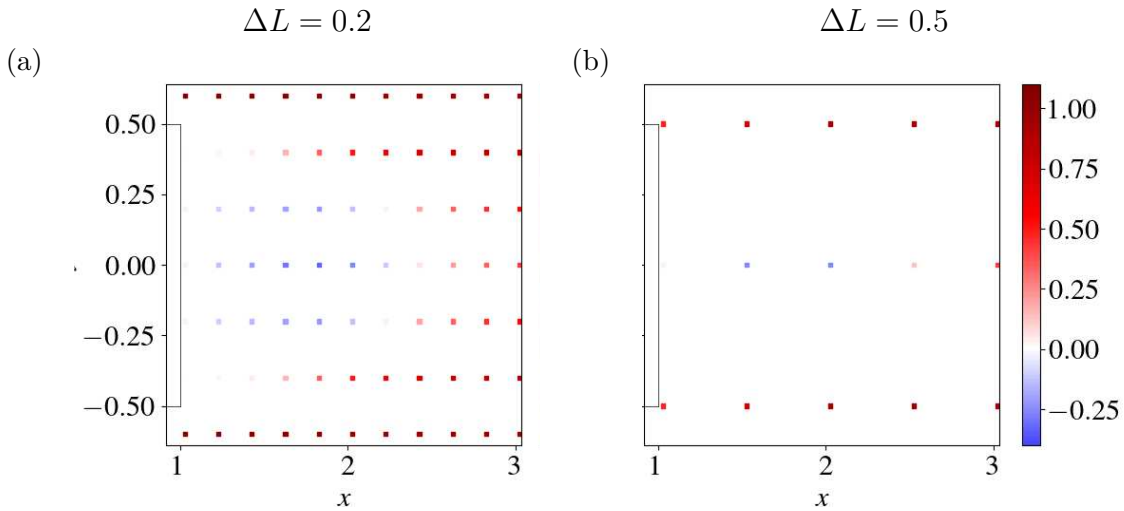


Figure 8.3: Coarse pointwise measurements of the streamwise component of the mean velocity field used for flow reconstruction with PINNs, sampled at two different resolutions.

The streamwise SPOD velocity modes, at the selected resolved frequencies ($St = 0.04, 0.22$), are presented in Figure 8.2, with the magnitude of each mode reaching approximately 5% of free-stream velocity (as calculated using Eq. (4.8b)). Using these SPOD modes, as with the mean flow, pointwise velocity measurements can be sampled, to train the PINN-DA-HBNS (and PINN-DA-HBNS-SA) models. In this chapter, the PINN-DA-HBNS models are trained using the quasi-linear formulation of the HBNS equations, resolving $N = 1$ low-frequency harmonics.

8.1.2 Data sampling

As with all previous chapters, the sparse pointwise velocity measurements, from which the flow is reconstructed, are sampled on a square grid, tested at different resolutions. In this chapter, whilst primarily flow reconstructions using data sampled at $\Delta L = 0.2$ are presented, use of $\Delta L = 0.5$ spacing is also discussed (with additional results at this spacing presented in Appendix D.2). The coarse pointwise data sampling is presented in Figure 8.3. Even at the finer resolution ($\Delta L = 0.2$), in Figure 8.3(a), many flow details have been lost. This becomes more apparent at the coarser resolution, at which becomes impossible to define the shape of the recirculations in the bluff body wake. It should be noted, as in previous chapters, whilst data sparsity is represented with pointwise measurements extracted from the high resolution PIV data, a real sparse PIV measurements would represent some average across the pixel area

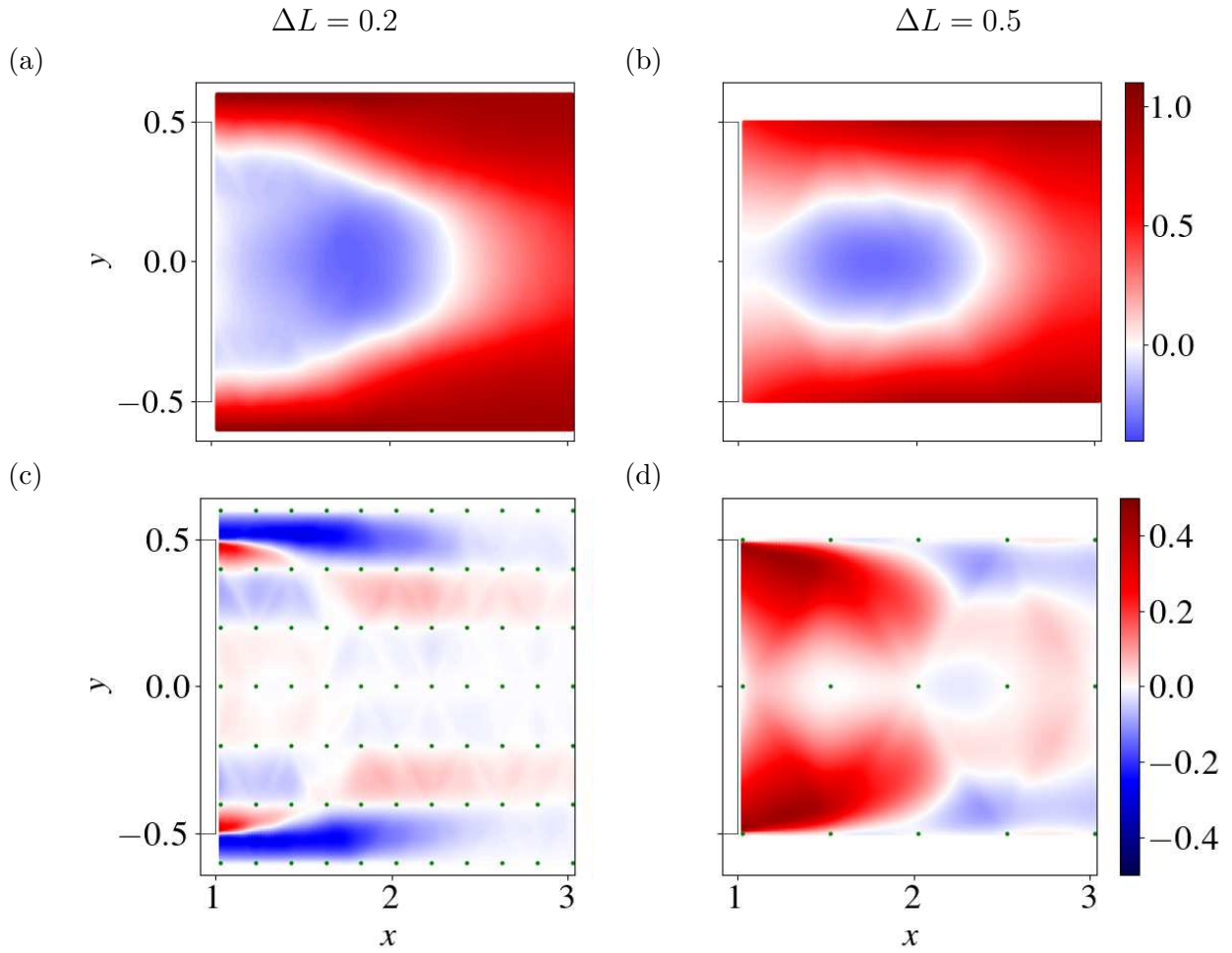


Figure 8.4: Cubic spline interpolation of streamwise PIV velocity data from two different data sampling resolutions ($\Delta L = 0.2, 0.5$). (Top) cubic spline interpolated reconstructions. (Bottom) reconstruction error.

(instead of pointwise sampling). These sparse measurement samples show the significance of the loss in information when using resolution limited sampling and highlights the advantages of using data assimilation for flow reconstruction.

8.1.3 Cubic spline interpolation

Whilst this work has explored the use of data assimilation with physics constraints to reconstruct flow fields from data, it can be useful to compare this approach with a cubic spline interpolation, presented in Figure 8.4, to examine the advantage data assimilation has over this interpolation (cubic spline) approach without using any physics constraints. The cubic spline interpolation highlights the key limitations of this approach and further motivates the

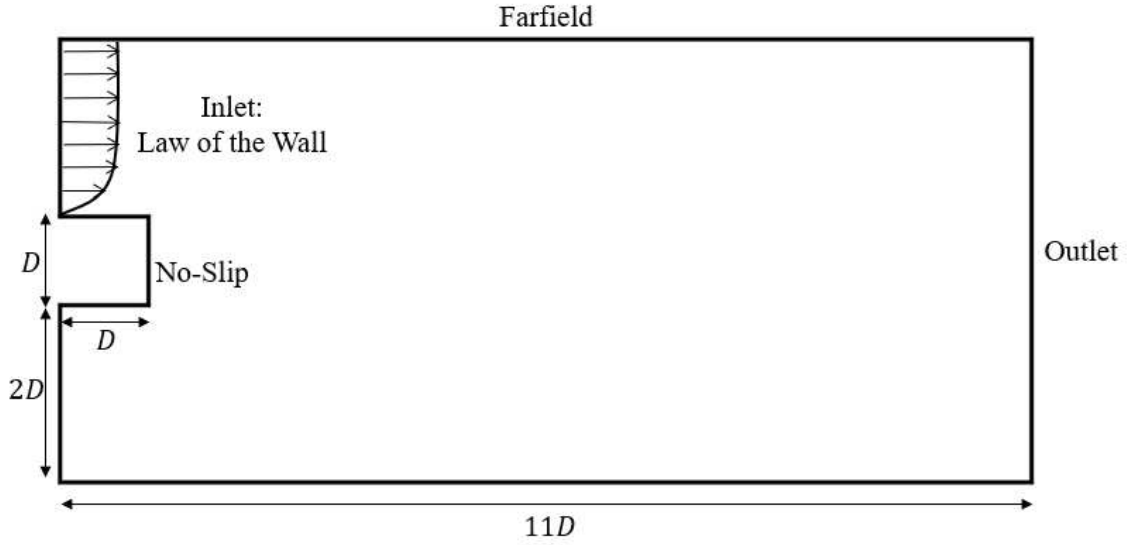
use of PINN-DA. Although interpolating from $\Delta L = 0.2$ data spacing seems to moderately resemble the true flow, Figure 8.4(a) shows that the cubic spline interpolation is not physical. This is demonstrated in the error field (Figure 8.4(c)), with large errors across the domain, especially in the shear layer. These problems become more apparent as data spacing becomes more coarse at $\Delta L = 0.5$, presented in Figure 8.4(b,d), where the cubic spline interpolation is highly inaccurate. These interpolations also demonstrate how the cubic spline based method is unable to extrapolate reliably outside of the data limits, such as in Figure 8.4(b,d), where the interpolation ends at the limits of the data bounds. As a consequence, a cubic spline interpolation would be unable to extract the full three-dimensional flow field from 2D2C velocity data. Similarly, a cubic spline interpolation, is unable to infer additional flow fields, such as pressure or the residual forcing terms, and thus other methods would be required for field inference. These significant issues present when using a cubic spline interpolation (poor reconstruction accuracy, lack of adherence to physics, limited extrapolation capability and inability to infer additional fields) are directly addressed by the PINN-DA formulation, which through adherence of governing laws and defining domain edges can, at least to some degree, extrapolate outside of data and infer new fields.

8.1.4 PINN Setup

Data assimilation of the bluff body wake flow using PINNs was completed on the truncated domain presented in Figure 8.5(a). This three-dimensional domain was defined with an inlet $1D$ upstream from the trailing edge of the bluff body and extended $10D$ downstream from it. The farfield boundary was defined with a cylinder of radius of $2.5D$. A velocity inlet was applied using a law of the wall velocity profile, whilst at the outlet a typical pressure outlet was enforced. Finally, free-stream conditions were enforced at the farfield and a no-slip condition at the bluff body walls.

To train the PINN-DA, boundary conditions and governing laws were enforced at randomly sampled points in the domain, as presented in Figure 8.5(b). The governing equations were enforced at 1250 collocation points per mini-batch whilst a further 640 boundary points (for

(a)



(b)

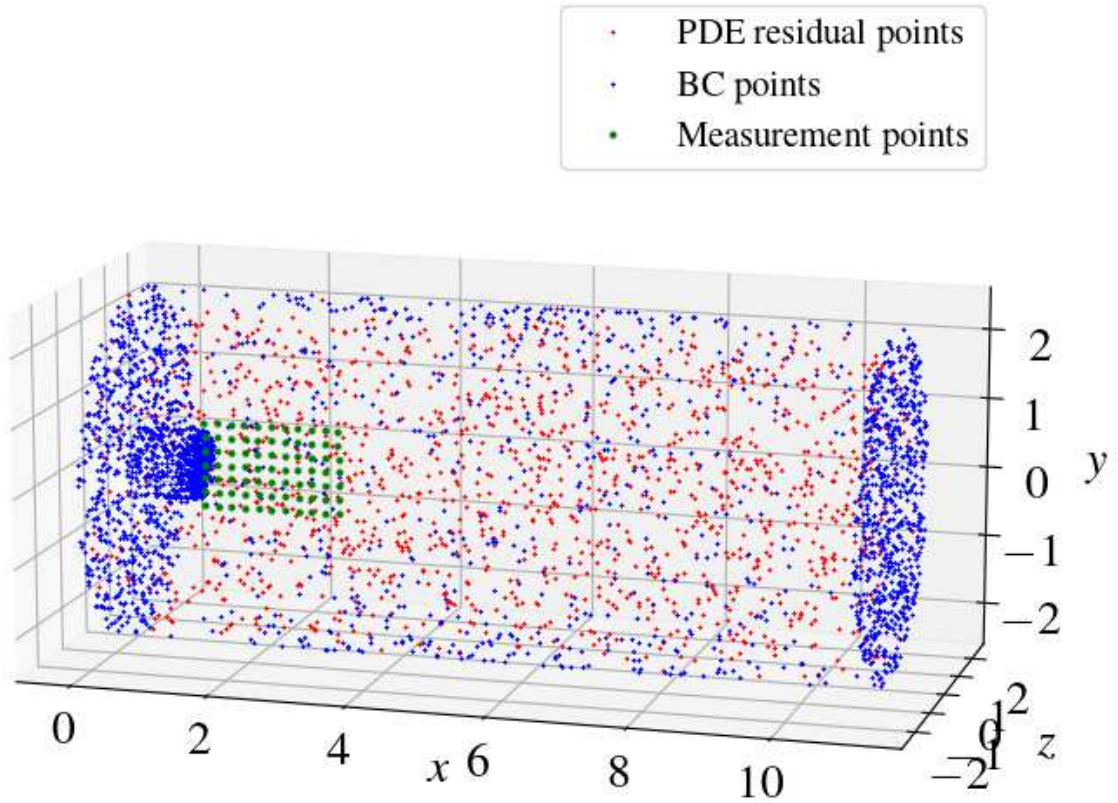


Figure 8.5: Optimisation domain used to train the PINNs for flow reconstruction of the bluff body wake from sparse PIV velocity measurements. (a) Truncated domain for reconstruction. (b) Three-dimensional domain showing the data sampling (at $\Delta L = 0.2$) and a single sampled mini-batch of collocation and boundary points.

each boundary condition) per mini-batch were used to enforce boundary conditions. In every mini-batch, these points were resampled with a uniform random distribution, with one such

sample shown in Figure 8.5(b). In every iteration of training, 2000 mini-batches were used. As in Chapter 7, an SDF spatial weighting was applied at the collocation points to the PDE loss, due to improved convergence. Figure 8.5(b) also shows the sparse measurement data points (sampled at $\Delta L = 0.2$) lying on the two-dimensional 2D2C PIV plane.

The PINN-DAs were constructed and trained within the NVIDIA Modulus framework (NVIDIA 2024), using a similar architecture to Chapter 7. Separate, parallel neural network were used for a) the velocity terms (mean velocity and first harmonic velocity components for PINN-DA-HBNS/HBNS-SA); b) pressure terms (mean combined pressure-potential forcing field and first harmonic combined pressure-potential forcing for PINN-DA-HBNS/HBNS-SA); c) residual corrective (solenoidal) forcing components ($\mathbf{R}_{n,s}$); and d) the eddy viscosity for the PINN-DA-SA/HBNS-SA approaches (where corrective forcing is decomposed into an eddy-viscosity and non eddy-viscosity part. Also note that for first resolved harmonic terms, a separate output is used for real and imaginary components (as before). Each network contained 9 layers, each with 512 nodes per layer, with each node using an adaptive swish activation function (as discussed in Chapter 7). The PINN-DA models were trained using the ADAM optimiser for 200,000 iterations, using an initial learning rate of 5×10^{-4} and a learning rate decay of 0.95. The PINNs were trained using a single NVIDIA RTX 6000 GPU.

8.2 Results

In this section, the reconstruction of an axisymmetric, turbulent, bluff body wake flow from sparse experimental PIV data using PINN-DAs is presented. Four different PINN approaches (DA-Baseline, DA-SA, DA-HBNS resolved at either $St = 0.04, 0.22$, DA-HBNS-SA resolved at either $St = 0.04, 0.22$) are evaluated. These can be summarised by comparing the Reynolds forcing vector, $\mathbf{f}$ for the different formulations:

- PINN-DA-Baseline: $\mathbf{f} = \nabla\phi_0 + \mathbf{R}_{0,s}$
- PINN-DA-SA: $\mathbf{f} = \nabla\phi_0 + \mathbf{R}_{0,s} + \nabla(\nu_t \nabla_{\mathbf{s}} \hat{\mathbf{u}}_0)$

- PINN-DA-HBNS: $\mathbf{f} = \nabla\phi_0 + \mathbf{R}_{0,s} + \overline{\hat{\mathbf{u}}_1} \cdot \nabla\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_1 \cdot \nabla\overline{\hat{\mathbf{u}}_1}$
- PINN-DA-HBNS-SA: $\mathbf{f} = \nabla\phi_0 + \mathbf{R}_{0,s} + \nabla(\nu_t \nabla_S \hat{\mathbf{u}}_0) + \overline{\hat{\mathbf{u}}_1} \cdot \nabla\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_1 \cdot \nabla\overline{\hat{\mathbf{u}}_1}$

Firstly, it is demonstrated that whilst PINN-DA can be successfully to reconstruct a turbulent flow, use of experimental data presents a key challenge: sensor miscalibration leading to spatial misalignment. The effects of this are mitigated using the embedded transformation approach, proposed and validated in Chapter 5. Secondly, DA-SA, DA-HBNS and DA-HBNS-SA are compared with DA-Baseline, showing improved reconstruction capacity and thus validating the previous findings from Chapters 6 and 7 on a more complex case. Finally, it is demonstrated that use of PINN-DA methods enables the reconstruction of the full pressure field, away from possible measurement locations.

Mean-flow reconstruction accuracy is evaluated using two primary metrics: an L_2 based and L_∞ based error. These are defined using the following:

$$\varepsilon_2 = \frac{1}{N_P} \sum_{p=0}^{N_P} \|\tilde{\mathbf{u}}_{0,p} - \hat{\mathbf{u}}_{0,p}^{PIV}\|^2, \quad (8.1a)$$

$$\varepsilon_\infty = \max \|\tilde{\mathbf{u}}_0 - \hat{\mathbf{u}}_0^{PIV}\|. \quad (8.1b)$$

The ε_2 error describes the total reconstruction accuracy, whilst ε_∞ highlights the maximum error and are evaluated over the PIV plane.

8.2.1 Use of transformation embedded PINN-DA on spatially misaligned experimental measurement data

In this section, it will be shown how using data assimilation to reconstruct the flow from the experimental data poses a challenge, if there is sensor miscalibration, which leads to introduction of spatial misalignment bias to the underlying enforced measurement data. The effect of these problems and subsequent application of the embedded transformation approach (see Chapter 5) to mitigate them are then presented.

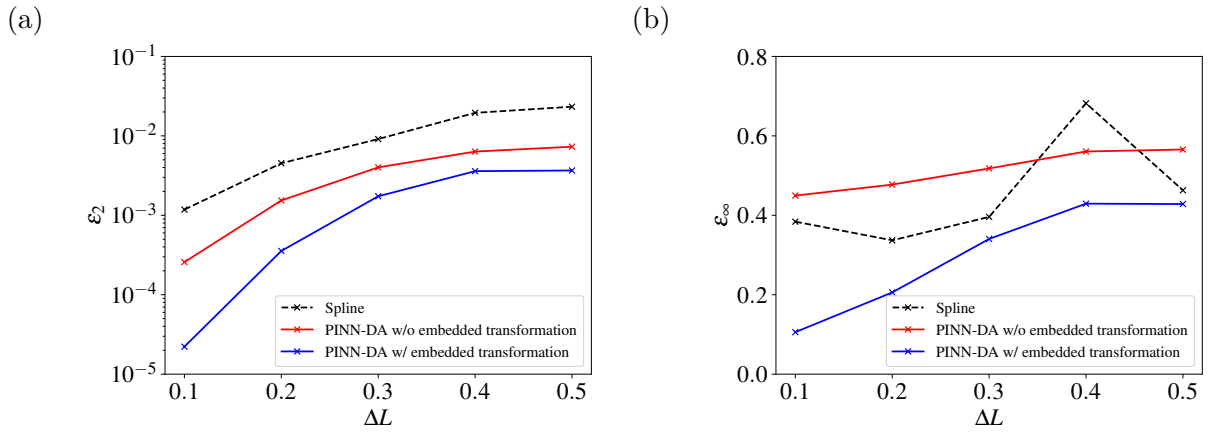


Figure 8.6: Reconstruction error averaged across all the PINN-DA formulations with and without the use of the embedded transformation approach for mitigating spatial measurement misalignment across different data resolutions. (a) L_2 -based error and (b) L_∞ -based error.

A comparison of the mean-flow reconstruction error for PINN-DA with the use of the embedded transformation approach (as proposed in Chapter 5) and without its use (as in Chapters 6 and 7) across a range of equidistant data-spacings is presented in Figure 8.6. Additionally included is the reconstruction accuracy of the cubic spline interpolation, as previously presented in Figure 8.4 (here error is only evaluated within the bounds of the measurement data due to its inability to extrapolate outside of this region). Irrespective of use of the embedded transformation to correct measurements, one can reconstruct the mean flow more accurately (ε_2 reconstruction error) using PINN-DA than with a cubic spline interpolation. By constraining data assimilation with known physics (governing equations), one can use data assimilation to reconstruct the mean flow to a higher accuracy than a typical interpolation method, for a turbulent, three-dimensional flow using sparse 2D2C experimental PIV data.

Use of the embedded transformation approach, to correct for spatial sensor misalignment, results in a reduction in mean velocity reconstruction error (Figure 8.6(a)) relative to PINN-DA without it, showing how use of spatially-misaligned, biased measurements can have a significant effect on flow reconstruction accuracy with data assimilation. This is most apparent at the finest data resolutions before the gap between approaches narrows as data resolution coarsens. This indicates that as the amount of data reduces, the embedded transformation approach is less able to capture discrepancies between the data and the enforced governing equations and boundary conditions. The effect of the embedded transformation is most apparent in the reduction in

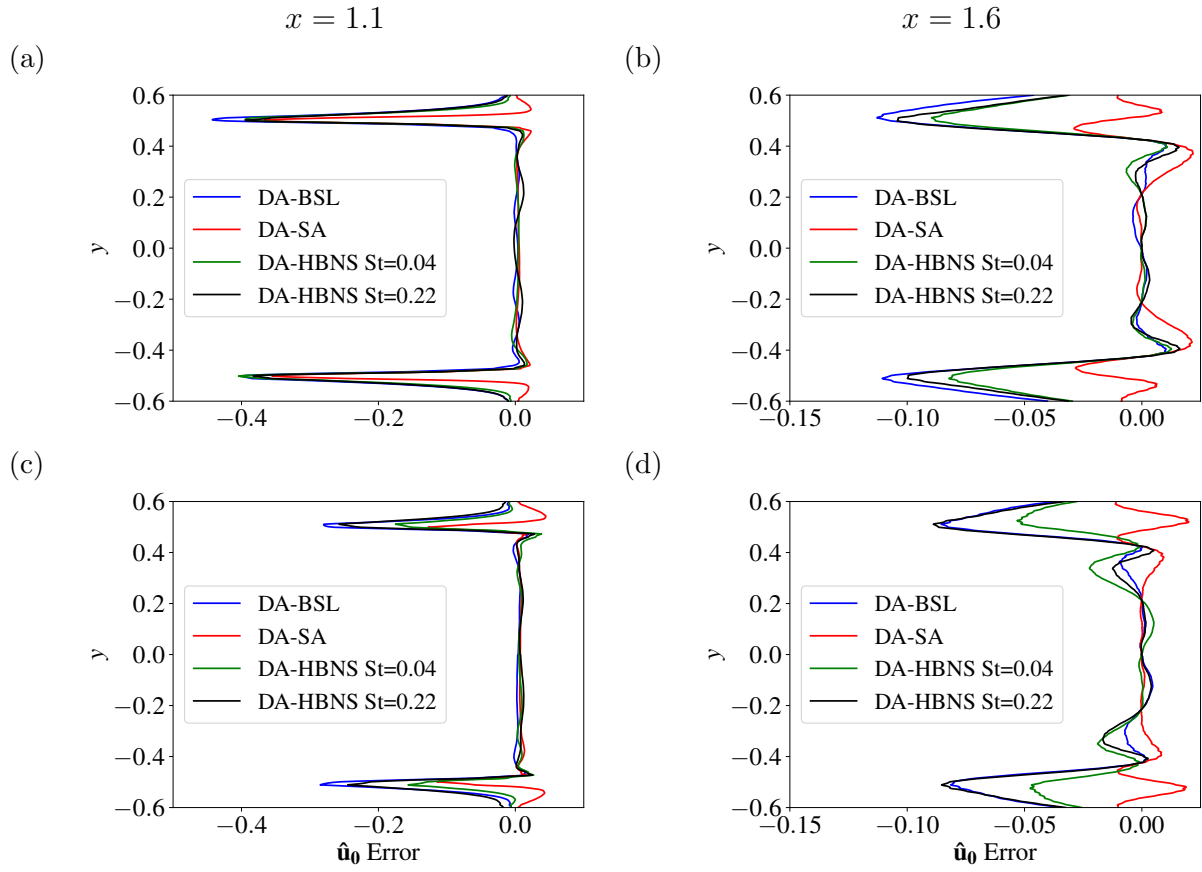


Figure 8.7: Reconstructed streamwise mean velocity error profiles for different PINN-DA formulations, evaluated at two different streamwise locations. (Top) without transformation embedded PINN-DA. (Bottom) with transformation embedded PINN-DA.

maximum error, shown in Figure 8.6(b). At the finest resolution, this reduction approaches up to approximately 35% free-stream. The reduction in mean velocity reconstruction error from use of the embedded transformation approach indicates there is a source of spatial misalignment error affecting the underlying PIV data, highlighting the challenges of flow reconstruction from experimental data, whilst also demonstrating how use of correction techniques can mitigate these misalignment effects for data assimilation methods.

Source of spatial misalignment from experimental data

The mean velocity error profiles for PINN-DA without embedded transformation using data spacing $\Delta L = 0.2$, presented in 8.7(a,b), show the underlying source and effect of spatial measurement misalignment, such as the high values of ε_∞ . At the station closest to the wall ($x = 1.1$), (Figure 8.7(a)), there is a “spike” in error centred around $y = 0.5$, the height of the

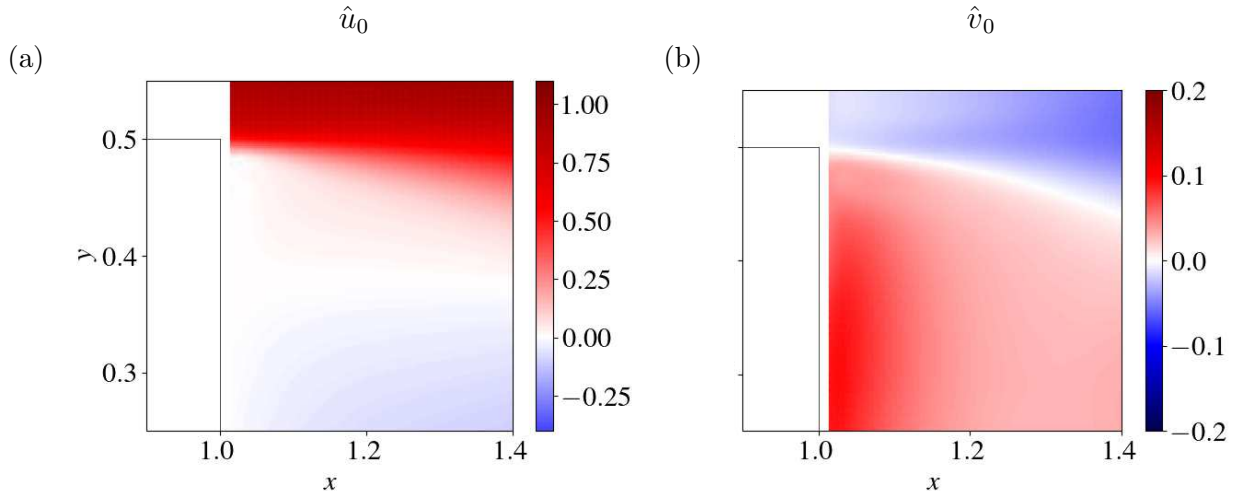


Figure 8.8: True PIV field in the vicinity of the bluff body trailing edge, where reconstruction error is highest. This shows the spatial misalignment between the PIV field and the trailing edge (and thus boundary conditions).

bluff body trailing edge, which is consistent across PINN-DA formulations. However, further downstream, at $x = 1.6$ (Figure 8.7(b)), the error decays and one can observe divergence between approaches. Although some of the error in the immediate vicinity of the trailing edge must be attributed to underlying reconstruction error of the PINN-DA methodology, the high magnitude and consistency across formulations suggests that there is a fundamental problem with the underlying experimental PIV data, used to constrain the PINN-DA optimisation.

The key spatial measurement misalignment (of the PIV plane) can be found through examination of the measured PIV velocity fields in the region around the bluff body trailing edge, presented in Figure 8.8, highlighting the source of the error observed in Figure 8.7(a). The streamwise mean velocity component, in Figure 8.8(a), demonstrates how the velocity profile of the shear layer from the PIV data does not align well with the trailing edge of the bluff body. Furthermore, despite approaching a no-slip boundary at the bluff body backplate (i.e. $x = 1.0$), the mean vertical velocity component, in Figure 8.8(b), has not yet begun to decrease to 0. Whilst the velocity may decrease sharply in the gap between the PIV plane (starting at $x = 1.0145$) and the wall, there is also a high velocity reconstruction error across all PINN-DA formulations, indicating another possible source of spatial misalignment in the experimental data. The highlighted spatial misalignments can be attributed to miscalibration of the PIV plane and thus mislabelling of the experimental data measurement locations. Whilst this sys-

tematic measurement error, caused by small spatial misalignment of the PIV plane, is small and does not affect qualitative analysis of the PIV velocity fields, for flow reconstruction using constrained data assimilation, it can have a significant effect, especially where mismatches in boundary conditions and high flow gradients occur (as seen in Figure 8.7(a)).

The improvement in reconstruction accuracy, through use of the transformation embedded PINN-DAs, as previously shown in Figure 8.6, demonstrates both the significance of spatial miscalibration of the PIV plane, but also how the proposed approach can identify and correct experimental spatial measurement misalignment. Embedding PINNs with an transformation matrix (as seen in Chapter 5) leverages the knowledge that the measurement data must obey both the governing equations and boundary conditions, to learn parameters of the transformation matrix, spatially correcting the measurements (if required), as demonstrated in the following section.

Effect of applying embedded transformation matrix on measured coordinates

Figure 8.9, show how three different PINN-DA formulations (PINN-DA-Baseline, PINN-DA-SA and PINN-DA-HBNS), learn and apply spatial corrections to the measurement locations (using data spacing $\Delta L = 0.2$), through use of transformation embedded PINNs. Across all PINN-DA formulations, the learnt transformation is qualitatively consistent, with application of a small translation in the positive x -direction (away from the wall) and a small percentage stretch in the y -direction (approximately 2.5%), compared with the nominally measured locations. This transformation of the PIV data correlates with the expected spatial misalignment, discussed in Section 8.2.1 (see Figure 8.8), as the stretching better aligns the shear layer to the trailing edge and the translation increases space for the velocity to reduce to zero near the back plate.

Despite the qualitative similarity in learnt transformations between different PINN-DA formulations, there are quantitative differences, seen in the learnt translation. PINN-DA-Baseline translates the measurements the largest amount (0.18), followed by PINN-DA-HBNS (0.09) and finally PINN-DA-SA learning the smallest translation (0.07). This behaviour is consistent with the specific formulations of the governing equations and results in previous chapters. The

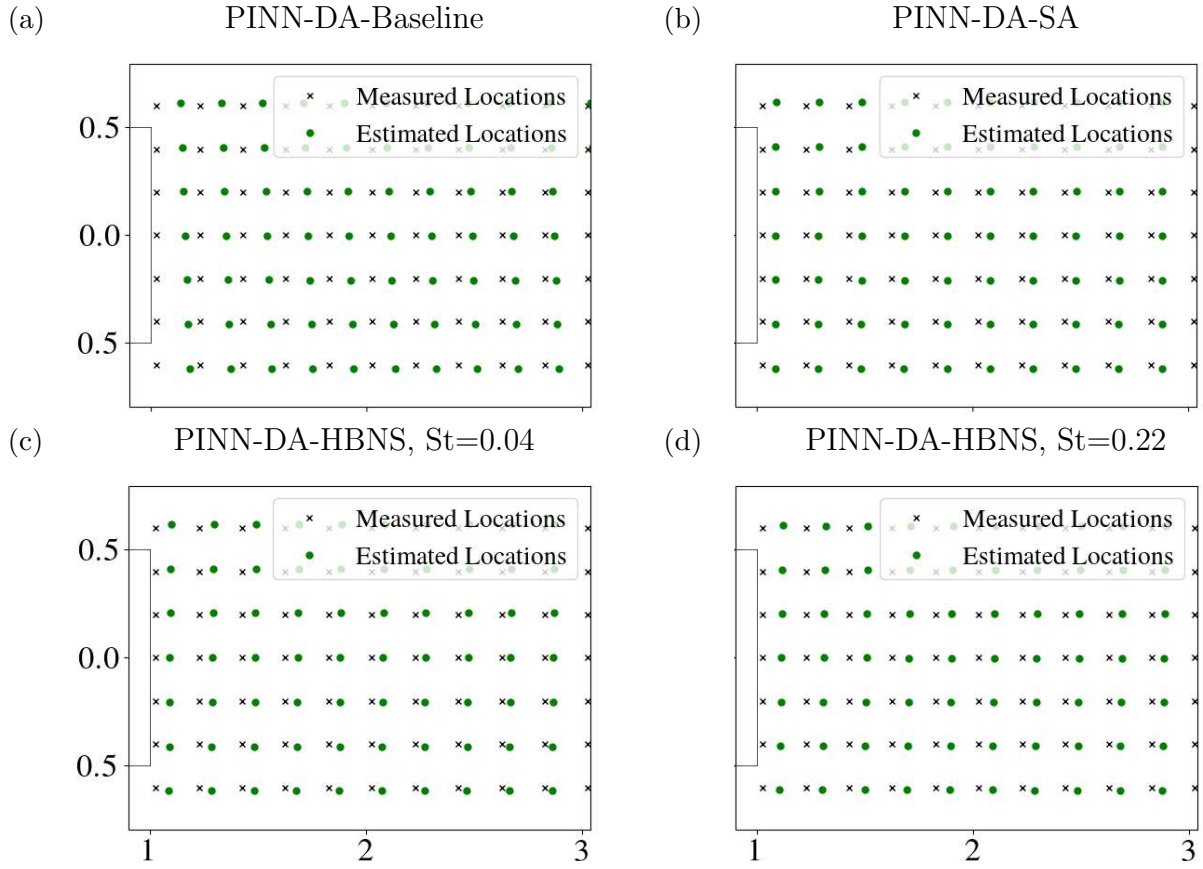


Figure 8.9: Nominal and transformed measurement coordinates showing the effect of using embedded transformation (restricted to the xy plane) during training of the PINNs when data spacing is $\Delta L = 0.2$.

PINN-DA-Baseline formulation overpredicts the thickness of boundary layers (see Chapter 6), which is corrected by use of turbulence model augmentation. It follows that, in order to minimise PDE loss, the learnt spatial corrections to the measurement data, observes the same trend. Consequently the translation correction of the measurement data for the PINN-DA-Baseline formulation is overpredicted relative to the PINN-DA-SA translation correction, as observed in Figures 8.9(a,b) (whilst a similar argument can be made for PINN-DA-HBNS). This comparison shows that whilst the learnt spatial corrections to measurements are consistent and robust across formulations, they are sensitive to the choice of PINN-DA formulation, as enforcement of different governing equations is satisfied by slightly different corrections.

As seen in Figure 8.6, as data coarsens, the reduction in reconstruction error from the embedded transformation approach diminishes. This indicates that accurate learning of corrections for spatially misaligned data is dependent on sufficient measurements being provided, as shown

using $\Delta L = 0.5$ in Appendix D.2.

Finally, the effect of correcting spatially misaligned measurements with the embedded transformation approach, is demonstrated with the streamwise velocity error profiles presented in Figure 8.7(c,d). Examining the error profiles closest to the trailing edge (at $x = 1.1$), shows how applying spatial corrections to measurements significantly reduces maximum error magnitude. The maximum error reduces from approximately 40% of free-stream when not using the embedded transformation (Figure 8.7(a)) to between 15-30%, when the embedded transformation is used (Figure 8.7(c)), depending on PINN-DA formulation. These improvements propagate downstream ($x = 1.6$), as shown in Figure 8.7(b,d), although the reduction in reconstruction error is smaller. Comparing the mean velocity error profiles, with and without embedded transformations, emphasise how use of miscalibrated, spatially misaligned measurements affect flow reconstruction and the effectiveness of the proposed correction method, for mitigating these effects and improving reconstruction accuracy of data assimilation methods.

8.2.2 Flow Reconstruction using embedded transformation for spatially misaligned data

In the following section, the reconstruction of the turbulent bluff body wake using transformation embedded PINN-DAs are evaluated. Firstly, flow reconstruction using the three PINN-DA formulations (Baseline, SA, HBNS) are presented and compared. This includes a breakdown of the results showing how different PINN-DA formulations affect reconstruction accuracy and how this is caused by difference in inferred forcing. Secondly, PINN-DA-HBNS-SA, where turbulence model augmentation is combined with HBNS equations is tested (as presented in Eq. (3.27)). Finally, use of PINN-DA to reconstruct the bulk pressure field away from the wall, where it can not typically be measured, is presented. In this section, the reconstructed velocity and error fields are presented in the measured, non-transformed coordinate space i.e. $\mathbf{x}_m$, ensuring comparability between the PINN-DA approaches with and without embedded transformation.

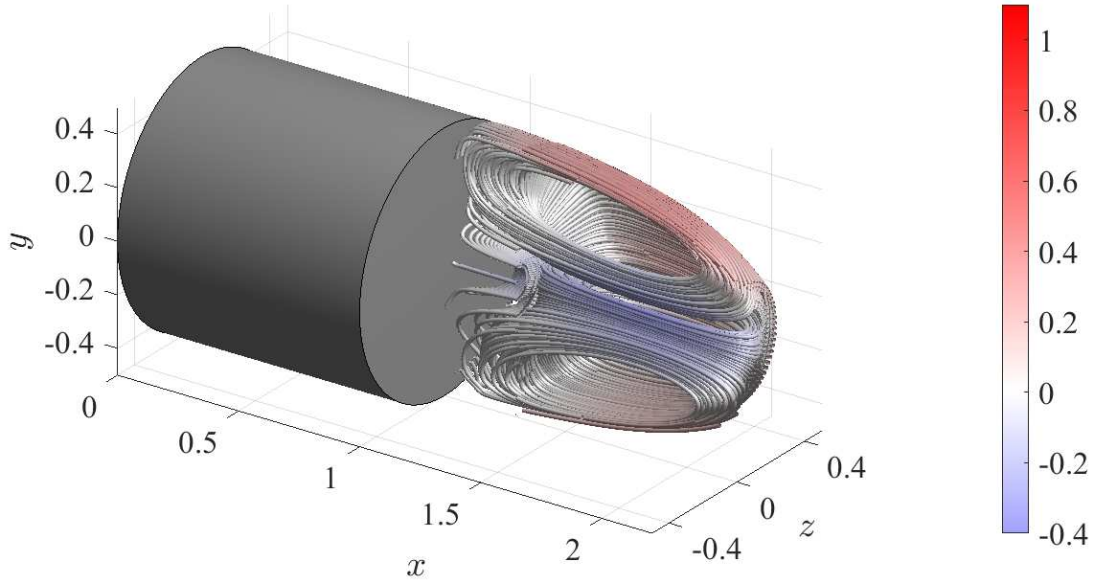


Figure 8.10: 3D velocity streamlines (coloured with streamwise mean velocity, as in Pavia et al. (2019)) for mean-flow reconstruction using PINN-DA-Baseline trained at data spacing $\Delta L = 0.2$, showing the annular recirculation of the full 3D flow.

Three-dimensional flow reconstruction capacity using PINN-DA

The PINN-DA framework is able to extrapolate data from sparse, planar 2D2C PIV data, in order to reconstruct the three-dimensional flow, as presented in Figure 8.10. In this streamline flow field, one can demonstrate flow features documented in existing literature (Gentile et al. 2016, Pavia et al. 2019, Rigas et al. 2014). Primarily, one can see how the flow separates off the trailing edge of the bluff body before developing into an annular recirculation vortex structure. The reconstructed streamline flow demonstrates how using two-dimensional data alone, PINN-DA (here PINN-DA-Baseline) is able to qualitatively reproduce key, physical three-dimensional flow features.

Flow reconstruction using PINN-DA-Baseline and PINN-DA-HBNS

The mean velocity reconstruction and error fields for PINN-DA-Baseline and PINN-DA-HBNS (resolved at either $St = 0.04, 0.22$) with data spacing $\Delta L = 0.2$, is presented in Figure 8.11. Whilst all PINN-DA approaches predict the shape of the wake and recirculation length well, the largest error region, irrespective of formulation, is observed in the shear layer, as the flow

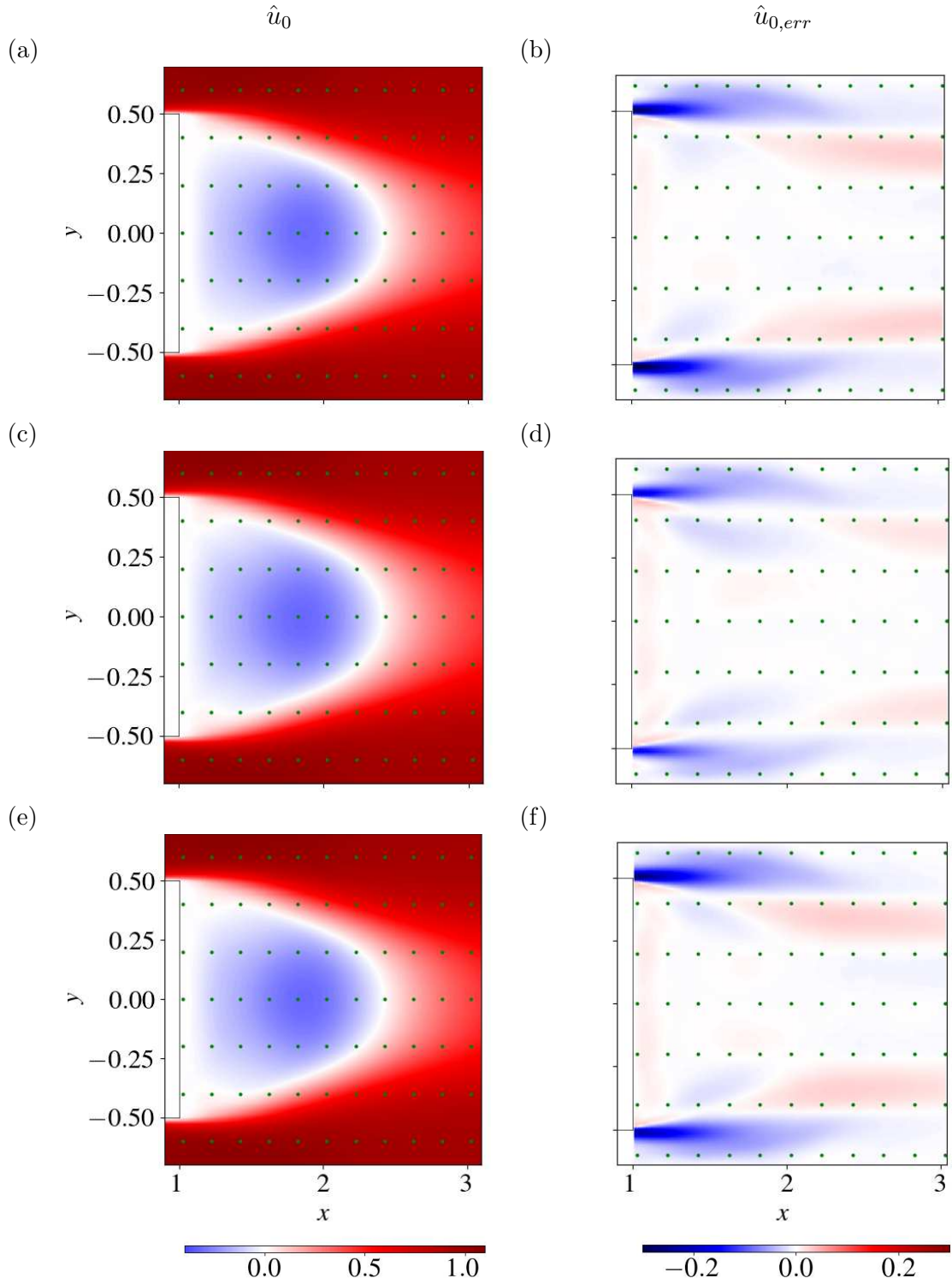


Figure 8.11: PINN reconstruction using different methodologies for data spacing $\Delta L = 0.2$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-HBNS resolved at $St=0.22$

separates off the trailing edge of the body (as seen previously in Figure 8.7). Given that this separated region is both the largest source of reconstruction error and also the region, where correcting spatially misaligned measurements have the most significant effect, this result

stresses the importance of using the embedded transformation approach. Even after applying measurement corrections to enable better enforcement of the governing equations, boundary conditions and data loss, there is still a significant error in the shear layer (albeit with a reduced magnitude). Irrespective of PINN-DA formulation, this result shows that PINN-DA is able to reconstruct the mean flow well, and that the error topology is qualitatively similar across formulation.

Figure 8.11(c-f) shows the effects of resolving a single low-frequency harmonic scale, by constraining PINNs with the HBNS equations and (SPOD-determined) harmonic velocity modes, in addition to demonstrating how the choice of resolved frequency ($St = 0.04$ (c,d), $St = 0.22$ (e,f)) impacts reconstruction accuracy of the bluff body wake mean flow. By using PINN-DA-HBNS to reconstruct the flow, resolved at the low-frequency, $St = 0.04$ (Figure 8.11(c,d)), one can observe improvements in the mean-flow reconstruction at the initial separation and the shear layer relative to PINN-DA-Baseline ($\varepsilon_2 = 8.5 \times 10^{-4}$ for PINN-DA-Baseline compared with $\varepsilon_2 = 3.15 \times 10^{-4}$ for PINN-DA-HBNS, resolved at $St = 0.04$). This is in contrast to the PINN-DA-HBNS flow reconstruction resolved at the vortex-shedding frequency, $St = 0.22$ (Figure 8.11(e,f)). In this case, the mean-flow reconstruction has marginally smaller error than PINN-DA-Baseline ($\varepsilon_2 = 7.08 \times 10^{-4}$). It should also be mentioned that resolving a low-frequency harmonic (Fourier) mode enables the prediction of some unsteadiness in the flow. The mean velocity reconstruction using PINN-DA-HBNS, shows that resolution of a low-frequency harmonic scale (extracted with SPOD), does improve the reconstruction accuracy of the more complex bluff body wake, but that i) the effect initially appears smaller than in Chapter 7, and ii) the relative improvement is dependant on choice of resolved frequency.

The resolved forcing components ($\mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1) = \overline{\hat{\mathbf{u}}_1} \cdot \nabla \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1}$), presented in Figure 8.12, show the resolved low-frequency turbulent scales affect the mean-flow equations and subsequently the PINN-DA-HBNS reconstruction accuracy. In Figure 8.11(c,d), it was shown how resolving the flow at $St = 0.04$ resulted in a reduction of reconstruction error in the shear layer near the trailing edge of the bluff body. The corresponding resolved forcing component highlights the source of this improvement in Figure 8.12(a,b). Resolving the turbulent scales at $St = 0.04$ (as captured in the first SPOD mode), results in resolved forcing field, primarily act-

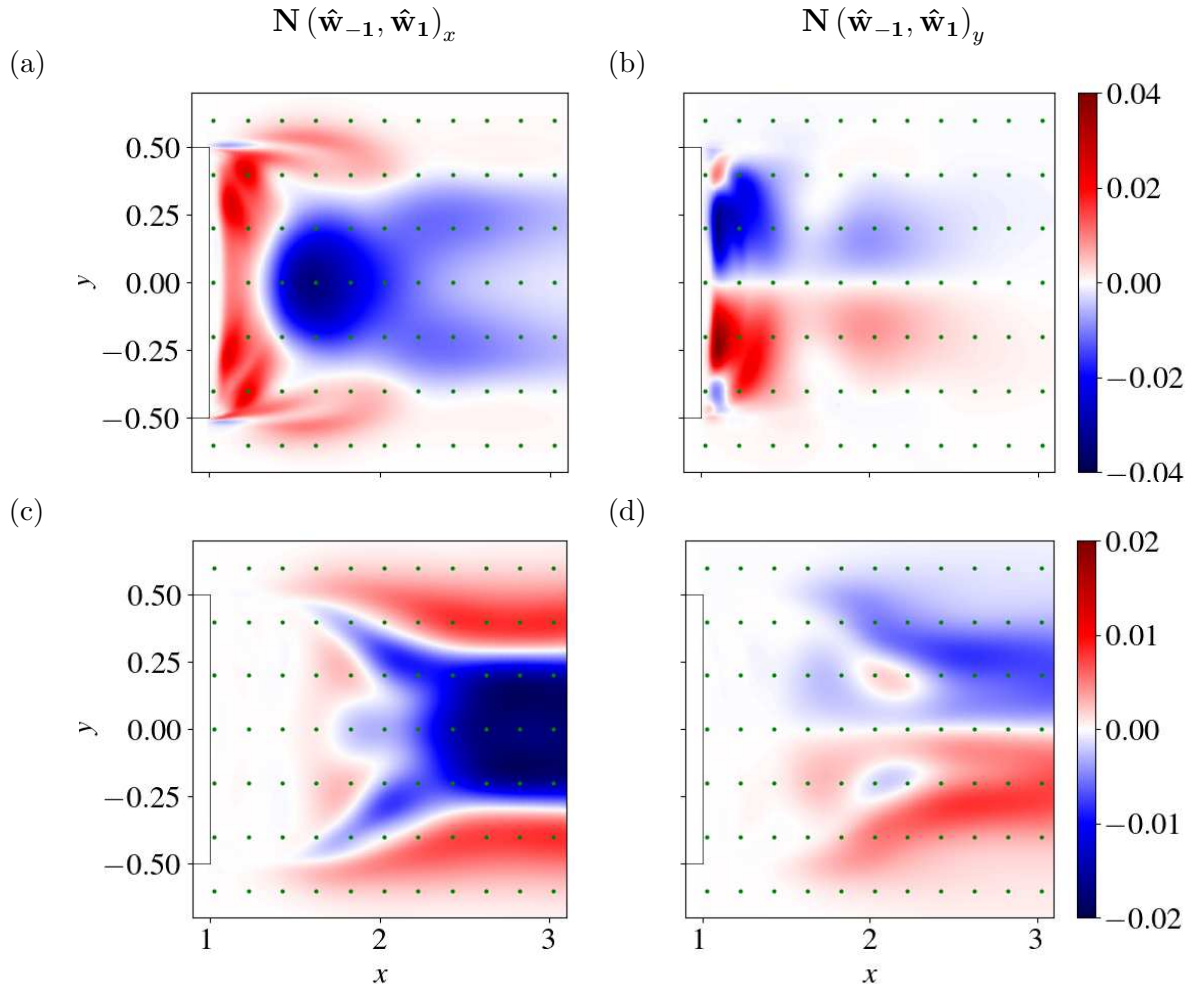


Figure 8.12: Resolved forcing components ((left) streamwise, (right) vertical), $\mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1) = \overline{\hat{\mathbf{u}}_1 \cdot \nabla \hat{\mathbf{u}}_1} + \hat{\mathbf{u}}_1 \cdot \nabla \hat{\mathbf{u}}_1$, from PINN-DA-HBNS flow reconstructions, resolved at two different low-frequency harmonic scales. (a,b) - $St = 0.04$, (c,d) - $St = 0.22$

ing in the initial separation and into the shear layer. There are also higher magnitude resolved forcing features around the wall and into the recirculation bubble ($1.0 \leq x \leq 2.0$). Similarly, there is a correlation between the more limited improvements in reconstruction accuracy for PINN-DA-HBNS resolved at the vortex-shedding frequency $St = 0.22$ (Figure 8.11(e,f)) and the corresponding resolved forcing components (Figure 8.12(c,d)). When resolving the turbulent scales at the vortex-shedding frequency (as extracted from the first SPOD mode), there is almost no noticeable resolved forcing in the critical initial separation area or near to the wall at all. Instead, most of the resolved forcing appears downstream ($x \geq 1.5$). As a consequence, in the region where reconstruction error is highest (i.e. separation and shear layer), one only observes marginal reductions in reconstruction accuracy. Analysis of the resolved forcing fields shows that resolving PINN-DA-HBNS at $St = 0.04$ has a greater effect on resolved forcing near

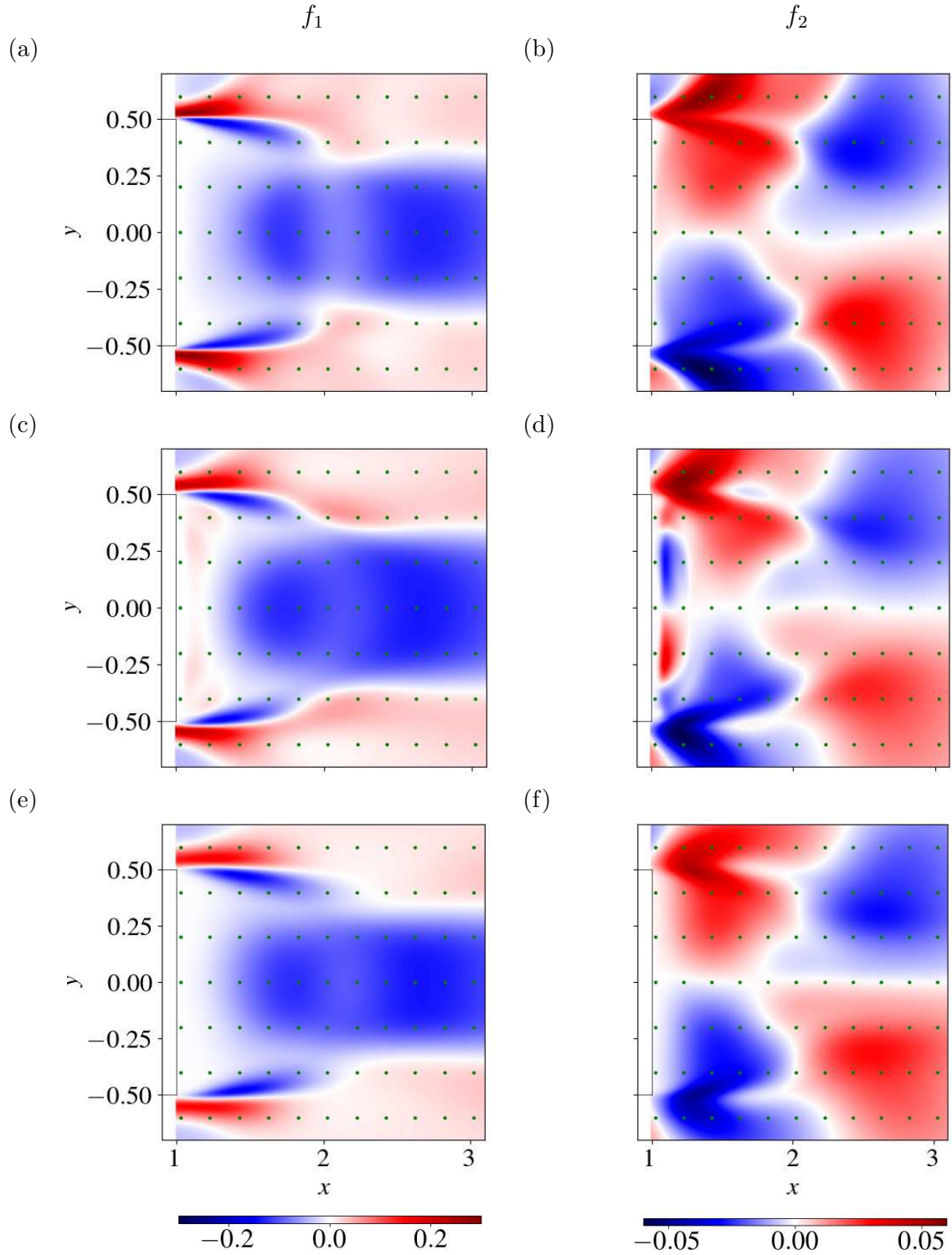


Figure 8.13: Composite forcing terms for PINN-DA reconstructed flows using (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St = 0.04$. (e,f) PINN-DA-HBNS resolved at $St = 0.22$. The composite forcing is defined as $\mathbf{f} = \mathbf{R}_{0,s} + \mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1)$, where $\mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1) = \overline{\hat{\mathbf{u}}_1} \cdot \nabla \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1} = 0$ for PINN-DA-Baseline.

the separation, where the mean velocity reconstruction error is highest, compared to $St = 0.22$, which therefore leads to a reduction in reconstruction error, as shown in Figure 8.11.

The full composite inferred forcing, presented in Figure 8.13, explains both how the use of PINN-DA-HBNS to resolve a low-frequency harmonic improves the mean-flow reconstruction accuracy compared to PINN-DA-Baseline for the bluff body wake, and why these observed improvements are relatively small compared to those seen in Chapter 7. The first key observation is that the composite forcing for both PINN-DA-Baseline and PINN-DA-HBNS resolved at $St = 0.22$ (Figure 8.13(a,b) compared with (e,f)) are similar. Despite using the HBNS equations and mean and (SPOD determined) harmonic velocity measurement data to resolve some of the low-frequency turbulent scales, there is little difference between the combined resolved and corrective ($\mathbf{R}_{0,s}$) forcing, which subsequently explains the similarity in reconstruction error. This indicates that the corrective forcing component, $\mathbf{R}_{0,s}$, has the capacity to capture some of the unresolved turbulent scales (including those captured with the first SPOD mode at the vortex-shedding frequency), and therefore even without the explicit resolution of these scales using the HBNS equation, PINN-DA-Baseline is able to reconstruct the mean flow to similar accuracy to PINN-DA-HBNS resolved at $St = 0.22$.

Unlike resolving the flow at the vortex-shedding frequency ($St = 0.22$), using the PINN-DA-HBNS formulation to resolve the flow at $St = 0.04$ has a much more noticeable effect on the total forcing of the momentum equations ($\mathbf{f}$) and therefore reconstruction error compared to PINN-DA-Baseline. Figure 8.13(c,d) demonstrates clear differences between PINN-DA-Baseline (and PINN-DA-HBNS at $St = 0.22$) and PINN-DA-HBNS at $St = 0.04$, most apparent in the near wall region, initial separation and in the shear layer. This correlates with the region that reductions in mean-flow reconstruction error are observed. It also reemphasises that choice of resolved frequency (and therefore resolved turbulent scales) is critical to performance of PINN-DA-HBNS. By resolving the flow at $St = 0.04$, a more significant fraction of the low-frequency turbulent scales are resolved, which the corrective forcing term for PINN-DA-Baseline is unable to capture (unlike when resolving at $St = 0.22$) and this subsequently leads to improvements in reconstruction accuracy.

However, comparison of the relative forcing magnitudes for resolved (Figure 8.12) and total composite (Figure 8.13) forcing demonstrates the limitations of the current PINN-DA-HBNS approach, specifically how SPOD is used to extract the harmonic modes ($\hat{\mathbf{u}}_1$). The resolved

forcing component, from the PINN-DA-HBNS reconstruction of (SPOD) harmonic velocity mode, has a magnitude up to approximately 10% of the total, composite forcing, and thus its effect on mean-flow reconstruction accuracy is restricted (this is in contrast to Chapter 7 where the magnitude approaches 80-90%). This is a direct result of how SPOD has been used to describe the flow fluctuations and represent the first low-frequency resolved harmonic velocity, $\hat{\mathbf{u}}_1$. Firstly, as described in Section 8.1.1, the energy of the bluff body wake (fluctuations) can be distributed across a broadband spectrum of frequencies. By choosing to represent the first harmonic with an SPOD mode from a single discrete frequency, substantial energy (and resolved turbulent scales) are neglected from neighbouring bins. Furthermore, using only the first SPOD mode at the given frequency, even if that mode has the highest energy content, discounts the effects of all the higher order modes at the selected frequency. Representing the first low-frequency resolved harmonic, $\hat{\mathbf{u}}_1$, with a single SPOD at a single discrete frequency neglects the effects of neighbouring frequency bins and high order SPOD modes, and thus a small fraction of the low-frequency scales are actually resolved, limiting the effectiveness of the proposed PINN-DA-HBNS framework to improve mean-flow reconstruction accuracy (compared with PINN-DA-Baseline).

Mean-flow reconstruction using turbulence model augmentation (PINN-DA-SA and PINN-DA-HBNS-SA)

In the previous section, as in Chapter 6, the highest mean-flow reconstruction error for the bluff body wake was observed in the separation and shear layer, indicating use of turbulence model augmentation would be equally beneficial here. This section explores the use of turbulence model augmentation by itself (as in Ch. 6), as well as the effect of combining turbulence model augmentation with HBNS in PINN-DA-HBNS-SA (formalised in Eq. (3.27)).

Figure 8.14 shows the effect on mean-flow reconstruction of adding an SA turbulence model to the PINN-DA formulation both with and without HBNS, on the bluff body wake. Firstly, one can consider the performance of the PINN-DA-SA formulation (Figure 8.14(a,b)) with respect to PINN-DA-Baseline (Figure 8.11(a,b)). Augmenting the system with the SA turbulence

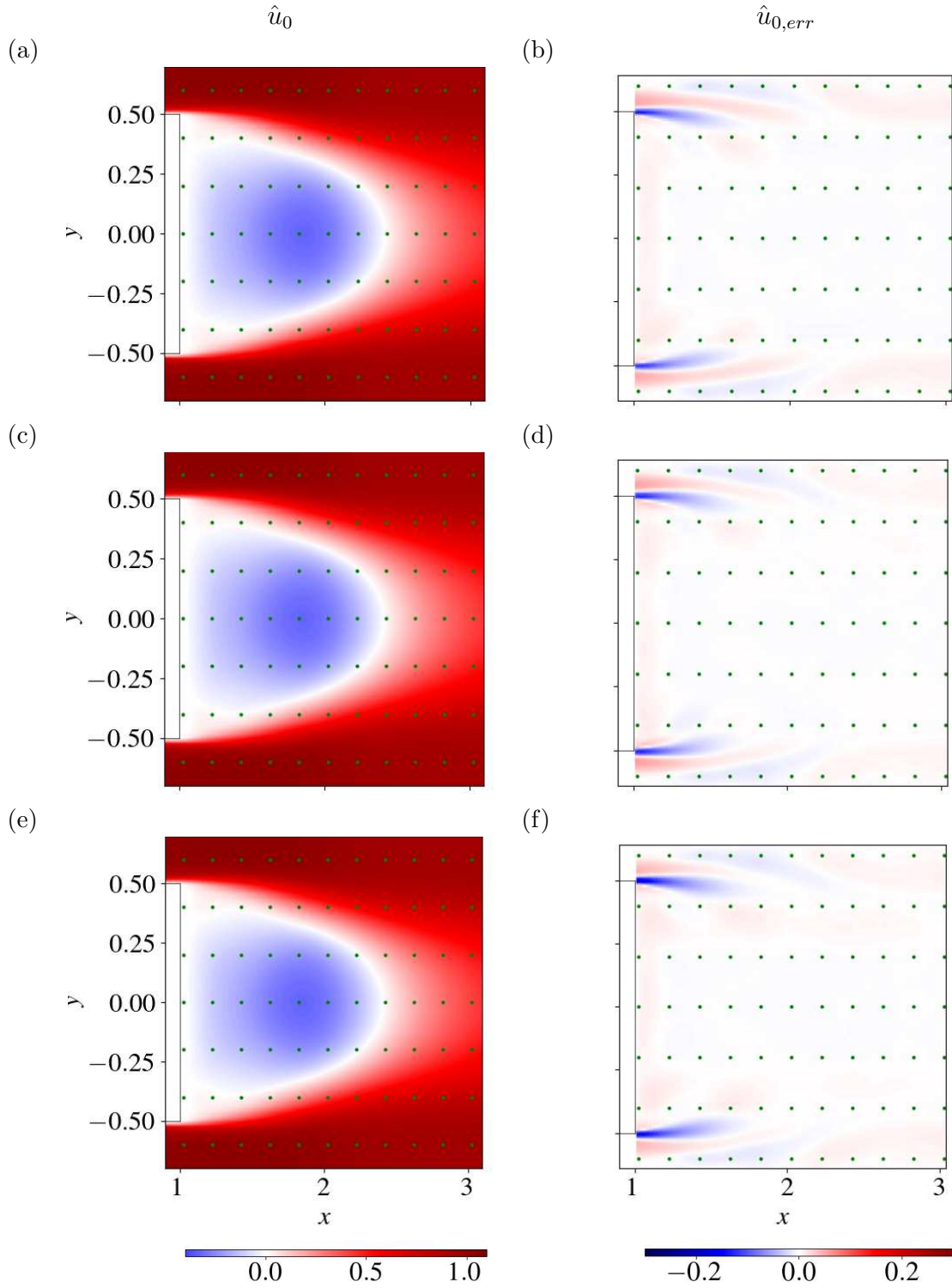


Figure 8.14: PINN reconstruction using different methodologies for data spacing $\Delta L = 0.2$, when embedded transformation method is being used. (a,b) PINN-DA-SA, (c,d) PINN-DA-HBNS-SA resolved at $St=0.04$, (e,f) PINN-DA-HBNS-SA resolved at $St=0.22$

model has a significant effect on the reconstructed mean flow in the shear layer. Chapter 6 demonstrated how PINN-DA-Baseline struggled to capture high flow gradients and boundary layers. Consistent with this prior chapter, adding the SA turbulence model, designed for

flows with adverse pressure gradients improves the prediction accuracy of the boundary layer development along the bluff body wall, in addition to further improvement in prediction of the initial separation. It follows thus, that as the prediction of the velocity profile through the boundary layer and at the separation of the bluff body improves, this reduction in mean velocity reconstruction error propagates downstream and is observed through the shear layer. This is validated in the mean velocity error profiles, presented in Figure 8.7(c). In the vicinity of the wall, the peak mean velocity reconstruction error between PINN-DA-Baseline and PINN-DA-SA reduces by approximately 50% (or 15% of free-stream). Further downstream, presented in Figure 8.7(d), use of PINN-DA-SA results in up to a 25% reduction of maximum error relative to PINN-DA-Baseline. Given the shear layer has the largest mean velocity reconstruction error, use of PINN-DA-SA has a stark effect on total reconstruction error (total reconstruction error is reduced to $\varepsilon_2 = 8.14 \times 10^{-5}$). Similarly, PINN-DA-SA demonstrates a greater reduction in mean velocity reconstruction error over PINN-DA-Baseline compared to the best PINN-DA-HBNS formulation (resolved at $St = 0.04$), as the physics knowledge provided with the turbulence model augmentation has a larger effect on the area of greatest error (boundary layer growth and separation) compared with resolution of the selected low-frequency scale, despite the introduction of modelling error from use of an empirical model.

The reconstructed velocity fields, presented in Figure 8.14(c-f), suggests that combining turbulence model augmentation (PINN-DA-SA) with resolution of some low-frequency scales (PINN-DA-HBNS) in the PINN-DA-HBNS-SA formulation does not lead to further reductions in reconstruction error, compared with PINN-DA-SA. One can observe that, whilst the PINN-DA-HBNS-SA formulation (irrespective of resolved frequency) reconstructs mean flow better than the equivalent PINN-DA-HBNS reconstruction, it is no better than the base PINN-DA-SA formulation, further indicating the physics added by the resolved turbulent scales can also be predicted through use of the SA turbulence model.

The composite forcing for the turbulence model augmented formulations (DA-SA and DA-HBNS-SA), presented in Figure 8.15, highlights the effects of using a turbulence model augmented formulation. The PINN-DA-SA composite forcing (Figure 8.15(a,b)) is almost identical to the equivalent PINN-DA-HBNS-SA composite forcing (Figure 8.15(c-f)), which subsequently

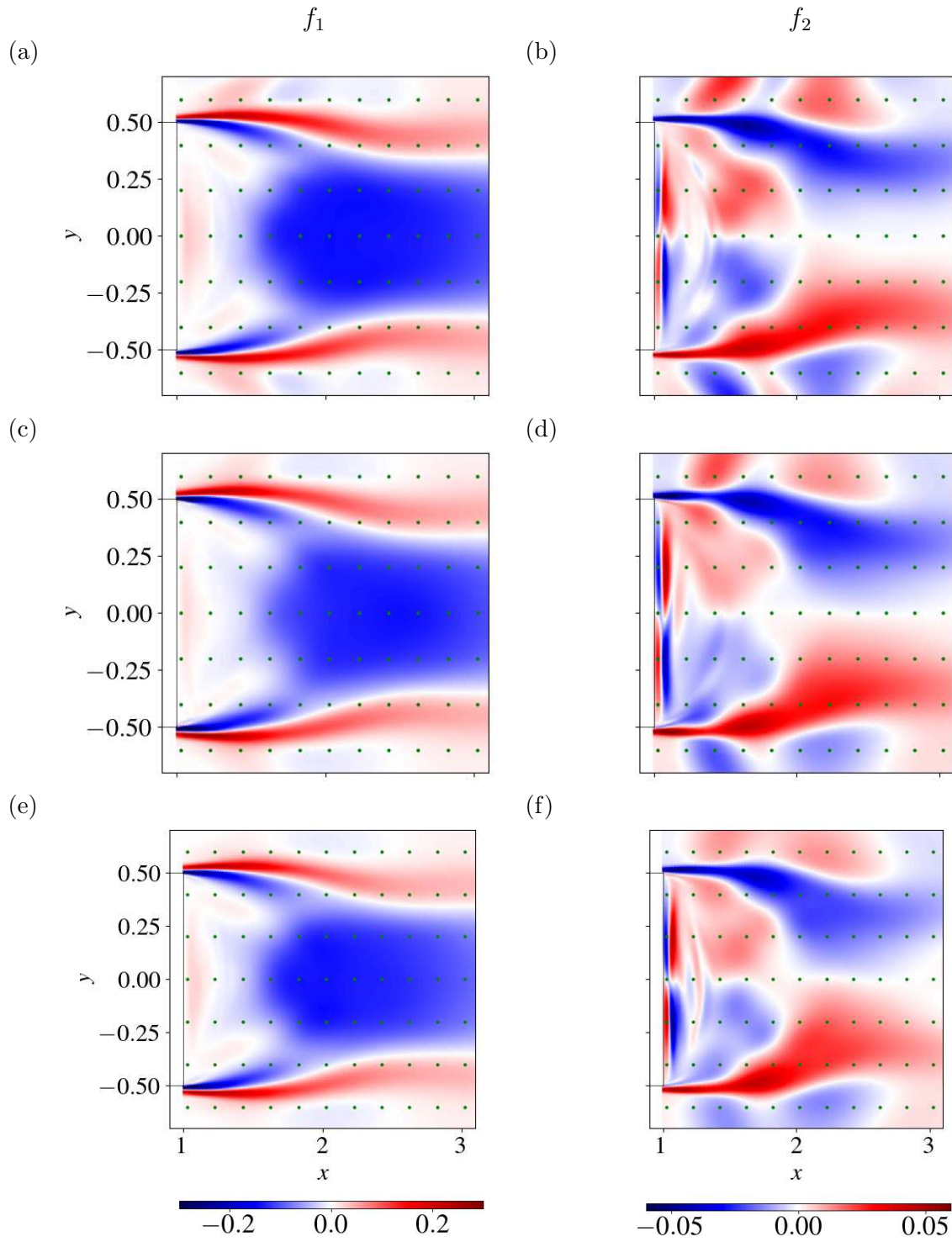


Figure 8.15: Composite forcing terms for PINN-DA reconstructed flows using (a,b) PINN-DA-SA, (c,d) PINN-DA-HBNS-SA resolved at $St = 0.04$. (e,f) PINN-DA-HBNS-SA resolved at $St = 0.22$. The composite forcing is defined as $\mathbf{f} = \mathbf{R}_{0,s} + \nabla(\nu_t \nabla_s \hat{\mathbf{u}}_0) + \mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1)$, where $\mathbf{N}(\hat{\mathbf{w}}_{-1}, \hat{\mathbf{w}}_1) = \overline{\hat{\mathbf{u}}_1 \cdot \nabla \hat{\mathbf{u}}_1} + \hat{\mathbf{u}}_1 \cdot \nabla \overline{\hat{\mathbf{u}}_1} = 0$ for PINN-DA-SA.

results in similarly reconstructed mean velocity fields. Furthermore, comparing these turbulence model augmented composite forcing fields with the corresponding PINN-DA-Baseline fields in Figure 8.13(a,b), explains the reduction in mean velocity reconstruction error in the

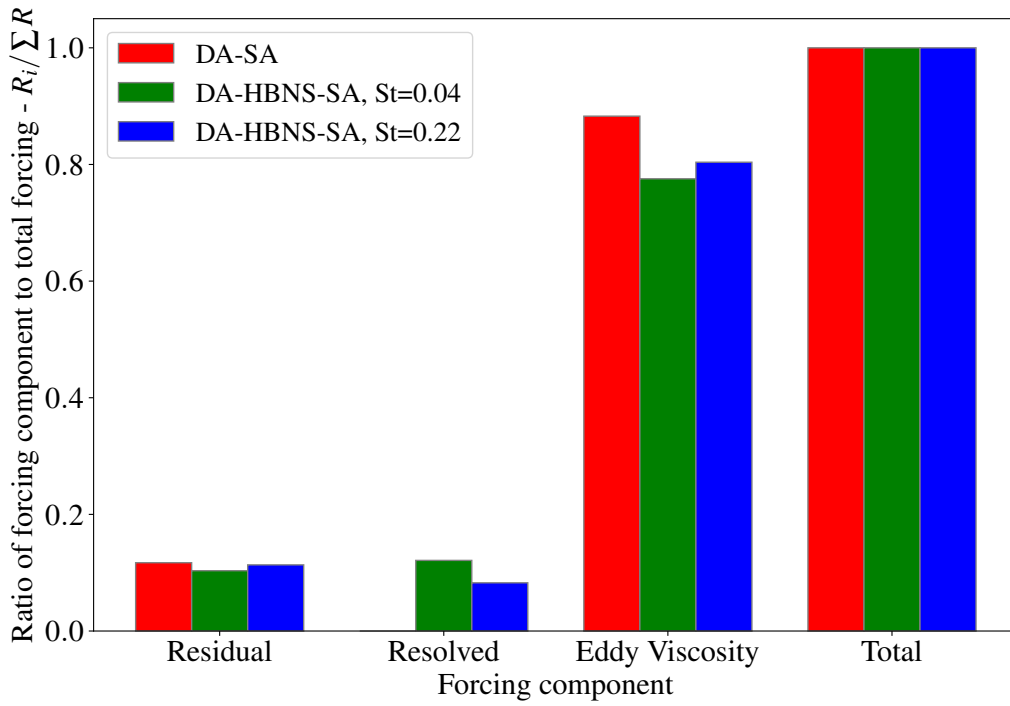


Figure 8.16: Breakdown of relative contribution of corrective ($\mathbf{R}_{0,s}$), resolved ($\hat{\mathbf{u}}_1 \cdot \nabla \hat{\mathbf{u}}_1$) and modelled components $\nabla \cdot (\nu_t \nabla_s \hat{\mathbf{u}}_0)$ to total forcing for PINN-DA-SA and PINN-DA-HBNS-SA.

critical shear layer region. The reduction in mean velocity reconstruction error in the shear layer, is coupled with an increase in composite forcing in the same area. This direct correlation shows that the addition of the modelled eddy-viscosity component of forcing provides significant physical constraints (and information) even for the more complex, turbulent bluff body wake, resulting in improvements in mean-flow reconstruction accuracy, in a manner consistent with previous cases (Ch. 6).

By breaking down the relative contribution of the three individual forcing components (modelled, resolved and corrective), as presented in Figure 8.16, one can explain the indifference between PINN-DA-SA and PINN-DA-HBNS-SA. Although resolving some of flow using the HBNS equations (and harmonic velocity data from SPOD) captures some of the low-frequency turbulent scales in the resolved forcing component (see second column of Figure 8.16), this corresponds to an equivalent reduction in modelled eddy viscosity (column 3) and possibly some corrective (column 1) forcing components. The same argument can be made for the similar reconstruction of PINN-DA-Baseline and PINN-DA-HBNS resolved at $St = 0.22$. Effectively, the resolved forcing term is only capturing information that has already been determined with

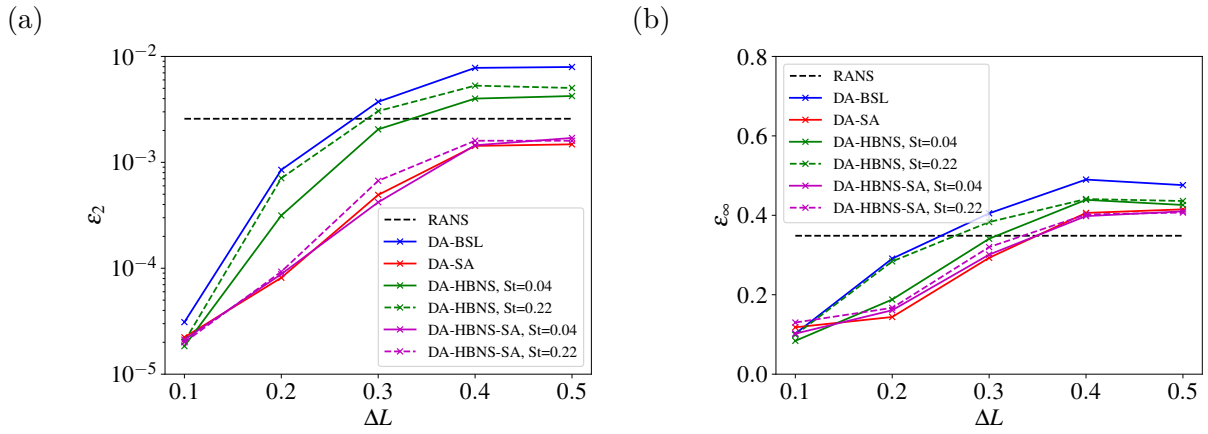


Figure 8.17: Reconstruction error for different PINN formulations using the embedded transformation approach showing (a) L_2 error and (b) L_∞ across data resolutions.

the modelled or corrective term. This is more apparent as the PINN-DA-SA formulation significantly reduces the mean velocity error in the key shear layer region. As discussed prior, neglecting neighbouring discrete frequencies and higher order SPOD modes, by only using data from a single SPOD at a single discrete frequency to represent $\hat{\mathbf{u}}_1$, means the resolved term is not sufficiently large to capture significant flow behaviour, for which the corrective or modelled term can not compensate.

Reconstruction error using different data resolutions

Whilst in Figure 8.6, it was shown how use of transformation embedded PINN-DA improved reconstruction accuracy compared to PINN-DA without it, Figure 8.17 shows how mean-flow reconstruction accuracy, across a range of data resolutions, is affected by choice of PINN-DA formulation. The L_2 -based reconstruction error for all approaches, irrespective of formulation, increases monotonically as data resolution coarsens, as presented in Figure 8.17(a). However, one can see that the turbulence model augmented approaches (PINN-DA-SA and PINN-DA-HBNS-SA) outperform all other PINN formulations, in terms of L_2 -based reconstruction accuracy. For example, the turbulence model augmented formulations (PINN-DA-SA/PINN-DA-HBNS-SA) reconstruct the flow to the same accuracy as the PINN-DA-Baseline formulation at $\Delta L = 0.2$, indicating the significant improvement from modelling a component of forcing, most obviously in the shear layer, as seen previous. Furthermore, as data coarsens the difference

between approaches (turbulence model augmented and PINN-DA-Baseline) grows. However, between PINN-DA-SA and PINN-DA-HBNS-SA, there is little difference, confirming that the modelled eddy-viscosity component outweighs the contribution of the resolved forcing, as shown in Figure 8.16.

As seen in Figure 8.11, use of the PINN-DA-HBNS formulation, does provide a small reduction in reconstruction error relative to PINN-DA-Baseline across the resolution range, albeit to a lesser degree than the turbulence model augmented approaches, showing that resolution of some low-frequency scales does have a positive effect on reconstruction performance. Specifically, PINN-DA-HBNS, resolved at the lower $St = 0.04$ frequency, improves reconstruction of the mean velocity relative to PINN-DA-Baseline across all data resolutions, whilst PINN-DA-HBNS, resolved at $St = 0.22$, reduces reconstruction error to a lesser degree. These results suggest that whilst resolving some low-frequency scales does have some benefit, the applied approach (resolving a single SPOD mode at a discrete frequency) is limited in the amount of “information” it provides to the solution and additional research to increase the amount of scales resolved is required to maximise the benefit of the PINN-DA-HBNS (or even the PINN-DA-HBNS-SA) formulation.

The L_∞ -based error, presented in Figure 8.17(b), identifies how different PINN-DA formulations at different resolutions affect flow reconstruction accuracy in the previously identified high error region i.e. the shear layer near the initial separation at the trailing edge. One can observe how the ε_∞ error increases as data coarsens from around 10% free-stream at the finest resolutions to around 40%. This significant increase can be attributed to two factors. Firstly, there is a sensitivity to data resolution but also measurement location which directly affects capture of the shear layer. Secondly, as discussed prior, as data coarsens, the embedded transformation approach becomes less effective and converges to use of PINN-DA without embedded transformation. Therefore, the improvements seen from using embedded transformations are reduced at coarser resolutions, as less information about the flow is available from which to learn spatial corrections to the measurements (as demonstrated in Appendix D.2).

As expected, the turbulence model augmented formulations have the lowest ε_∞ error as the

profile of the shear layer is most accurately captured. The second effect is that use of the transformation embedded PINN-DA is also affected by PINN-DA formulation, with the PINN-DA-SA/DA-HBNS-SA formulations demonstrating the most consistent improvement from its use. Whilst the turbulence model augmented approaches demonstrate the lowest ε_∞ error due to best capture of the shear layer, this is followed by the PINN-DA-HBNS and then PINN-DA-Baseline approaches.

Comparison of PINN-DA with RANS-SA

Figure 8.17(a) also shows the relative performance of PINN-DA with respect to a RANS-SA simulation. For data spacing approximately $\Delta L \leq 0.3$, all data assimilation approaches reconstruct the flow at least as well as a RANS simulation, if not better. At the finest data resolutions tested, PINN-DA improves ε_2 mean velocity reconstruction error by two orders of magnitude. However, as data coarsens, only the turbulence model augmented approaches perform better than the RANS-SA simulation. The comparison between maximum error for RANS-SA and PINN-DA behaves similarly to the total mean velocity reconstruction. At the finest resolutions, the maximum error is three times smaller than the RANS-SA simulation. At the coarsest resolutions, the PINN-DA behaves more similarly to RANS-SA. PINN-DA approaches, which best capture the shear layer, perform best relative to RANS-SA with respect to both total and maximum mean velocity reconstruction error.

Figure 8.18 compares the RANS-SA simulation reconstruction error with the more accurate PINN-DA-SA reconstruction at the coarsest data resolution ($\Delta L = 0.5$). Note that whilst the RANS-SA error subtracts the measured PIV field from the RANS-SA prediction, the PINN-DA-SA is first transformed into measured coordinate space. Firstly, one can see that the RANS-SA and PINN-DA-SA reconstruction are qualitatively similar, with highest error in the shear layer and the initial separation. However the PINN-DA-SA more accurately captures the wake further downstream (i.e. $x \geq 2.0$). This indicates that the turbulence model augmented PINN-DA formulation is similarly representing the flow as the RANS-SA simulation but that constraining the solution with data and allowing corrections to the modelled forcing improves

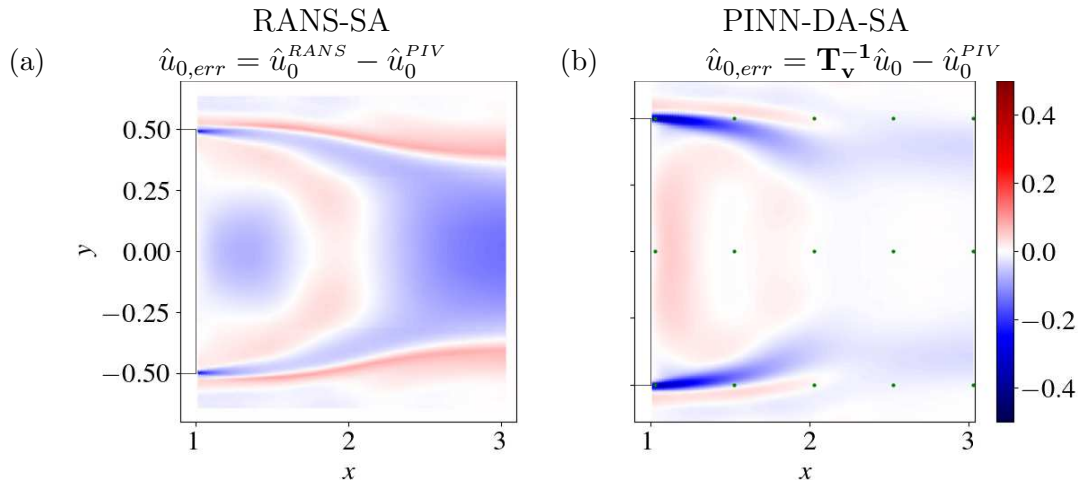


Figure 8.18: Mean streamwise velocity reconstruction error for PINN-DA-SA using data spacing $\Delta L = 0.5$, compared with RANS-SA simulation. Note how (a) the PIV data is directly subtracted from the RANS-SA prediction (i.e. both terms have underlying error), whilst (b) the PINN-DA-SA prediction is first transformed into measured coordinate space (corresponding to the measured PIV field).

reconstruction capacity.

However, at this coarser resolution, the RANS-SA simulation better captures the shear layer, resulting in similar total mean velocity reconstruction accuracy (as discussed in Section 8.2.2). This can be attributed to several factors. Firstly, as seen in other studies, PINN-DA is more data dependant than other approaches (PINNs are often not always effective solvers for the forward problem, such as a RANS-SA simulation). As the data coarsens, the optimisation problem becomes ill-posed, as there are insufficient constraints on the solution, which can be mitigated by use of the turbulence model augmented approach. Secondly, as seen in Figure 8.6, the embedded transformation approach becomes less effective when less data is provided. As a result, the influence of misaligned data on mean velocity reconstruction accuracy, as seen in Section 8.2.1, grows. This indicates use of PINN-DA is more accurate than RANS, provided sufficient data is used to constrain the data assimilation problem, but that the addition of a physical constraint (such as turbulence model augmentation) reduces data dependence.

The final factor affecting PINN-DA accuracy, relative to RANS-SA are the higher flow gradients and smaller scales in the shear layer compared with the recirculation. Whilst use of turbulence model augmentation improves the capture of the shear layer, the current architecture may still be insufficient to capture these different scales for this high Reynolds number case. Use of archi-

tectures, such as Fourier embedding or conservative PINNs (Jagtap, Kharazmi & Karniadakis 2020) (which use different PINNs for different flow regions have been shown to improve the capacity of PINNs to accurately predict flows. Successful implementation of these alternative architecture may further reduce reconstruction error, especially at the coarsest data resolutions, by better representing the different scales.

8.2.3 Pressure reconstruction of three-dimensional bluff body wake using PINN-DA

Alternatively, one can use PINN-DA to infer the pressure field in the bulk three-dimensional domain, as presented in Figure 8.19. To extract the pressure field, the PINN-DA-Baseline formulation with Reynolds stress terms was used (see Table 3.1). Instead of inferring the corrective forcing, representing the divergence of the Reynolds stress tensor (as in the typical PINN-DA-Baseline formulation), the individual Reynolds stresses were predicted. This is the same as the formulation used in Section 6.4.2. Additionally, data for the Reynolds stresses, in addition to mean velocity, from the time-resolved 2D2C PIV data were provided (at the fine PIV resolution). Using this approach, the PINN-DA inferred wall pressure was accurately predicted within 10% of the experimental wall pressure tap data, correctly capturing the shape of the pressure profile (including the two pressure valleys, as a result of the vortical recirculation structures).

One can also examine the bulk pressure field away from the wall, which is typically difficult to measure. Firstly, in the PIV plane, presented in Figure 8.19(b), one can clearly identify two low pressure regions at the centre of the (annular) recirculation vortex. Further afield, in the three-dimensional pressure field, presented in Figure 8.19(c), one can observe the low pressure centre of the annular recirculation vortex, previously seen in Figure 8.10. As mentioned prior, these flow features are consistent with the analysis completed in other literature (Gentile et al. (2016), Pavia et al. (2019), Rigas et al. (2014)). This pressure reconstruction result validates the use of PINN-DA, to infer the three-dimensional pressure field, where it is otherwise impossible to measure, from 2D2C experimental PIV measurements and can be a useful tool for engineering

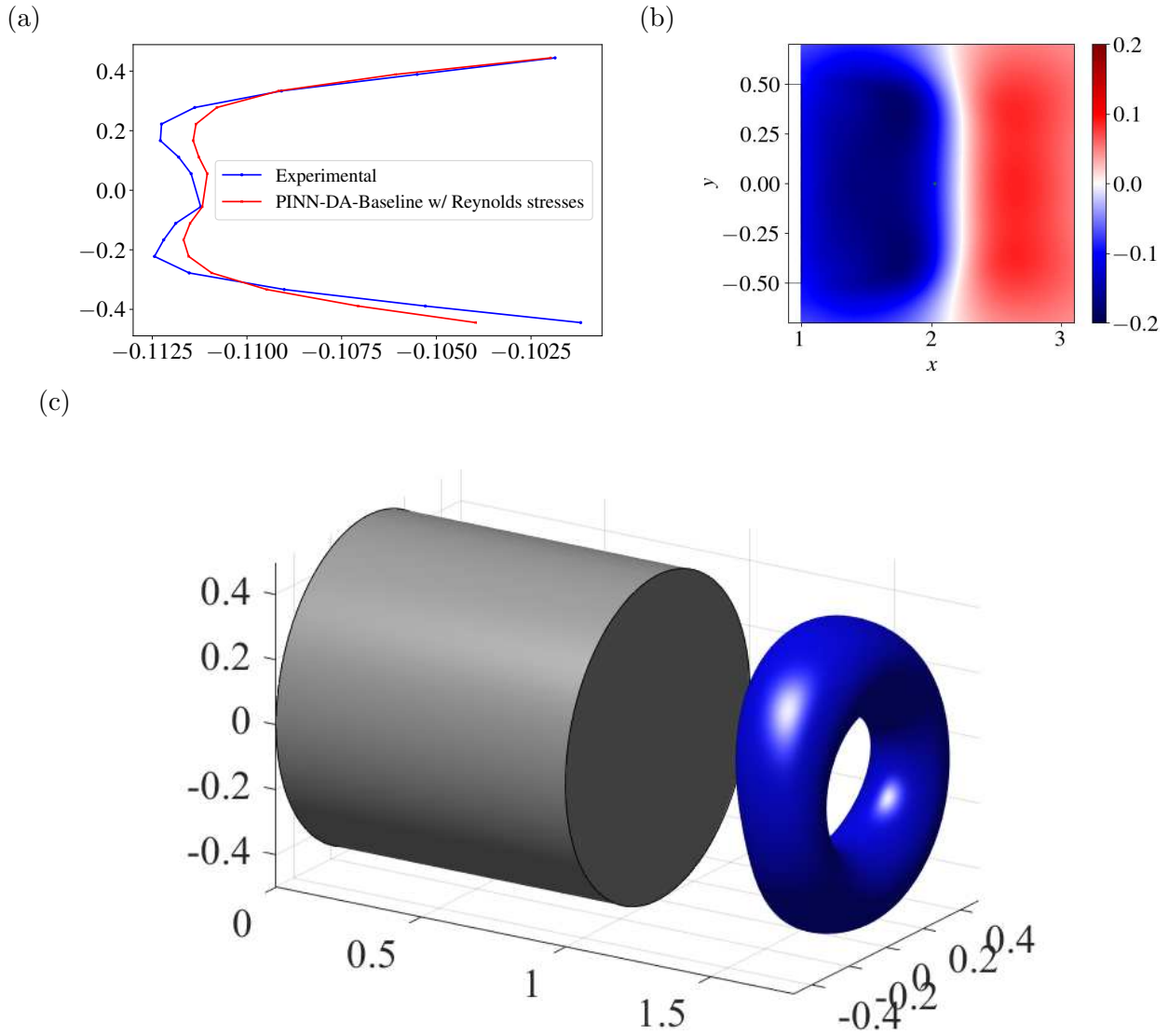


Figure 8.19: Reconstructed pressure field using PINN-DA-Baseline with non-sparse velocity and Reynolds stress data: a) Pressure prediction at wall. b) Pressure across PIV plane. c) 3D iso-pressure surface at $C_p = -0.19$.

use with aerodynamic design.

8.3 Conclusion

The successful use of PINN-DA to reconstruct an axisymmetric bluff body wake from sparse 2D2C experimental PIV measurements was presented. In particular, the effect of choice of PINN-DA formulation (Baseline, SA, HBNS, HBNS-SA) was explored. This chapter demonstrates the use of PINNs on a more realistic engineering use case.

Firstly, it was shown that miscalibration of experimental setup and resultant spatial misalignment of data can have a significant outcome on the accuracy of data-assimilation methods to reconstruct flows and therefore must be accounted for when using experimental data. The use of an affine transformation matrix embedded into the PINN-DA architecture to learn spatial corrections for measurements, successfully reduced reconstruction error in highly sensitive regions (near the bluff body) by discovering that the PIV plane required an approximately 2.5% stretch and a 0.07 shift downstream, resulting in a reduction of mean velocity reconstruction error.

Secondly, it was shown that choice of PINN-DA formulation has a big effect on the capacity of PINN-DA to reconstruct the bluff body wake from sparse 2D2C experimental PIV measurements. Consistent with the results seen in Chapter 6, the turbulence model augmented approaches had the highest reconstruction accuracy. Compared to PINN-DA-Baseline, PINN-DA-SA (and PINN-DA-HBNS-SA) reduced reconstruction error in key areas, such as the initial separation and the shear layer, due to the effect of modelling a component of forcing.

Secondly, use of the HBNS formulation, whilst successful (compared to PINN-DA-Baseline), has a smaller effect than the turbulence model augmented formulation. This is attributed to the relatively small magnitude of resolved forcing (from the low-frequency harmonic scales), explaining why the PINN-DA-SA and PINN-DA-HBNS-SA formulations reconstruct the mean flow to the same accuracy. The small effect of resolved low-frequency scales is a result of the use of SPOD, to extract (and constrain the solution with) only a single harmonic mode at a single discrete frequency. This single harmonic (SPOD) mode only captures a small fraction of the low-frequency scales, neglecting the effects of higher order SPOD modes and more importantly the effects of neighbouring frequencies of the broadband spectrum flow. In order to maximise the use of the HBNS formulation, other methods of capturing the low-frequency scales must be considered (such as those explored in Franceschini et al. (2021)).

Finally, providing both planar mean velocity and Reynolds stress data with the PINN-DA-Baseline formulation, the bulk pressure field was successfully inferred, validated against experimental wall pressure tap data. This provides vital information for the aerodynamic design

process and thus would be an invaluable tool to an engineer.

8.3.1 Future Work

Whilst this chapter has demonstrated use of PINN-DA with a more realistic engineering use case, the flow case remains axisymmetric and thus expansion to more complex three-dimensional flows, such as in Steinfurth & Weiss (2024), is important. Whilst this paper reconstructs the flow of a more complex, three-dimensional separated flow, this is completed with more data across several planes.

Furthermore, whilst PINN-DA-HBNS was applied successfully with limited success, maximising the effect of resolving scales has many potential benefits, such as reducing the dependence on modelling flow scales, capturing some of the flow unsteadiness and improving flow reconstruction accuracy. Therefore, this requires the development of an alternative approach (to the SPOD method in this chapter) to increase the effect of the resolved scales.

Finally, whilst the effect of a specific source of experimental error (spatial misalignment of measurements) was investigated, the effect of other sources of experimental error and noise, such as those seen in Kelshaw & Magri (2024), was not explored. Further studies are required, to evaluate the robustness of the presented PINN-DA mean flow reconstruction results to other sources of experimental error and noise.

Chapter 9

Conclusion

9.1 Summary

Throughout this thesis, the use of PINNs to reconstruct flow fields have been investigated, in order to achieve the goal of demonstrating their capacity to extract new information from a high-fidelity dataset for use with aerodynamic design, as described in Section 1.1.2. To this aim, four key objectives were presented in Section 2.4, targeting key steps in proving their utility as a tool in the engineering process and therefore guided the subsequent research.

The first objective was to compare PINN-DA with existing data assimilation methods, and in particular a variational adjoint-based data assimilation, showing how they perform on the same test case and whether the techniques developed for one method (adjoint-based variational) are transferable to the other (PINN). This was the primary focus of Chapter 6. In this chapter, both PINN-DA and variational-DA were applied to the the same case, a turbulent periodic hill flow (as well as a retrospective comparison for the reconstruction of a laminar cylinder flow, completed by Foures et al. (2014) and Sliwinski & Rigas (2022)). It was shown that, whilst the constrained optimisation is objectively the same, there are slight differences in implementation which affect reconstruction performance. The spatial discretisation and regularisation method employed with the variational-DA introduce errors that are avoided by using the PINN-DA formulation. As a result, for the presented case, PINN-DA reconstructed the flow more accurately

than the equivalent variational-DA across a range of data resolutions.

The second objective was to investigate how the formulation of the governing equations affect the capacity of PINNs to reconstruct flows, by adding and comparing different prior physics constraints and data. Across this work, three main formulations of the governing equations were tested: PINN-DA-Baseline, PINN-DA-SA and PINN-DA-HBNS, with the former first used in Sliwinski & Rigas (2022) and used as the benchmark. In Chapter 6, the turbulence model augmentation formulation of PINNs (PINN-DA-SA) was presented, showing a significant improvement in reconstruction accuracy of separations and boundary layers for turbulent flows (consistent with the design purpose of the SA model). Alternatively, in Chapter 7, modelling error from constraining the flow solution with an empirical turbulence model was avoided, as additional physics was added through resolution of low-frequency harmonic flow scales through the use of the HBNS equations (PINN-DA-HBNS), and tested on a laminar circular cylinder. Here, it was shown that resolving additional harmonics improved the capacity of PINN-DA-HBNS to reconstruct flow compared to PINN-DA-Baseline, with improvements in wake reconstruction and skin-friction prediction. Also, in Chapter 8, a brief investigation into combining both turbulence model augmentation and HBNS methodologies (PINN-DA-HBNS-SA) was presented, although more fine tuning is required to maximise its performance. In particular, further research is required, to evaluate how to increase the effect of the resolved harmonics and investigate if this formulation is more suitable for certain flow cases, where the turbulent spectrum is centred around a more dominant mode. These results clearly showed that choice of physics constraints have arguably the largest effect on reconstruction performance, with the addition of physics through use of turbulence model augmentation or resolving flow harmonics, proving to improve flow reconstruction accuracy.

The third objective was to expand the use of a trained PINN model beyond flow reconstruction by making use of the PINN-inferred fields. This was covered in the second half of Chapter 7. A data-driven closure model for the HBNS equations was successfully built, mapping local resolved velocities (extracted from a pre-trained PINN-DA-HBNS) to residual closure terms at the same location. This approach can be used to build closure models from limited sparse high-fidelity (experimental) data. Therefore, use of the proposed framework, shows how the

application of PINN-DA can be extended, by using the super-resolved and inferred fields, to build closure models.

The final objective was to demonstrate how PINNs can be used in a realistic engineering scenario, with the aim of applying PINN-DA to a turbulent, three-dimensional flow using experimental data (instead of “synthetic” data). This was achieved in Chapter 8, in which the mean flow of a turbulent axisymmetric bluff body wake from 2D2C experimental PIV data was completed. Firstly, it was shown that experimental data can have systematic spatial misalignment errors from sensor miscalibration, which affect data assimilation, but that the proposed transformation embedded PINN-DA can mitigate these effects. Furthermore, it was shown that, whilst all the approaches presented in Chapters 6 and 7 can be successfully used to reconstruct the full flow field even from sparse 2D2C PIV measurements, PINN-DA-SA/HBNS-SA performed best, followed by PINN-DA-HBNS and finally PINN-DA-Baseline. However, additional work is required to maximise the PINN-DA-HBNS framework for (broadband spectrum) turbulent flow cases. It was also shown that the modified PINN-DA-Baseline used in Chapter 6, can also be used to accurately determine the mean pressure in the bulk flow, where it is typically impossible to measure (without interference). These results show that PINN-DAs are capable of being used for realistic engineering flow scenarios and problems (turbulent flow, experimental data, sensor miscalibration), and, to this aim, a toolbox of different formulations and solutions have been built, to tackle an array of challenges.

9.2 Applications

The research in this work has several potential applications, both short and long-term. Some of these will be discussed here.

Considering flow reconstruction, as already discussed, the PINN-DA framework is already shown to be robust when used with realistic engineering flows. Whilst research should continue to explore the limitations of the existing framework, in its current form it can already be used to, at the very least, augment the aerodynamic design process by providing additional

information about flows that would otherwise be unavailable to engineers. This includes super-resolution of flows, where data-availability is limited and inferring bulk mean pressure away from the walls, a particular critical field. The results in this thesis suggest that the PINN-DA framework is arguably close to maturity for more wide-spread use with flow reconstruction problems.

The proposed framework to account and correct for spatially misaligned measurement error from sensor miscalibration (transformation embedded PINN-DA), is another useful tool for experimental methods. This approach was successfully able to both mitigate the effects of sensor miscalibration for PIV data, whilst also quantitatively describing the effects of spatial misalignment. Given that spatial misalignment of sensors can occur for many different experimental methods across a range of fields, the transformation embedded PINN-DA can be applied, for any system where data obeys a set of governing equations, and the effects of miscalibration are known. The results presented show that the transformation embedded PINN-DAs, can be successfully used in conjunction with experimental methods, potentially extending beyond fluid problems (although more research to validate its robustness across different problems would be required).

Finally, a framework for using PINN-DA to build HBNS closure models from sparse data was presented. Provided a sufficiently generalised closure model can be built, it may enable a wider use of HBNS numerical simulations, allowing the resolution of low-frequency turbulent scales without neglecting the effect of high-frequency behaviour. Whilst this approach demonstrated a way to build data-driven HBNS closures with PINN-DA, it could also be extended to the more established data-driven turbulence model field (with methods such as FIML).

9.3 Future Work

Whilst the key objectives of this thesis were met, completing this research also presented several key directions to develop the work further. Firstly, development and further validation of the existing approaches on realistic engineering use cases is crucial. Whilst Chapter 8 evalu-

ates the reconstruction against the high-fidelity 2D2C PIV data, to validate the full 3D flow quantitatively, the PINN-DA reconstructed third velocity component must also be validated against high-fidelity data. Similarly applying the PINN-DA to a fully three-dimensional (not just axisymmetric) flow, as in Steinfurth & Weiss (2024), would evaluate whether 2D2C PIV data is sufficient to reconstruct a highly three-dimensional flow.

Another consideration, especially for turbulence model augmented PINNs, is use of an approach that applies separate models and/or equations in different flow regimes, such as conservative PINNs (Jagtap, Kharazmi & Karniadakis 2020) or Fourier embedded PINNs. Given that turbulent flow are made of different flow scales and different turbulence models have ideal design regimes, it follows that selecting the best model for a given region (or for a different flow scale) would minimise the required corrections and improve reconstruction accuracy.

One of the key areas for future development is application of the PINN-DA-HBNS to reconstruct more complex flows. Whilst the PINN-DA-HBNS provided more significant improvements with the laminar circular cylinder, these were more modest for the bluff body wake. This was mostly attributed to the relatively low contribution of the discrete, single frequency resolved harmonic mode to the overall solution. It follows, therefore, that to improve the performance of the PINN-DA-HBNS for broadband flows, an alternative approach to the presented SPOD method to extract modes is required. This may include methods, such as resolving more harmonics, applying some modal energy maximisation term to the cost function, or using an approach more similar to Franceschini et al. (2021). Similarly, the success of the PINN-DA-HBNS approach may depend on the suitability of the flow case. To test the dependence on flow cases, future work must include applying the PINN-DA-HBNS framework on turbulent flows, where the turbulent spectrum is centred around a dominant mode (e.g. square circular cylinder flow).

A brief mention should be given to the continuing evolution of PINNs, especially with regards to new architectures and model features. Whilst many of these were tested in this research, the focus of this research was on the effects of the formulation of the underlying physics. However a comprehensive review of the different architectures and their strengths and weaknesses for flow reconstruction problems may highlight potential avenues to improve the reconstructions

with the existing formulations, such as the convolutional-based approaches from Kelshaw & Magri (2024) and Mo & Magri (2024).

One critical component for wide-spread use of PINN-DA for mean flow reconstruction in industry, is to evaluate and quantify the robustness of the researched approaches. Whilst a method for mitigating one source of experimental measurement error (spatial misalignment) was proposed, other error sources such as systematic error (bias or epistemic uncertainty) and noise also exist. The next step would be to demonstrate the robustness of the PINN-DA methodologies to these other sources of error. Similarly, a more comprehensive analysis of sensitivity of reconstruction accuracy to hyper-parameters would demonstrate further robustness of the PINN-DA framework. Additionally, it was seen in each of the flow cases that enforcing data at certain spatial locations had a larger effect on flow reconstruction accuracy than others. It follows, that there is some optimal sensor placement. A key development for experimental application is the design of approaches, which remove redundant sensors or arrange a fixed number of sensors, such that one maximises reconstruction accuracy, as in Mons & Marquet (2021). Therefore, to validate the robustness of the proposed methodologies to various factors, such as experimental error sources, hyper-parameter sensitivity and data selection and sensor placement, future work requires the completion of uncertainty quantification.

Finally, another important step is the continued development of the HBNS closure model. Here, future work is two-fold. Firstly, whilst this research built a model that demonstrated some generalisability (same flow case but higher Reynolds number) for a fully generalisable model, this must be built across many different flow cases and regimes. Similarly, the importance of input features was shown and thus an additional focus on feature selection (such as tensor basis), as in many data-driven turbulence models, will likely improve the capacity of the closure model. The second aspect is to actually apply the closure models to a computational solver. To demonstrate the usability of an HBNS closure model, it must be implemented with a numerical solver first. Given some of the challenges faced by existing data-driven RANS turbulence models, it is likely similar challenges will need to be solved for the HBNS closure model.

Bibliography

Agarwal, K., Ram, O., Wang, J., Lu, Y. & Katz, J. (2021), ‘Reconstructing velocity and pressure from noisy sparse particle tracks using constrained cost minimization’, *Exp. Fluids* **62**, 1–20.

Agrawal, A. & Koutsourelakis, P. (2024), ‘A probabilistic, data-driven closure model for RANS simulations with aleatoric, model uncertainty’, *J. Comput. Phys.* **508**, 112982.

Angriman, S., Cobelli, P., Mininni, P. D., Obligado, M. & Clark Di Leoni, P. (2023), ‘Assimilation of statistical data into turbulent flows using physics-informed neural networks’, *Eur. Phys. J. E* **46**(3), 13.

Beetham, S. & Capecelatro, J. (2020), ‘Formulating turbulence closures using sparse regression with embedded form invariance’, *Phys. Rev. Fluids* **5**(8), 084611.

Belligoli, Z., Dwight, R. P. & Eitelberg, G. (2021), ‘Reconstruction of turbulent flows at high Reynolds numbers using data assimilation techniques’, *AIAA J.* **59**(3), 855–867.

Brener, B. P., Cruz, M. A., Thompson, R. L. & Anjos, R. P. (2021), ‘Conditioning and accurate solutions of Reynolds average Navier–Stokes equations with data-driven turbulence closures’, *J. Fluid Mech.* **915**.

Brenner, O., Plogmann, J., Piroozmand, P. & Jenny, P. (2024), ‘A variational data assimilation approach for sparse velocity reference data in coarse RANS simulations through a corrective forcing term’, *Comput. Methods Appl. Mech. Eng.* **427**, 117026.

URL: <https://www.sciencedirect.com/science/article/pii/S0045782524002822>

- Brooks, A. N. & Hughes, T. J. R. (1982), ‘Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations’, *Comput. Methods Appl. Mech. Eng.* **32**(1-3), 199–259.
- Brunton, S. L., Noack, B. R. & Koumoutsakos, P. (2020), ‘Machine learning for fluid mechanics’, *Annu. Rev. Fluid Mech.* **52**, 477–508.
- Cantwell, C. D., Moxey, D., Comerford, A., Bolis, A., Rocco, G., Mengaldo, G., De Grazia, D., Yakovlev, S., Lombard, J. E., Ekelschot, D. et al. (2015), ‘Nektar++: An open-source spectral/hp element framework’, *Comput. Phys. Commun.* **192**, 205–219.
- Cato, A. S., Volpiani, P. S., Mons, V., Marquet, O. & Sipp, D. (2023), ‘Comparison of different data-assimilation approaches to augment RANS turbulence models’, *Comput. Fluids* **266**, 106054.
- Chandramouli, P., Mémin, E. & Heitz, D. (2020), ‘4D large scale variational data assimilation of a turbulent flow with a dynamics error model’, *J. Comput. Phys.* **412**, 109446.
- Chen, K. K., Tu, J. H. & Rowley, C. W. (2012), ‘Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses’, *J. Nonlinear Sci.* **22**, 887–915.
- Cherroud, S., Merle, X., Cinnella, P. & Gloerfelt, X. (2022), ‘Sparse Bayesian Learning of Explicit Algebraic Reynolds-Stress models for turbulent separated flows’, *Int. J. Heat Fluid Flow* **98**, 109047.
- Clark Di Leoni, P., Agarwal, K., Zaki, T. A., Meneveau, C. & Katz, J. (2023), ‘Reconstructing turbulent velocity and pressure fields from under-resolved noisy particle tracks using physics-informed neural networks’, *Exp. Fluids* **64**(5), 95.
- De Ryck, T., Jagtap, A. D. & Mishra, S. (2023), ‘Error estimates for physics-informed neural networks approximating the Navier–Stokes equations’, *IMA J. Numer. Anal.* .
- de Zordo-Banliat, M., Dergham, G., Merle, X. & Cinnella, P. (2024), ‘Space-dependent turbulence model aggregation using machine learning’, *J. Comput. Phys.* **497**, 112628.

- Deck, S., Gand, F., Brunet, V. & Ben Khelil, S. (2014), ‘High-fidelity simulations of unsteady civil aircraft aerodynamics: stakes and perspectives. Application of zonal detached eddy simulation’, *Philos. Trans. R. Soc. A* **372**(2022), 20130325.
- Du, Y., Wang, M. & Zaki, T. A. (2023), ‘State estimation in minimal turbulent channel flow: A comparative study of 4DVar and PINN’, *Int. J. Heat Fluid Flow* **99**, 109073.
- Duraisamy, K., Iaccarino, G. & Xiao, H. (2019), ‘Turbulence modeling in the age of data’, *Annu. Rev. Fluid Mech.* **51**, 357–377.
- Edeling, W. N., Cinnella, P. & Dwight, R. P. (2014), ‘Predictive RANS simulations via Bayesian model-scenario averaging’, *J. Comput. Phys.* **275**, 65–91.
- Eivazi, H., Tahani, M., Schlatter, P. & Vinuesa, R. (2022), ‘Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes equations’, *Phys. Fluids* **34**(7), 075117.
- Fabre, D., Citro, V., Ferreira Sabino, D., Bonnefils, P., Sierra, J., Giannetti, F. & Pigou, M. (2018), ‘A practical review on linear and nonlinear global approaches to flow instabilities’, *Appl. Mech. Rev.* **70**(6), 060802.
- Foures, D. P. G., Dovetta, N., Sipp, D. & Schmid, P. J. (2014), ‘A data-assimilation method for Reynolds-averaged Navier–Stokes-driven mean flow reconstruction’, *J. Fluid Mech.* **759**, 404–431.
- Franceschini, L., Sipp, D. & Marquet, O. (2020), ‘Mean-flow data assimilation based on minimal correction of turbulence models: Application to turbulent high Reynolds number backward-facing step’, *Phys. Rev. Fluids* **5**(9), 094603.
- Franceschini, L., Sipp, D. & Marquet, O. (2021), ‘Mean-and unsteady-flow reconstruction with one or two time-resolved measurements’, *arXiv preprint arXiv:2102.03839*.
- Gao, H., Sun, L. & Wang, J. X. (2021), ‘Super-resolution and denoising of fluid flow using physics-informed convolutional neural networks without high-resolution labels’, *Phys. Fluids* **33**(7), 073603.

- Gatski, T. B. & Speziale, C. G. (1993), ‘On explicit algebraic stress models for complex turbulent flows’, *J. Fluid Mech.* **254**, 59–78.
- Gentile, V., Schrijer, F. F. J., Van Oudheusden, B. W. & Scarano, F. (2016), ‘Low-frequency behavior of the turbulent axisymmetric near-wake’, *Phys. Fluids* **28**(6).
- Ghosh, S., Mons, V., Sipp, D. & Schmid, P. J. (2024), ‘A robust computational framework for variational data assimilation of mean flows with sparse measurements corrupted by strong outliers’, *J. Comput. Phys.* **508**, 113008.
URL: <https://www.sciencedirect.com/science/article/pii/S0021999124002572>
- Girimaji, S. S. (2023), ‘Turbulence closure modeling with machine learning approaches: A perspective’, *arXiv preprint arXiv:2312.14902*.
- Grabe, C., Jäckel, F., Khurana, P. & Dwight, R. P. (2023), ‘Data-driven augmentation of a RANS turbulence model for transonic flow prediction’, *Int. J. Numer. Methods Heat Fluid Flow* **33**(4), 1544–1561.
- Han, J., Zhou, X. & Xiao, H. (2023), ‘An equivariant neural operator for developing nonlocal tensorial constitutive models’, *J. Comput. Phys.* **488**, 112243.
- Hasanuzzaman, G., Eivazi, H., Merbold, S., Egbers, C. & Vinuesa, R. (2023), ‘Enhancement of PIV measurements via physics-informed neural networks’, *Meas. Sci. Technol.* **34**(4), 044002.
- He, X., Tan, J., Rigas, G. & Vahdati, M. (2022), ‘On the explainability of machine-learning-assisted turbulence modeling for transonic flows’, *Int. J. Heat Fluid Flow* **97**, 109038.
URL: <https://www.sciencedirect.com/science/article/pii/S0142727X22001084>
- Hecht, F. (2012), ‘New development in FreeFem++’, *J. Numer. Math.* **20**(3-4), 251–265.
URL: <https://freefem.org/>
- Hernández, C. G., Yang, Q. & Hwang, Y. (2022), ‘Generalised quasilinear approximations of turbulent channel flow. Part 2. Spanwise triadic scale interactions’, *J. Fluid Mech.* **944**, A34.

- Holland, J. R., Baeder, J. D. & Duraisamy, K. (2019), Towards integrated field inversion and machine learning with embedded neural networks for RANS modeling, *in* ‘AIAA Scitech 2019 Forum’, American Institute of Aeronautics and Astronautics, p. 1884.
- Hornik, K., Stinchcombe, M. & White, H. (1989), ‘Multilayer feedforward networks are universal approximators’, *Neural Netw.* **2**(5), 359–366.
- Hu, Z., Jagtap, A. D., Karniadakis, G. E. & Kawaguchi, K. (2022), ‘When Do Extended Physics-Informed Neural Networks (XPINNs) Improve Generalization?’, *SIAM J. Sci. Comput.* **44**(5), A3158–A3182.
- Israel, D. M. (2023), ‘The myth of URANS’, *J. Turbul.* **24**(8), 367–392.
- Jagtap, A. D. & Karniadakis, G. E. (2020), ‘Extended Physics-Informed Neural Networks (XPINNs): A Generalized Space-Time Domain Decomposition Based Deep Learning Framework for Nonlinear Partial Differential Equations’, *Commun. Comput. Phys.* **28**(5), 2002–2041.
- Jagtap, A. D., Kawaguchi, K. & Karniadakis, G. E. (2020), ‘Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks’, *Proc. Math. Phys. Eng. Sci.* **476**(2239), 20200334.
- Jagtap, A. D., Kharazmi, E. & Karniadakis, G. E. (2020), ‘Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems’, *Comput. Methods Appl. Mech. Eng.* **365**, 113028.
- Jagtap, A. D., Mao, Z., Adams, N. & Karniadakis, G. E. (2022), ‘Physics-informed neural networks for inverse problems in supersonic flows’, *J. Comput. Phys.* **466**, 111402.
- Jagtap, A. D., Mitsotakis, D. & Karniadakis, G. E. (2022), ‘Deep learning of inverse water waves problems using multi-fidelity data: Application to Serre–Green–Naghdi equations’, *Ocean Eng.* **248**, 110775.
- Jiang, H. & Cheng, L. (2021), ‘Large-eddy simulation of flow past a circular cylinder for Reynolds numbers 400 to 3900’, *Phys. Fluids* **33**(3).

- Kaiser, T. L. & Oberleithner, K. (2021), ‘Modeling the Transport of Fuel Mixture Perturbations and Entropy Waves in the Linearized Framework’, *J. Eng. Gas Turbines Power* **143**(11), 111001.
- Kelshaw, D. & Magri, L. (2024), ‘Physics-constrained convolutional neural networks for inverse problems in spatiotemporal partial differential equations’, *arXiv preprint arXiv:2401.10306*.
- Kochenderfer, M. J. & Wheeler, T. A. (2019), *Algorithms for optimization*, MIT Press.
- Lagaris, I. E., Likas, A. & Fotiadis, D. I. (1998), ‘Artificial neural networks for solving ordinary and partial differential equations’, *IEEE Trans. Neural Netw.* **9**(5), 987–1000.
- Laizet, S. & Li, N. (2011), ‘Incompact3d: A powerful tool to tackle turbulence problems with up to O (10⁵) computational cores’, *Int. J. Numer. Methods Fluids* **67**(11), 1735–1757.
- Launder, B. E., Reece, G. J. & Rodi, W. (1975), ‘Progress in the development of a Reynolds-stress turbulence closure’, *J. Fluid Mech.* **68**(3), 537–566.
- Launder, B. E. & Sharma, B. I. (1974), ‘Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc’, *Lett. Heat Mass Transf.* **1**(2), 131–137.
- Lu, L., Meng, X., Mao, Z. & Karniadakis, G. E. (2021), ‘DeepXDE: A deep learning library for solving differential equations’, *SIAM Rev.* **63**(1), 208–228.
- Lumley, J. L. (1970), *Stochastic Tools in Turbulence*, Academic Press.
- Lumley, J. L. (1979), Computational Modeling of Turbulent Flows, in C. Yih, ed., ‘Advances in Applied Mechanics’, Vol. 18, Elsevier, pp. 123–176.
URL: <https://www.sciencedirect.com/science/article/pii/S0065215608702667>
- Luo, S., Vellakal, M., Koric, S., Kindratenko, V. & Cui, J. (2020), Parameter identification of RANS turbulence model using physics-embedded neural network, in ‘International Conference on High Performance Computing’, Springer, pp. 137–149.
- Magri, L. & Doan, N. A. K. (2020), *Physics-Informed Data-Driven Prediction of Turbulent Reacting Flows with Lyapunov Analysis and Sequential Data Assimilation*, Springer International Publishing, Cham, pp. 177–196.

- Mao, Z., Jagtap, A. D. & Karniadakis, G. E. (2020), ‘Physics-informed neural networks for high-speed flows’, *Comput. Methods Appl. Mech. Eng.* **360**, 112789.
- Marquet, O., Mons, V., Zauner, M. & Leclaire, B. (2022), Turbulent mean flow estimation with state observer assimilation of velocity measurements in RANS models, *in* ‘TSFP 12’.
- McKeon, B., Comte-Bellot, G., Foss, J., J. W., Scarano, F., Tropea, C., Meyers, J., Lee, J., Cavone, A., Schodl, R., Koochesfahani, M., Andreopoulos, Y., Dahm, W., Mullin, J., Wallace, J., Vukoslavčević, P., Morris, S., Pardyjak, E. & Cuerva, A. (2007), *Velocity, Vorticity, and Mach Number*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 215–471.
- Mishra, A., Iaccarino, G. & Duraisamy, K. (2015), ‘Epistemic uncertainty in statistical Markovian turbulence models’, *CTR Annu. Res. Briefs* **2015**, 183–195.
- Mishra, S. & Molinaro, R. (2022), ‘Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs’, *IMA J. Numer. Anal.* **42**(2), 981–1022.
- Mo, Y. & Magri, L. (2024), ‘Reconstructing unsteady flows from sparse, noisy measurements with a physics-constrained convolutional neural network’, *arXiv preprint arXiv:2409.00260*.
- Mo, Y., Traverso, T. & Magri, L. (2024), ‘Decoder Decomposition for the Analysis of the Latent Space of Nonlinear Autoencoders With Wind-Tunnel Experimental Data’, *arXiv preprint arXiv:2404.19660*.
- Molnar, J. P. & Grauer, S. J. (2022), ‘Flow field tomography with uncertainty quantification using a Bayesian physics-informed neural network’, *Meas. Sci. Technol.* **33**(6), 065305.
- Molnar, J. P., Venkatakrishnan, L., Schmidt, B. E., Sipkens, T. A. & Grauer, S. J. (2023), ‘Estimating density, velocity, and pressure fields in supersonic flows using physics-informed BOS’, *Exp. Fluids* **64**(1), 14.
- Mons, V. & Marquet, O. (2021), ‘Linear and nonlinear sensor placement strategies for mean-flow reconstruction via data assimilation’, *J. Fluid Mech.* **923**, A1.

- Mons, V., Marquet, O., Leclaire, B., Cornic, P. & Champagnat, F. (2022), ‘Dense velocity, pressure and Eulerian acceleration fields from single-instant scattered velocities through Navier–Stokes-based data assimilation’, *Meas. Sci. Technol.* **33**(12), 124004.
URL: <https://dx.doi.org/10.1088/1361-6501/ac8dac>
- NASA (2015), ‘Operations & Maintenance Cost Study for NASA Facilities: Final Report for Wind Tunnels’, <https://www.nasa.gov/wp-content/uploads/2023/06/nasa-wind-tunnels-final-report-20150813-tagged.pdf>. [Accessed 10-06-2024].
- Nocedal, J. (1980), ‘Updating Quasi-Newton Matrices With Limited Storage’, *Math. Comput.* **35**, 773–782.
- NVIDIA (2024), ‘Modulus - A Neural Network framework’.
URL: <https://developer.nvidia.com/modulus>
- Nóvoa, A., Noiray, N., Dawson, J. R. & Magri, L. (2024), ‘A real-time digital twin of azimuthal thermoacoustic instabilities’.
URL: <https://arxiv.org/abs/2404.18793>
- Nóvoa, A., Racca, A. & Magri, L. (2024), ‘Inferring unknown unknowns: Regularized bias-aware ensemble Kalman filter’, *Comput. Methods Appl. Mech. Engrg.* **418**, 116502.
URL: <https://www.sciencedirect.com/science/article/pii/S0045782523006266>
- Olshanskii, M., Lube, G., Heister, T. & Löwe, J. (2009), ‘Grad–div stabilization and subgrid pressure models for the incompressible Navier–Stokes equations’, *Comput. Methods Appl. Mech. Eng.* **198**, 3975–3988.
- Parish, E. J. & Duraisamy, K. (2016), ‘A paradigm for data-driven predictive modeling using field inversion and machine learning’, *J. Comput. Phys.* **305**, 758–774.
- Parish, E. J. & Duraisamy, K. (2017), ‘Non-Markovian closure models for large eddy simulations using the Mori-Zwanzig formalism’, *Phys. Rev. Fluids* **2**(1), 014604.
- Pavia, G., Varney, M., Passmore, M. & Almond, M. (2019), ‘Three dimensional structure of the unsteady wake of an axisymmetric body’, *Phys. Fluids* **31**(2).

- Pioch, F., Harmening, J. H., Müller, A. M., Peitzmann, F., Schramm, D. & el Moctar, O. (2023), ‘Turbulence modeling for physics-informed neural networks: Comparison of different RANS models for the backward-facing step flow’, *Fluids* **8**(2), 43.
- Plogmann, J., Brenner, O. & Jenny, P. (2024a), ‘Spectral adjoint-based assimilation of sparse data in unsteady simulations of turbulent flows’, *arXiv preprint arXiv:2405.20160*.
- Plogmann, J., Brenner, O. & Jenny, P. (2024b), ‘Variational assimilation of sparse time-averaged data for efficient adjoint-based optimization of unsteady RANS simulations’, *Comput. Methods Appl. Mech. Eng.* **427**, 117052.
- Pope, S. B. (2000), *Turbulent Flows*, Cambridge University Press.
- Raissi, M., Perdikaris, P. & Karniadakis, G. E. (2019), ‘Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations’, *J. Comput. Phys.* **378**, 686–707.
- Raynaud, G., Houde, S. & Gosselin, F. P. (2022), ‘ModalPINN: An extension of physics-informed Neural Networks with enforced truncated Fourier decomposition for periodic flow reconstruction using a limited number of imperfect sensors’, *J. Comput. Phys.* **464**, 111271.
- Rigas, G., Oxlade, A. R., Morgans, A. S. & Morrison, J. F. (2014), ‘Low-dimensional dynamics of a turbulent axisymmetric wake’, *J. Fluid Mech.* **755**, R5.
- Rigas, G. & Schmid, P. J. (2022), Data-driven closure of the harmonic-balanced Navier-Stokes equations in the frequency domain, in ‘Proceedings of the Summer Program’, pp. 67–76.
- Rodriguez, S. (2019), *LES and DNS Turbulence Modeling*, Springer International Publishing, Cham, pp. 197–223.
- Saredi, O., Ramesh, N. T., Sciacchitano, A. & Scarano, F. (2021), ‘State observer data assimilation for RANS with time-averaged 3D-PIV data’, *Comput. Fluids* **8**, 104827.
- Schmid, P. J. (2010), ‘Dynamic mode decomposition of numerical and experimental data’, *J. Fluid Mech.* **656**, 5–28.

- Schmidt, O. T. & Towne, A. (2019), ‘An efficient streaming algorithm for spectral proper orthogonal decomposition’, *Comput. Phys. Commun.* **237**, 98–109.
- Sharma, P., Evans, L., Tindall, M. & Nithiarasu, P. (2023), ‘Stiff-PDEs and physics-informed neural networks’, *Arch. Comput. Methods Eng.* **30**(5), 2929–2958.
- Shukla, K., Jagtap, A. D. & Karniadakis, G. E. (2021), ‘Parallel physics-informed neural networks via domain decomposition’, *J. Comput. Phys.* **447**, 110683.
- Singh, A. P. (2018), A framework to improve turbulence models using full-field inversion and machine learning, PhD thesis, University of Michigan.
- Singh, A. P. & Duraisamy, K. (2016), ‘Using field inversion to quantify functional errors in turbulence closures’, *Phys. Fluids* **28**, 045110.
- Sliwinski, L. & Rigas, G. (2022), ‘Mean flow reconstruction of unsteady flows using Physics-Informed Neural Networks’, *Data-Centric Eng.* **2021**, 1–18.
- Smagorinsky, J. (1963), ‘General circulation experiments with the primitive equations: I. The basic experiment’, *Mon. Weather Rev.* **91**(3), 99–164.
- Spalart, P. & Allmaras, S. (1992), A one-equation turbulence model for aerodynamic flows, *in* ‘30th Aerospace Sciences Meeting and Exhibit’, p. 439.
- Steinfurth, B. & Weiss, J. (2024), ‘Assimilating experimental data of a mean three-dimensional separated flow using physics-informed neural networks’, *Phys. Fluids* **36**(1).
- Symon, S., Dovetta, N., McKeon, B. J., Sipp, D. & Schmid, P. J. (2017), ‘Data assimilation of mean velocity from 2D PIV measurements of flow over an idealized airfoil’, *Exp. Fluids* **58**(5), 1–17.
- Symon, S. & McKeon, B. J. (2018), Experimental flow reconstruction using resolvent analysis and data-assimilation, *in* ‘21st Australasian fluid mechanics conference, Adelaide, Australia’.
- Symon, S., Sipp, D., Schmid, P. J. & McKeon, B. J. (2020), ‘Mean and unsteady flow reconstruction using data-assimilation and resolvent analysis’, *AIAA J.* **58**(2), 575–588.

- Tayeh, C., Mons, V. & Marquet, O. (2023), ‘Data assimilation of turbulent separated flows using single synthetic and experimental wall-pressure data’. working paper or preprint.
URL: <https://hal.science/hal-04251656>
- Towne, A., Schmidt, O. T. & Colonius, T. (2018), ‘Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis’, *J. Fluid Mech.* **847**, 821–867.
- van der Meer, R., Oosterlee, C. W. & Borovikh, A. (2022), ‘Optimally weighted loss functions for solving pdes with neural networks’, *J. Comput. Appl. Math.* **405**, 113887.
- Vinuesa, R. (2021), ‘High-fidelity simulations in complex geometries: Towards better flow understanding and development of turbulence models’, *Results Eng.* **11**, 100254.
- Volpiani, P. S., Meyer, M., Franceschini, L., Dandois, J., Renac, F., Martin, E., Marquet, O. & Sipp, D. (2021), ‘Machine learning-augmented turbulence modeling for RANS simulations of massively separated flows’, *Phys. Rev. Fluids* **6**(6), 064607.
- von Saldern, J. G. R., Reumschüssel, J. M., Kaiser, T. L., Sieber, M. & Oberleithner, K. (2022), ‘Mean flow data assimilation based on physics-informed neural networks’, *Phys. Fluids* **34**(11), 115129.
- von Saldern, J. G., Reumschüssel, J. M., Kaiser, T. L., Schmidt, O. T., Jordan, P. & Oberleithner, K. (2023), Self-consistent closure modeling for linearized mean field methods, in ‘AIAA AVIATION 2023 Forum’, p. 4351.
- Wallin, S. & Johansson, A. V. (2000), ‘An explicit algebraic Reynolds stress model for incompressible and compressible turbulent flows’, *J. Fluid Mech.* **403**, 89–132.
- Wang, H., Liu, Y. & Wang, S. (2022), ‘Dense velocity reconstruction from particle image velocimetry/particle tracking velocimetry using a physics-informed neural network’, *Phys. Fluids* **34**(1).
- Wang, J. X., Wu, J. L. & Xiao, H. (2017), ‘Physics-informed machine learning approach for

- reconstructing Reynolds stress modeling discrepancies based on DNS data', *Phys. Rev. Fluids* **2**(3), 034603.
- Wang, S., Teng, Y. & Perdikaris, P. (2021), 'Understanding and mitigating gradient flow pathologies in physics-informed neural networks', *SIAM J. Sci. Comput.* **43**(5), A3055–A3081.
- Wang, S., Wang, H. & Perdikaris, P. (2021), 'On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks', *Comput. Methods Appl. Mech. Eng.* **384**, 113938.
- Wang, S., Yu, X. & Perdikaris, P. (2022), 'When and why PINNs fail to train: A neural tangent kernel perspective', *J. Comput. Phys.* **449**, 110768.
- Weatheritt, J. & Sandberg, R. (2016), 'A novel evolutionary algorithm applied to algebraic modifications of the RANS stress–strain relationship', *J. Comput. Phys.* **325**, 22–37.
- Wilcox, D. C. (1988), 'Reassessment of the scale-determining equation for advanced turbulence models', *AIAA J.* **26**(11), 1299–1310.
- Wu, J. L., Wang, J. X. & Xiao, H. (2016), 'A Bayesian calibration–prediction method for reducing model-form uncertainties with application in RANS simulations', *Flow Turbul. Combust.* **97**(3), 761–786.
- Wu, J., Xiao, H., Sun, R. & Wang, Q. (2019), 'Reynolds-averaged Navier–Stokes equations with explicit data-driven Reynolds stress closure can be ill-conditioned', *J. Fluid Mech.* **869**, 553–586.
- Xiao, H. & Cinnella, P. (2019), 'Quantification of model uncertainty in RANS simulations: A review', *Prog. Aerosp. Sci.* **108**, 1–31.
- Xiao, H., Wu, J. L., Laizet, S. & Duan, L. (2020), 'Flows over periodic hills of parameterized geometries: A dataset for data-driven turbulence modeling from direct simulations', *Comput. Fluids* **200**, 104431.

- Xu, S., Sun, Z., Huang, R., Guo, D., Yang, G. & Ju, S. (2023), ‘A practical approach to flow field reconstruction with sparse or incomplete data through physics informed neural network’, *Acta Mech. Sin.* **39**(3), 322302.
- Yang, L., Meng, X. & Karniadakis, G. E. (2021), ‘B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data’, *J. Comput. Phys.* **425**, 109913.
- Yarlanki, S., Rajendran, B. & Hamann, H. (2012), Estimation of turbulence closure coefficients for data centers using machine learning algorithms, *in* ‘13th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems’, pp. 38–42.
- Zauner, M., Mons, V., Marquet, O. & Leclaire, B. (2022), ‘Nudging-based data assimilation of the turbulent flow around a square cylinder’, *J. Fluid Mech.* **937**, A38.
- Zeng, Q., Bryngelson, S. H. & Schäfer, F. (2022), ‘Competitive Physics Informed Networks’, *arXiv preprint arXiv:2204.11144* .
- Zhou, K. & Grauer, S. J. (2023), ‘Flow reconstruction and particle characterization from inertial Lagrangian tracks’, *arXiv preprint arXiv:2311.09076* .
- Zhu, T. (2023), The turbulent axisymmetric bluff body wake: analysis and control through experiment and simulation, PhD thesis, Imperial College London.

Appendix A

Additional equations and derivations

A.1 Spalart-Allmaras equations

The full definition of the SA transport equations, as mentioned in Equation (3.9), is described here.

A.1.1 Formulation of Spalart-Allmaras equations

The Spalart-Allmaras equations are solved for $\tilde{\nu}$. Then, the eddy viscosity, ν_t , is obtained as

$$\nu_t = \begin{cases} 0 & \text{if } \tilde{\nu} < 0, \\ f_{v1}\tilde{\nu} & \text{if } \tilde{\nu} \geq 0, \end{cases} \quad (\text{A.1})$$

where

$$f_{v1} = \frac{\chi^3}{\chi^3 + c_{v1}^3}, \quad \chi = \frac{\tilde{\nu}}{\nu}. \quad (\text{A.2})$$

The SA transport equation takes the form

$$U_j \frac{\partial \tilde{\nu}}{\partial x_j} - S_p - S_{diff} - S_c - S_d = 0, \quad (\text{A.3})$$

where S_p, S_{diff}, S_c, S_d are production, diffusion, cross-diffusion and destruction terms respectively. These are defined as

$$S_p = \begin{cases} c_{b1} S \tilde{\nu} g_n & \text{if } \tilde{\nu} < 0, \\ c_{b1} \tilde{S} \tilde{\nu} & \text{if } \tilde{\nu} \geq 0, \end{cases} \quad S_d = \begin{cases} c_{w1} \frac{\tilde{\nu}^2}{d^2} & \text{if } \tilde{\nu} < 0, \\ -c_{w1} f_w \frac{\tilde{\nu}^2}{d^2} & \text{if } \tilde{\nu} \geq 0, \end{cases} \quad (A.4)$$

$$S_{diff} = \frac{1}{\sigma} \nabla \cdot (\eta \nabla \tilde{\nu}) = \frac{1}{\sigma} \left(\eta \frac{\partial^2 \tilde{\nu}}{\partial x_j^2} + \frac{\partial \eta}{\partial x_j} \frac{\partial \tilde{\nu}}{\partial x_j} \right), \quad S_c = \frac{c_{b2}}{\sigma} \|\nabla \tilde{\nu}\|^2 = \frac{c_{b2}}{\sigma} \left(\frac{\partial \tilde{\nu}}{\partial x_j} \right)^2.$$

where the auxillary functions and constants are defined as

$$\eta = \begin{cases} \nu(1 + \chi + \frac{1}{2}\chi^2) & \text{if } \tilde{\nu} < 0, \\ \nu(1 + \chi) & \text{if } \tilde{\nu} \geq 0, \end{cases} \quad S' = \|\nabla \times \mathbf{U}\|, \quad S = \sqrt{S'^2 + M^2} - M,$$

$$\bar{S} = \frac{\tilde{\nu} f_{v2}}{\kappa^2 d^2}, \quad f_{v2} = 1 - \frac{\chi}{1 + \chi f_{v1}}, \quad \tilde{S} = \max(10^{-10}, S + \bar{S}),$$

$$r' = \frac{\tilde{\nu}}{\tilde{S} \kappa^2 d^2}, \quad r = \begin{cases} r' & \text{if } 0 \leq r' \leq 10, \\ 0 & \text{otherwise,} \end{cases} \quad g = r + c_{w2}(r^6 - r),$$

$$f_w = g \left(\frac{1 + c_{w3}^6}{g^6 + c_{w3}^6} \right)^{\frac{1}{6}}, \quad g_n = 1 - \frac{1000\chi^2}{1 + \chi^2}. \quad (A.5)$$

The constants in the SA transport equations are $c_{v1} = 7.1$, $c_{v2} = 0.7$, $c_{v1} = 0.9$, $\kappa = 0.41$, $\sigma = \frac{2}{3}$, $c_{b1} = 0.1355$, $c_{b2} = 0.622$, $c_{w1} = \frac{c_{b1}}{\kappa^2} + \frac{1 + c_{b2}}{\sigma}$, $c_{w2} = 0.3$, $c_{w3} = 2$, $M = 10^{-5}$.

A.1.2 Wall-Distance multiplied Spalart-Allmaras Equations

To remove singularities, which cause problems for the PINN-DA-SA formulation at points evaluated at (or near) the wall (as discussed in Section 6.2.1), the modified SA equations take the form

$$d^2 \left(U_j \frac{\partial \tilde{\nu}}{\partial x_j} - S_{diff} - S_c \right) - S_p^N - S_d^N = 0, \quad (A.6)$$

where S_p^N, S_d^N are new modified production and destruction terms, respectively. These are defined as

$$S_p^N = d^2 S_p = \begin{cases} c_{b1} S^N \tilde{\nu} g_n & \text{if } \tilde{\nu} < 0, \\ c_{b1} \tilde{S}^N \tilde{\nu} & \text{if } \tilde{\nu} \geq 0, \end{cases} \quad S_d^N = d^2 S_d = \begin{cases} c_{w1} \tilde{\nu}^2 & \text{if } \tilde{\nu} < 0, \\ -c_{w1} f_w \tilde{\nu}^2 & \text{if } \tilde{\nu} \geq 0. \end{cases} \quad (\text{A.7})$$

The modified auxiliary functions are thus

$$\begin{aligned} S^N &= d^2 S = d^2 (\sqrt{S'^2 + M^2} - M), & \bar{S}^N &= d^2 \bar{S} = \frac{\tilde{\nu} f_{v2}}{\kappa^2}, \\ \tilde{S}^N &= d^2 \tilde{S} = \max(10^{-10}, d^2 (S + \bar{S})) = \max(10^{-10}, S^N + \bar{S}^N), & r' &= \frac{\tilde{\nu}}{\tilde{S}^N \kappa^2}. \end{aligned} \quad (\text{A.8})$$

A.2 Harmonic-Balanced Navier-Stokes equations

A.2.1 Derivation of HBNS equations

In this section, the derivation of the HBNS equations from the Navier-Stokes is described in more detail.

Substitution of Fourier series expansion

Firstly, one can substitute the Fourier series expansion of the steady state flow (Eq. (3.10)), truncated to M resolved harmonics, into the Navier-Stokes (Eq. (3.3)). The time derivative and linear operator become

$$\mathbf{M} \partial_t \mathbf{w} = \mathbf{M} \partial_t \left(\sum_{-M \leq n \leq M} e^{in\omega t} \hat{\mathbf{w}}_{\mathbf{n}} \right) = \sum_{-M \leq n \leq M} in\omega \mathbf{M} e^{in\omega t} \hat{\mathbf{w}}_{\mathbf{n}}, \quad (\text{A.9a})$$

$$\mathbf{L} \mathbf{w} = \mathbf{L} \left(\sum_{-M \leq n \leq M} e^{in\omega t} \hat{\mathbf{w}}_{\mathbf{n}} \right) = \begin{pmatrix} \nu \sum_n e^{in\omega t} \Delta \hat{\mathbf{u}}_{\mathbf{n}} & \sum_n e^{in\omega t} \nabla \hat{p}_n \\ \sum_n e^{in\omega t} \nabla \cdot \hat{\mathbf{u}}_{\mathbf{n}} & 0 \end{pmatrix} = \sum_{-M \leq n \leq M} e^{in\omega t} \mathbf{L} \hat{\mathbf{w}}_{\mathbf{n}}. \quad (\text{A.9b})$$

Similarly, using the knowledge that $\frac{1}{2}\mathbf{N}(\mathbf{w}, \mathbf{w}) = \mathbf{u} \cdot \nabla \mathbf{u}$, substitution of the Fourier series expansion into the non-linear operator becomes:

$$\begin{aligned} \frac{1}{2}\mathbf{N}(\mathbf{w}, \mathbf{w}) &= \left(\sum_{-M \leq n_1 \leq M} e^{in_1 \omega t} \hat{\mathbf{u}}_{\mathbf{n}_1} \right) \cdot \left(\sum_{-M \leq n_2 \leq M} e^{in_2 \omega t} \nabla \hat{\mathbf{u}}_{\mathbf{n}_2} \right) \\ &= \sum_{n=n_1+n_2} \left(\sum_{-M \leq n_1 \leq M} \hat{\mathbf{u}}_{\mathbf{n}_1} \cdot \nabla \hat{\mathbf{u}}_{\mathbf{n}_2} \right) e^{in \omega t}, \end{aligned} \quad (\text{A.10})$$

where $n = n_1 + n_2$. Examining the coefficient term of the above series ($\sum_{-M \leq n_1 \leq M} \hat{\mathbf{u}}_{\mathbf{n}_1} \cdot \nabla \hat{\mathbf{u}}_{\mathbf{n}_2}$), and noting the symmetry of the non-linear operator, $\mathbf{N}$, this can be simplified to

$$\sum_{-M \leq n_1 \leq M} \hat{\mathbf{u}}_{\mathbf{n}_1} \cdot \nabla \hat{\mathbf{u}}_{\mathbf{n}_2} = \frac{1}{2} \sum_{-M \leq n_1 \leq M} \mathbf{N}(\hat{\mathbf{w}}_{\mathbf{n}_1}, \hat{\mathbf{w}}_{\mathbf{n}_2}). \quad (\text{A.11})$$

Additionally, by considering further symmetry of the condition, $n = n_1 + n_2$, one can apply another simplification by enforcing $n_1 \leq n_2$

$$\sum_{-M \leq n_1 \leq M} \hat{\mathbf{u}}_{\mathbf{n}_1} \cdot \nabla \hat{\mathbf{u}}_{\mathbf{n}_2} = \sum_{-M \leq n_1 \leq n_2 \leq M} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{\mathbf{n}_1}, \hat{\mathbf{w}}_{\mathbf{n}_2}). \quad (\text{A.12})$$

Lastly, terms related to $n_1, n_2 = 0$ are separated from the above series (A.12)), and noting that only M harmonics are resolved, the non-linear operator can be expressed as

$$\frac{1}{2}\mathbf{N}(\mathbf{w}, \mathbf{w}) = \sum_{-M \leq n \leq M} (\gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_{\mathbf{n}}) + \sum_{\substack{n=n_1+n_2 \\ -M \leq n_1 \leq n_2 \leq M \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{\mathbf{n}_1}, \hat{\mathbf{w}}_{\mathbf{n}_2})) e^{in \omega t}. \quad (\text{A.13})$$

Using the results from Eqs. (A.9) and (A.13) and substituting into Eq. (3.3), the Navier-Stokes equations become

$$\begin{aligned} \mathcal{NS} &\equiv \sum_{-M \leq n \leq M} ((in\omega \mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)) \hat{\mathbf{w}}_{\mathbf{n}} + \sum_{\substack{n=n_1+n_2 \\ -M \leq n_1 \leq n_2 \leq M \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{\mathbf{n}_1}, \hat{\mathbf{w}}_{\mathbf{n}_2})) e^{in \omega t} \\ &= 0 \end{aligned} \quad (\text{A.14})$$

Fourier-Galerkin projection

Firstly assume the flow solution, $\mathbf{w}(\mathbf{t})$ is periodic over some period, T . Using the Galerkin method of weighted residuals on Eq. (A.14) (i.e. $\mathcal{NS}$) the final HBNS equations can be derived, by multiplying $\mathcal{NS}$ by some (orthogonal) Fourier basis functions and integrating over the period

$$\int_0^T \sum_{-M \leq n \leq M} \mathcal{NS} e^{-in\omega t} dt = \sum_{-M \leq n \leq M} \hat{\mathbf{r}}_{\mathbf{n}}, \quad (\text{A.15})$$

where $\hat{\mathbf{r}}_{\mathbf{n}}$ is some residual term. Requiring that the residual, $\hat{\mathbf{r}}_{\mathbf{n}} = 0$ for all n , then

$$\begin{aligned} & \sum_{-M \leq m \leq M} [(in\omega \mathbf{M} + \mathbf{L} + \gamma_0^m \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)) \hat{\mathbf{w}}_{\mathbf{m}} \\ & + \sum_{\substack{m=m_1+m_2 \\ -M \leq m_1 \leq m_2 \leq M \\ m_1, m_2 \neq 0}} \gamma_{m_1}^{m_2} \mathbf{N}(\hat{\mathbf{w}}_{\mathbf{m}_1}, \hat{\mathbf{w}}_{\mathbf{m}_2})] \int_0^T \sum_{-M \leq n \leq M} e^{im\omega t} e^{-in\omega t} dt = \sum_{-M \leq n \leq M} \hat{\mathbf{r}}_{\mathbf{n}} = 0. \end{aligned} \quad (\text{A.16})$$

Due to orthogonality, $m = n$ and thus the final HBNS equations, as seen in Eq. (3.11) (and repeated here), can be defined.

$$(in\omega \mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \bullet)) \hat{\mathbf{w}}_{\mathbf{n}} + \sum_{\substack{n=n_1+n_2 \\ -M \leq n_1 \leq n_2 \leq M \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{\mathbf{n}_1}, \hat{\mathbf{w}}_{\mathbf{n}_2}) = 0. \quad (\text{A.17})$$

A.2.2 Division of low-frequency resolved and high-frequency unresolved scales for one resolved harmonic

As described in Section 3.1.2, one can divide the HBNS equations into low- and high-frequency dynamics. To demonstrate this, consider the HBNS equations, Eq. (3.16), truncated to $N = 2$ harmonics,

$$\begin{aligned} & \mathbf{L} \hat{\mathbf{w}}_0 + \frac{1}{2} \mathbf{N}(\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_0) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_1) + \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_2) \\ & + [\mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_3) + \mathbf{N}(\overline{\hat{\mathbf{w}}_4}, \hat{\mathbf{w}}_4) + \dots] = 0, \end{aligned} \quad (\text{A.18a})$$

$$\begin{aligned}
& (i\omega\mathbf{M} + \mathbf{L}) \hat{\mathbf{w}}_1 + \mathbf{N}(\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_1) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_2) \\
& + [\mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_4) + \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_3) + \dots] = 0,
\end{aligned} \tag{A.18b}$$

$$\begin{aligned}
& (2i\omega\mathbf{M} + \mathbf{L}) \hat{\mathbf{w}}_2 + \mathbf{N}(\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_2) + \frac{1}{2}\mathbf{N}(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1) \\
& + [\mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_4) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_3) + \dots] = 0,
\end{aligned} \tag{A.18c}$$

where the terms in $[\bullet]$ are the high-frequency truncated harmonics. One can see the effect of the non-linear operator, $\mathbf{N}$ and importantly how the high-frequency unresolved harmonics interact with the resolved low-frequency harmonics. By comparing Eq. (A.18) with the low-frequency projection from Eq. (3.17), one can describe and categorise the non-linear interactions.

$$\begin{aligned}
& \mathbf{L}\hat{\mathbf{w}}_0 + \underbrace{\frac{1}{2}\mathbf{N}(\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_0) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_1) + \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_2)}_{\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1]} \\
& + \left[\underbrace{\mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_3) + \mathbf{N}(\overline{\hat{\mathbf{w}}_4}, \hat{\mathbf{w}}_4) + \dots}_{\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h]} \right] = 0,
\end{aligned} \tag{A.19a}$$

$$\begin{aligned}
& (i\omega\mathbf{M} + \mathbf{L}) \hat{\mathbf{w}}_1 + \underbrace{\mathbf{N}(\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_1) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_2)}_{\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1]} \\
& + \left[\underbrace{\mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_4)}_{\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h]} + \underbrace{\mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_3)}_{\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h] + \mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_1]} + \dots \right] = 0,
\end{aligned} \tag{A.19b}$$

$$\begin{aligned}
& (2i\omega\mathbf{M} + \mathbf{L}) \hat{\mathbf{w}}_2 + \underbrace{\mathbf{N}(\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_2) + \frac{1}{2}\mathbf{N}(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1)}_{\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1]} \\
& + \left[\underbrace{\mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_4) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_3) + \dots}_{\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h] + \mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_1]} \right] = 0.
\end{aligned} \tag{A.19c}$$

One can see the interaction of i) low-frequency resolved scales on low-frequency resolved scales ($\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1]$), ii) high-frequency unresolved scales on high-frequency unresolved scales ($\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h]$), and iii) low-frequency resolved scales on high-frequency unresolved scales and vice versa ($\mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_h] + \mathcal{N}_{\mathcal{P}}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_1]$). These latter two interactions are captured in the residual term from Eq. (3.18), $\mathbf{R}_n$, as described in Section 3.1.2.

Whilst the above is shown for $N = 2$ resolved harmonics (and unresolved scales shown up till

$N = 4$), it can be repeated for any order of truncation.

Appendix B

Additional results for use of turbulence model augmented PINN-DA-SA

B.1 Hyper-parameter and loss weight selection

B.1.1 Tuning the model with a grid search approach

To select hyper-parameters (model architecture, activation function and learning rate) and loss weighting, a grid search method was used. Tables B.1 and B.2, show the sensitivity of final model loss to hyper-parameters. From these results, the architecture, as discussed in Section 6.2.1, was selected. A similar analysis was completed for the loss weighting. A summary of the key results are shown in Table B.3. Repeats of the final model were found to converge within 3.0% of the loss presented in Section 6.3 (after 3 repetitions).

Table B.1: Sensitivity of final model loss function, C , to activation functions and learning rate on a model with 7 layers and 50 nodes/layer relative to final choice (tanh and 5×10^{-4} learning rate)

Activation Functions					Learning Rate (with tanh activation)		
tanh	ReLU	sin	ELU	L-LAAF tanh	1×10^{-3}	5×10^{-4}	1×10^{-4}
-	+55.4%	+9.1%	+69.3%	+5.1%	-1.3%	-	+3.6%

Table B.2: Sensitivity on final model loss function, C , to architecture parameters (number of layers and nodes) relative to final architecture (7 layers, 50 nodes).

		Number of Layers				
		5	6	7	8	9
Number of Nodes/Layer	25	+19.7%	+15.6%	+9.9%	+7.5%	+5.2%
	50	+7.4%	+5.7%	-	+2.8%	+2.6%
	75	+6.3%	+5.5%	+2.8%	+4.7%	+6.3%

Table B.3: Sensitivity on final model loss function, C , to perturbations on weights from Table 6.1

λ^P		λ^D		λ^B	
$\times 10$	$\times 0.1$	$\times 10$	$\times 0.1$	$\times 10$	$\times 0.1$
+4.7%	+8.6%	+3.8%	+9.3%	+1.6%	+7.2%

A tanh activation was selected due to the lowest value of cost function. Whilst an initial learning rate of 1×10^{-3} demonstrated a slightly lower final loss value, this model was less repeatable than the chosen learning rate (5×10^{-4}). Similarly, the model architecture (nodes and layers) was selected for demonstrating the lowest final loss value. Whilst adding more layers and nodes performs similarly, the additional model complexity increases computational cost.

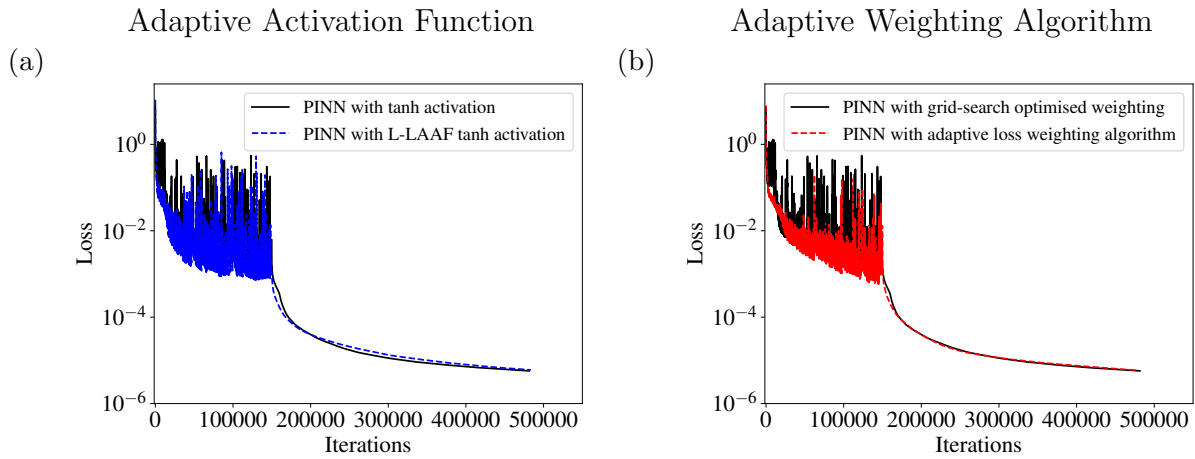


Figure B.1: Comparing the effect of convergence for PINN-DA-SA by applying (a) layerwise locally adaptive activation functions (L-LAAF) (b) an adaptive loss weighting algorithm.

B.1.2 Alternative tuning methodologies

Whilst a grid search approach was used to tune hyperparameters in this work, there have been many papers investigating the use of more robust methods. In this section, two approaches are investigated: adaptive activation functions and adaptive weights.

Adaptive activation functions, such as the layerwise locally adaptive activation function (L-LAAF) proposed in (Jagtap, Kawaguchi & Karniadakis 2020), have shown promise at both accelerating convergence of PINNs, as well as improving model accuracy. For every layer, L-LAAF introduces a learnable tuning parameter, a into the activation function to scale its gradient:

$$\sigma^*(x) = \sigma(ax), \quad (\text{B.1})$$

where σ^*, σ are the adaptive and base activation functions, respectively and x is some input into the activation function. The effect of applying L-LAAF (with tanh) to the periodic hill problem is presented in Table B.1 and Figure B.1(a). For this problem, there was a minimal effect on the converged loss value. There does appear to be some improvement in the ‘spikiness’ seen in the initial ADAM phase. However, during the second, long L-BFGS-B phase, this is no longer seen and the convergence speed is almost identical.

Adaptive weighting algorithms have been an increasingly promising area of research to make robust and reduce the difficulty of model weight tuning process. This is a particular challenge

for PINNs, where there are often many loss terms in the equations. Figure B.1(b) demonstrates the advantage of using the adaptive loss weighting algorithm, as implemented in (Sliwinski & Rigas 2022). Whilst the difference in the final loss between the model with optimised weights (from Table 6.1) and the model with adaptive weighting is minimal, the latter of these avoids all the intermediate grid-search steps used to find and select the optimal weights. This massively reduces the computational cost of finding the correct weighting and is thus the recommended approach moving forward.

B.2 Modification to PINNs for Spalart-Allmaras augmentation

Augmenting PINNs with the SA turbulence model introduced additional complexity, which required modifications to the equation formulation and neural network architecture to achieve convergence. Specifically, two key problems had to be solved for successful application of PINN-DA-SA: i) singularity in the SA transport equation at the wall (also mentioned in (Molnar et al. 2023)); ii) the effect of clipping negative values of $\tilde{\nu}$.

B.2.1 Singularity of the SA transport equation at the wall

As described in Section 6.2, evaluating the SA transport equation at the wall results in a singularity, which poses a problem for convergence of PINN-DA-SA, as shown in Figure B.2. Whilst the singularities can be avoided by not evaluating PDE loss at the wall, there remains a large imbalance between SA transport equation loss and other loss terms (such as x-momentum) in the near wall region, where large values of $\frac{1}{d^2}$ occur, which subsequently affect PINN-DA-SA optimisation. This can be seen by comparing residual loss when enforcing the enforcing the normal SA transport equation, seen in Figure B.2(c,d). After 1000 iterations, the SA transport residual approaches magnitudes of 10^7 near the wall (Figure B.2(d)) compared with approximately 10^{-1} for the corresponding x -momentum loss (Figure B.2(c)). This leads to a

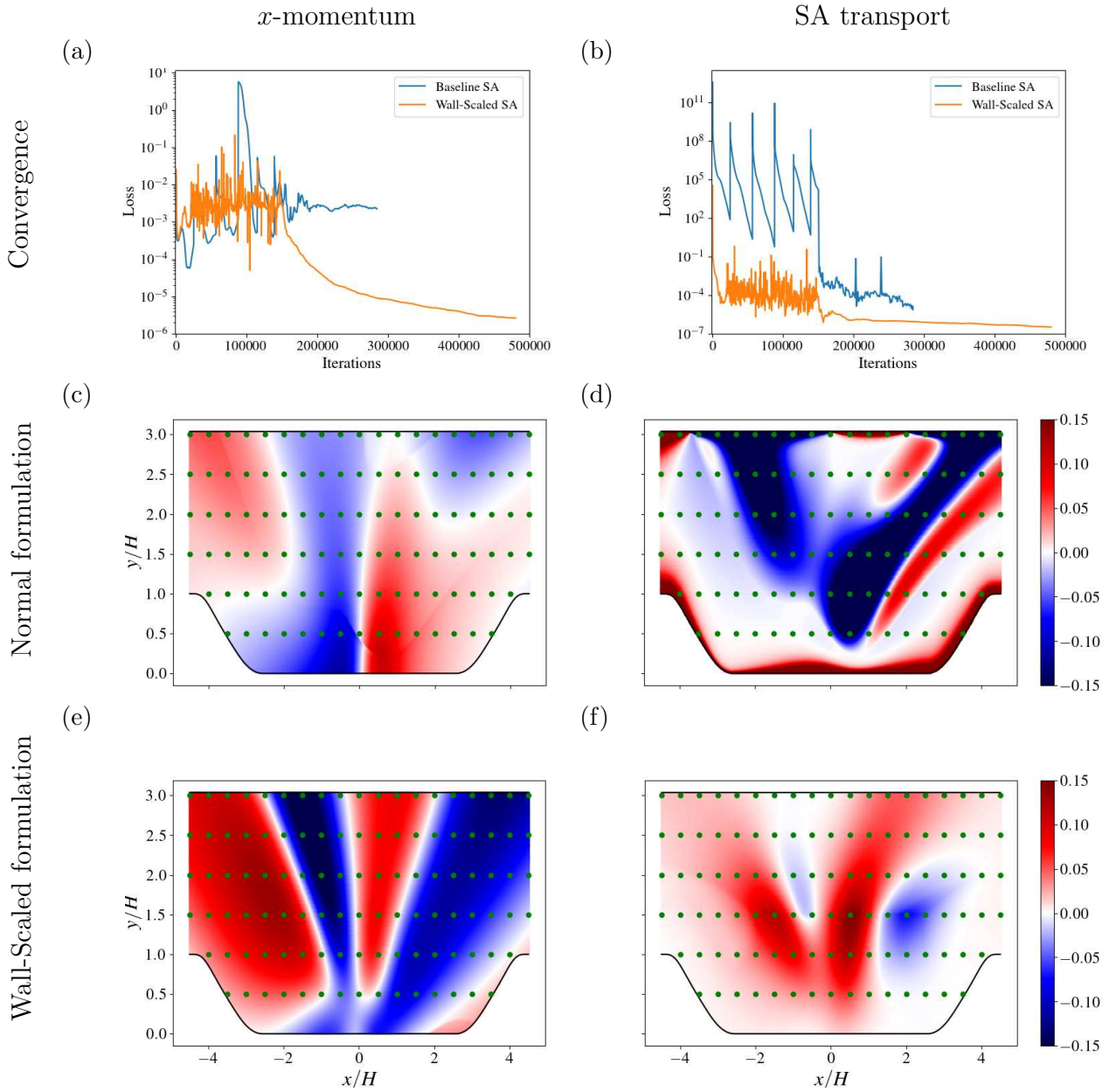


Figure B.2: Optimisation of PINN-DA-SA when using (middle) normal formulation of the SA transformation and (bottom) wall-scaled formulation (i.e. $\times d^2$). (a,b) convergence of x -momentum and SA transport loss using both formulations of the SA transport equation. (c,e) x -momentum residual field after 1000 iterations, using either the normal or wall-scaled SA respectively. (d,f) SA transport residual field after 1000 iterations, using either the normal or wall-scaled SA respectively.

poorer and more unstable convergence, as seen in Figure B.2(a,b). The poor convergence and loss imbalance when enforcing the normal SA transport equation (see Appendix A.1.1) leads to the solution of pseudo eddy viscosity, $\tilde{\nu}$, tending towards zero, leading to the same formulation and reconstruction accuracy of the PINN-DA-Baseline approach.

By multiplying the SA transport equation with d^2 to eliminate the singularity (see Appendix A.1.2), the singularities are removed and better scaling of the individual loss components is achieved, as see in Figure B.2(e,f). Whilst the x momentum residual field in Figure B.2(e) when enforcing wall-scaled SA is comparable to enforcing normal SA (see Figure B.2(c)), after 1000 iterations, the SA transport residual loss itself is now a similar order of magnitude to momentum (see Figure B.2(f)), achieving a 10^9 orders of magnitude reduction compared with Figure B.2(d). This is reflected in the convergence from Figure B.2(a,b). Enforcing the wall-scaled SA formulation, both stabilises the optimisation and achieves a lower final loss. Use of wall-scaled SA transport equation, both eliminates the singularities and better balances the individual loss components, leading to a better posed optimisation and thus successful convergence of PINN-DA-SA.

B.2.2 Clipping negative values of $\tilde{\nu}$

As with the previous section, admitting negative values of $\tilde{\nu}$ (even when the negative SA equation formulation is enforced), leads to a more complex optimisation problem. When negative values of $\tilde{\nu} \leq 0$ are clipped (and set to 0), the PINN-DA-SA converges to the trivial solution i.e. $\nu_t \approx 0$, and therefore becomes equivalent to the PINN-DA-Baseline formulation.

Using the PINN to output $\sqrt{\tilde{\nu}}$, and then squaring the output, enforces positive values of $\tilde{\nu}$, leading to better convergence of PINN-DA-SA. Removing clipping from the SA formulation, leads to a more stable optimisation (similar to the previous section), which is better posed, resulting in better convergence of the PINN-DA-SA.

B.3 Effect of regularisation on data assimilation techniques

As mentioned in Section 3.2, the forcing decomposition between modelled and corrective parts is not unique and thus the regularisation of corrective forcing is critical. This section will show

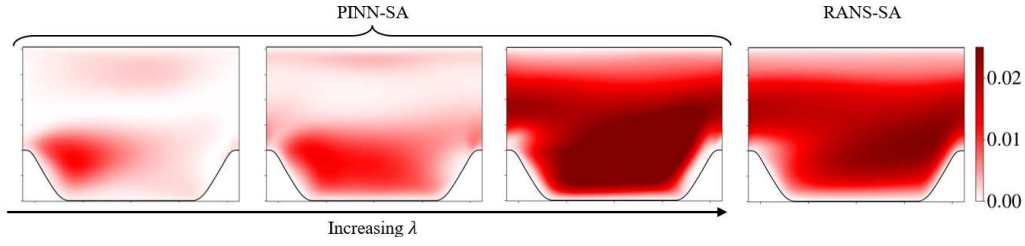


Figure B.3: Effect of regularisation parameter (solenoidal forcing magnitude penalisation), λ^R on PINN-SA solution for $\tilde{\nu}$. $\tilde{\nu}$ for increasing λ^R compared with the RANS-SA $\tilde{\nu}$ shows that increasing regularisation causes eddy viscosity to converge towards the RANS solution. The presented values of λ^R are 10^{-8} , 10^{-4} , 10^0 left to right

the effect of regularisation for both PINN-DA and variational approaches and highlight its importance.

B.3.1 Effect of solenoidal forcing regularisation on PINNs

For PINN-DA-SA, an L_2 penalisation of the solenoidal forcing magnitude was applied. During PINN-DA-SA hyper-parameter tuning, it was found that the final solution was strongly influenced by the regularisation of solenoidal forcing. Clearly as regularisation weight, λ^R , increases, the magnitude of solenoidal forcing is penalised more and thus there is a greater dependence on the modelled eddy-viscosity forcing component. In effect, the greater the value of λ , the closer the governing equations tend towards a pure RANS-SA problem. However, for the PINN-DA, enforcement of data loss at measurement points results in differences. This is demonstrated in Figure B.3. As solenoidal forcing magnitude is penalised more and more, the solution tends towards RANS-SA. However, this is an imperfect solution as the PINN-DA-SA is attempting to fit high-fidelity (mean) DNS data to the RANS-SA equations. This is most notable at the point around the rear hill apex. As a result, the residual PDE loss begins to increase at high λ^R . On the other hand, without penalisation ($\lambda^R = 0$), the PINN-DA-SA solution tends towards the PINN-DA-Baseline solution with a negligible contribution from the eddy-viscosity component ($\tilde{\nu} \approx 0$).

As the data coarsens, and thus the dependence on data reduces, the requirement for a modelled component becomes more important, as the corrective Reynolds forcing can be closed at fewer

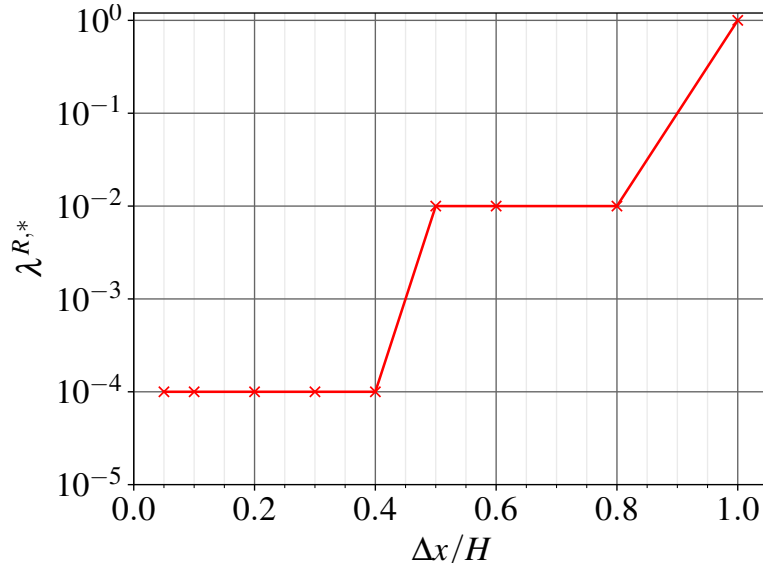


Figure B.4: The optimised regularisation parameter, $\lambda^{R,*}$, with respect to data resolution.

points using data. Less data means the solution will move closer to a RANS-SA solution. As a result, the coarser the data, the greater the value for the optimal λ^* , as the modelled forcing component becomes a greater component of the forcing. This is seen in Figure B.4.

B.3.2 Importance of regularisation for variational DA

For the variational DA, the choice of regularisation (of solenoidal forcing $f_{s,i}$) was more important. Three approaches were trialled: no regularisation; L_2 regularisation; H_1 regularisation.

Firstly, without regularisation, whilst data is well enforced (as seen in Figure B.5(a)) even compared with the final H_1 based approach, it results in a non-physical solenoidal forcing that appears strongly pointwise, clearly seen in Figure B.5(b), as large corrections are applied at the data locations. The final reconstruction is thus less physical. Additionally, regularisation is needed to enforce convergence to a unique decomposition.

For the L_2 regularisation approach, error is higher than either H_1 or no regularisation methods. This is due to the inefficiency of the approach in the variational framework. To achieve sufficient smoothing of the solution, avoiding the problems of the none regularised approach, large penalisation weighting is needed. However, this comes at the cost of poor data enforcement,

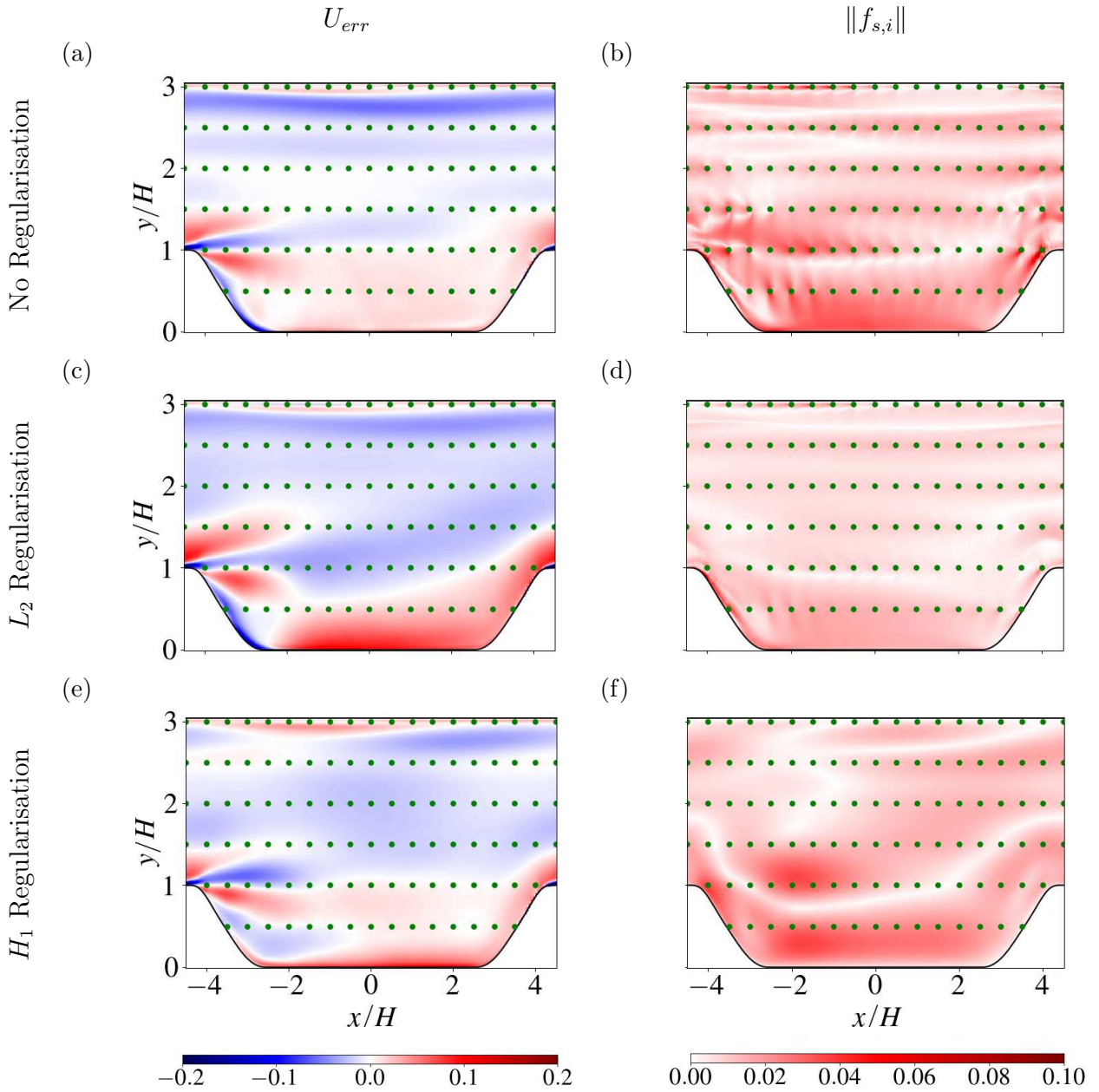


Figure B.5: Absolute U error (a,c,e) and corrective solenoidal forcing, $\|f_{s,i}\|$, (b,d,f) for variational-DA-SA method using different regularisation approaches: No regularisation (top), L_2 regularisation (middle) and H_1 regularisation (bottom).

as seen in Figure B.5(c).

For the variational-DA-SA, H_1 regularisation was selected to smooth out the solution. As demonstrated, for the variational approach, H_1 regularisation is a much more suitable approach. It allows much better matching of the data (as compared with L_2 regularisation), whilst also giving a more physical solution. As a result this was selected as the final regularisation method.

For PINN-DA-SA, good data enforcement and a smooth non-pointwise solenoidal forcing was always ensured. Regularisation is exclusively used to control the distribution between modelled and corrective forcing. Given its computational simplicity, L_2 was selected for PINN-DA-SA. This was not the case for variational-DA-SA and regularisation was needed to smooth the solution in addition to the distribution of modelled and corrective forcing.

Appendix C

Additional information for HBNS-constrained PINN-DA

C.1 Hyper-parameter selection

The selection of hyper-parameters had a substantial effect on the convergence of PINN-DA-HBNS. The following results show some of the isolated effects of hyper-parameters during the training process.

C.1.1 Activation Functions

Firstly, the effect of activation function choice is examined. In Chapter 6, a tanh activation function was used. This is a common choice for PINNs due to smoothness and differentiability. However, tanh activation functions can suffer from the “vanishing gradient” problem. Alternatively, the swish activation was tested, as presented in Figure C.1(a). One can see that use of swish activation results in a decrease of total loss by one order of magnitude compared with tanh activation.

Secondly, as discussed in Appendix B.1, use of adaptive activation can improve or accelerate convergence of PINN-DA. Whilst in Appendix B.1, the adaptive activation had a negligible

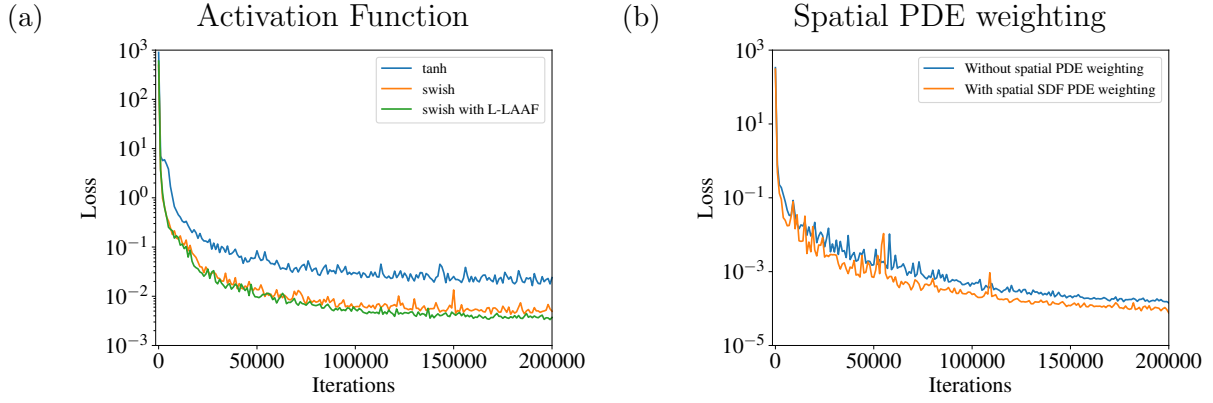


Figure C.1: Effect of hyper-parameter selection on convergence of PINN-DA-HBNS. (a) Effect of activation function choice comparing tanh, swish and swish with layerwise locally adaptive activation functions (L-LAAF). (b) Effect of spatial (SDF) weighting of residual PDE loss.

effect, for the PINN-DA-HBNS and using the NVIDIA Modulus implementation (as opposed to DeepXDE), swish with L-LAAF adaptive activation (Jagtap, Kawaguchi & Karniadakis 2020) results in a further reduction in total loss.

C.1.2 PDE weighting

The choice of residual loss weighting can also affect convergence of PINN-DA. For the PINN-DA-HBNS, it was shown that use of a spatial weight distribution function (as described in Sharma et al. (2023)), improved convergence, as presented in Figure C.1(b). The residual PDE loss, weighted by distance to the wall (SDF), reduces the contribution of PDE loss in the near wall region to the total loss. This subsequently prevents high PDE loss from high near wall flow gradients having too large of an effect on the total loss and thus convergence of PINN-DA-HBNS.

C.1.3 Architecture

In this chapter, the PINN-DA-HBNS architecture is formed from three separate neural networks, each mapping either velocity modes ($\hat{\mathbf{u}}_{\mathbf{n}}$), pressure fields ($((\hat{p} - \phi)_n)$) or corrective, residual forcing ($\mathbf{R}_{\mathbf{n},s}$). As seen across wider literature, by separating the different flow variables into different neural networks, one is able to improve the ability to capture the different scales and

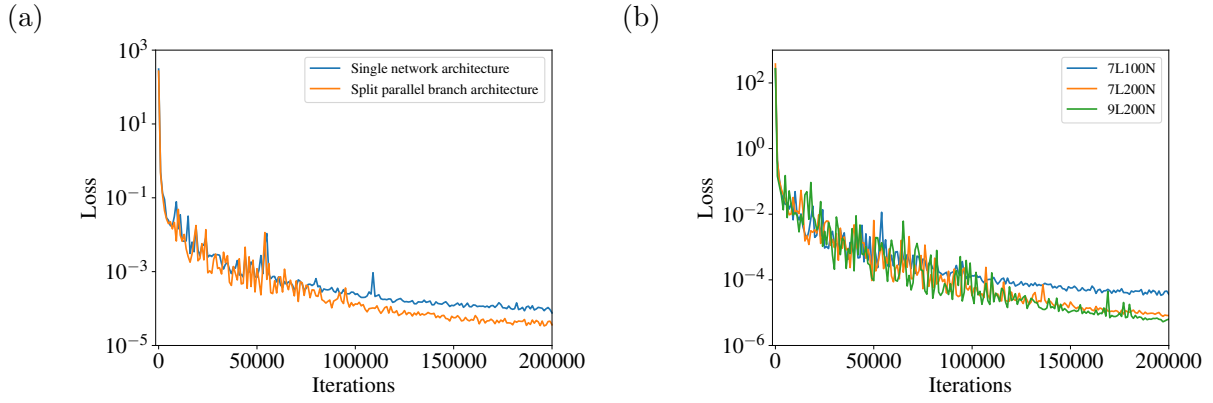


Figure C.2: The effect of PINN-DA-HBNS architecture on loss convergence showing the impact of (a) split parallel branches (as used in Chapter 7) and (b) different number of nodes and layers.

magnitudes of the different variables, improving convergence of total loss, as demonstrated in Figure C.2(a).

Finally, the number of neural network nodes and layers also affects the representative capacity of PINN-DAs as presented in Figure C.2(b). The three presented architectures (layers and node numbers) show that PINN-DA-HBNS was more sensitive to model width (nodes per layer) compared with model depth (number of layers).

C.2 PINN-DA-HBNS reconstructed fields

The PINN-DA-HBNS reconstructed velocity fields, both mean velocity and first harmonic strongly resemble the true fields (presented in Figure 4.3), irrespective of number of resolved harmonics, as seen in Figure C.3.

C.3 $N = 2$ HBNS closure model

Whilst an $N = 2$ HBNS closure model was successfully built, the relatively small magnitude of the residual closure terms (e.g. $\mathbf{R}_0$) is comparable to the noise and uncertainty of the data-driven closure model prediction. However, the results have been included here for completeness. The difference in training loss when building an $N = 2$ closure model, as presented in Figure

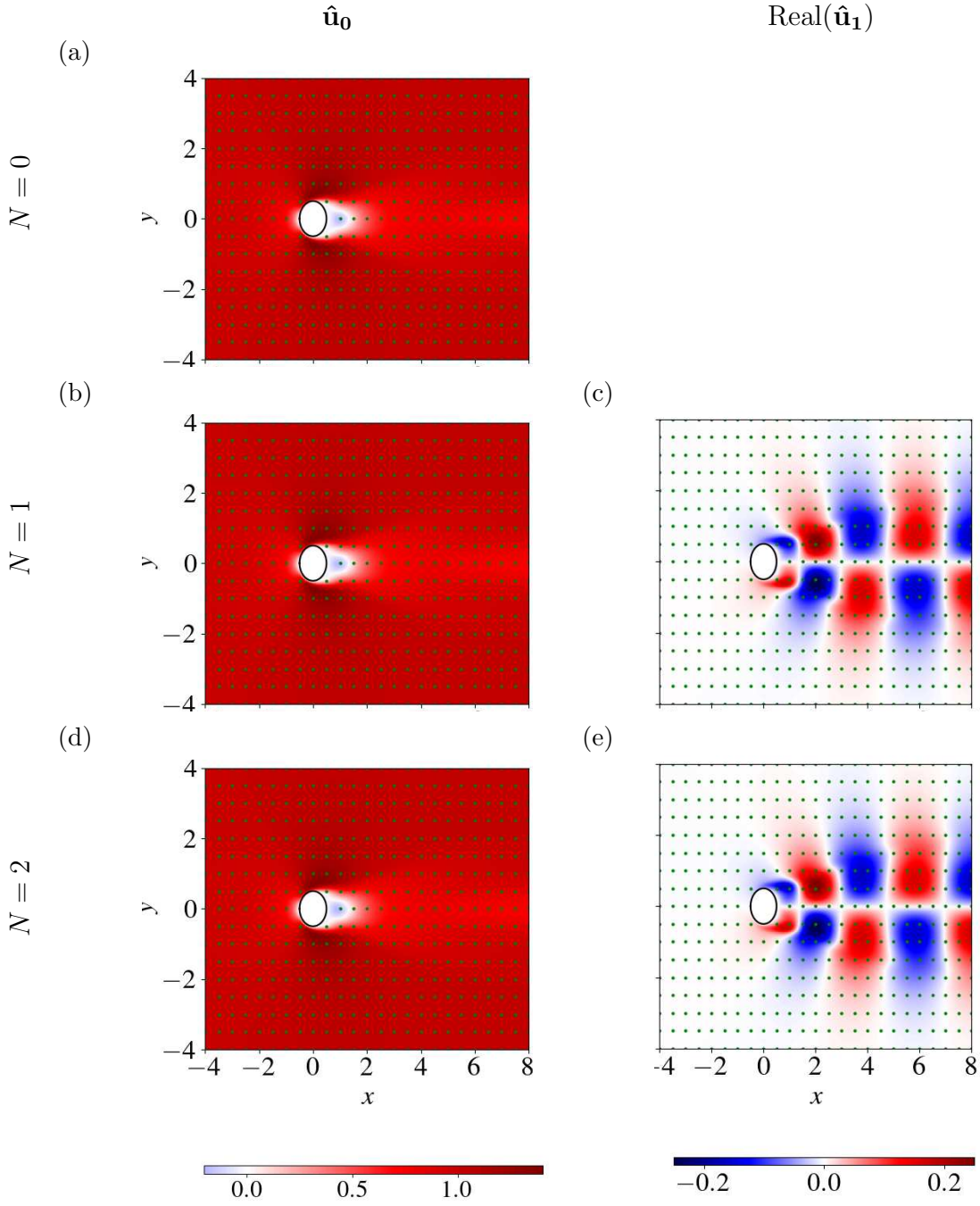


Figure C.3: Flow reconstruction for streamwise component of (a,b,d) mean flow ($\hat{\mathbf{u}}_0$) and (c,e) first velocity harmonic ($\hat{\mathbf{u}}_1$) for different number of resolved low-frequency harmonics, $N = 0, 1, 2$.

C.4, compared with an $N = 1$ closure model is significantly smaller than comparing an $N = 0$ to $N = 1$ HBNS closure model (shown in Section 7.3.2. This is consistent with the results from Section 7.2, where resolving $N = 2$ low-frequency harmonics had a much smaller effect than resolving $N = 1$ harmonics.

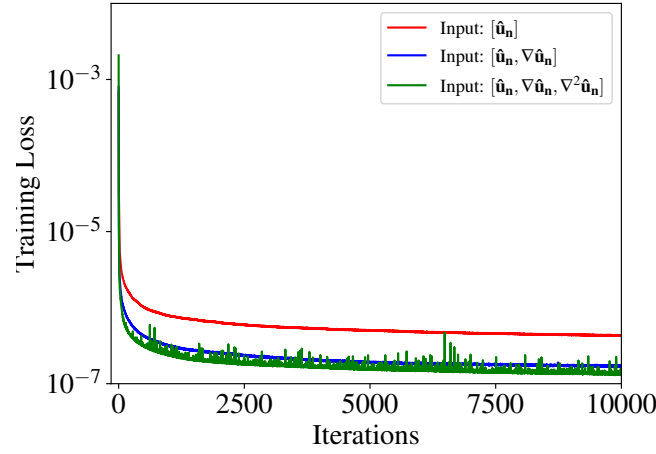


Figure C.4: Convergence of training loss during optimisation of closure model for $N = 2$ resolved harmonics.

The $N = 2$ closure model, presented in Figure C.5 shows similar trends in performance to the $N = 1$ closure model (e.g. effect of resolved velocity input features). However, one can clearly observe “noisiness” in the data-driven closure model prediction, attributed to the relatively small magnitude of the $N = 2$ closure term.

The streamwise error distribution from the $N = 2$ HBNS closure model, presented in Figure C.6, is also comparable to the $N = 1$ closure model (Figure 7.13(b)). As with other results, the effect of changing from an $N = 0$ HBNS closure model to $N = 1$ model is far more significant than the subsequent use of an $N = 2$ closure model.

In conclusion, whilst an $N = 2$ HBNS closure model has been constructed with reasonable accuracy, the small magnitude of the closure term results in a higher relative error from the data-driven closure model. In order to improve the accuracy of the $N = 2$ closure model techniques, such as input normalisation may need to be employed, in addition to application on cases where resolving the second harmonic has a larger effect on the flow.

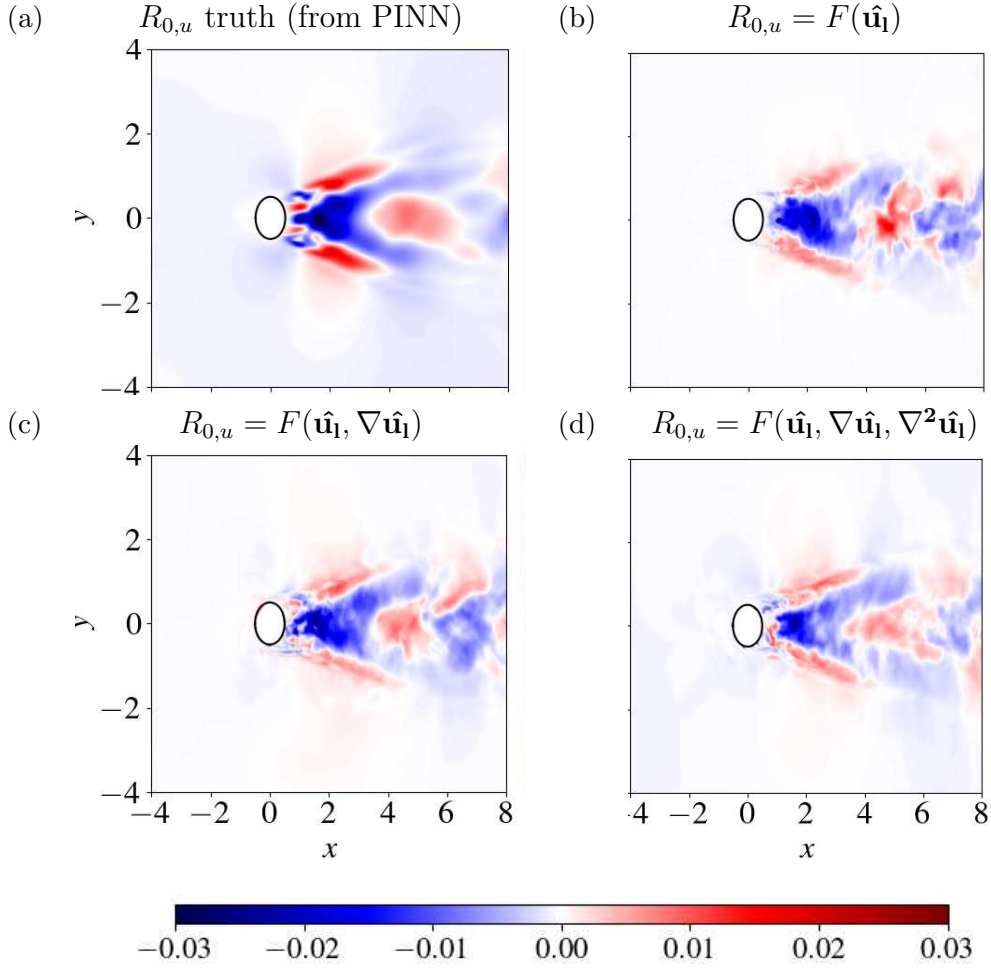


Figure C.5: True (PINN-DA-HBNS inferred) (a) and closure model predictions for $R_{0,u}$ field when trained with (b) velocity variables ($\mathbf{R}_n = F(\hat{\mathbf{u}}_1)$), (c) velocity gradients ($\mathbf{R}_n = F(\hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_1)$), (c) velocity laplacian ($\mathbf{R}_n = F(\hat{\mathbf{u}}_1, \nabla \hat{\mathbf{u}}_1, \nabla^2 \hat{\mathbf{u}}_1)$). This is compared for $N = 2$ resolved harmonics. Predictions are for $Re = 300$ after training on $Re = 150, 200$.

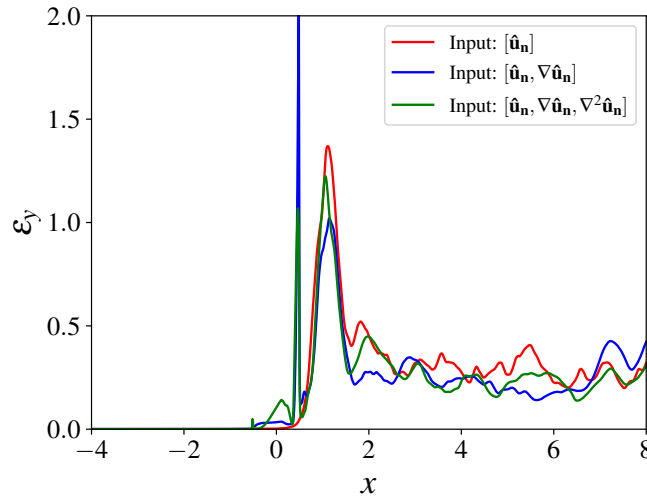


Figure C.6: Streamwise error distribution for predicted residual closure terms $\mathbf{R}_n$, for HBNS closure model resolved with $N = 2$ low-frequency harmonics. This is evaluated for the testing case, $Re = 300$.

Appendix D

Additional results for application of PINN-DA on experimental turbulent bluff wake

D.1 Flow reconstruction without embedded transformation approach

In this section the reconstruction results for PINN-DA, using data spacings $\Delta L = 0.2, 0.5$, without the use of the embedded transformation approach have been presented in Figures D.1 and D.2, to show the effect of the embedded transformation on reconstruction.

The $\Delta L = 0.2$ case highlights how in the shear layer, error magnitude approaches 50% free-stream velocity (compared to the worst performing PINN-DA with embedded transformation having maximum error of 30% free-stream velocity as seen in Figure 8.11). At the coarser spacing (Figure D.2), the shear layer remains the highest error region although greater error occurs in the recirculation region. The PINN-DA reconstruction without embedded transformation demonstrates how, near the trailing edge of the bluff body, there is a high sensitivity caused by misaligned measurements, which has a peak error magnitude at least 20% free-stream velocity

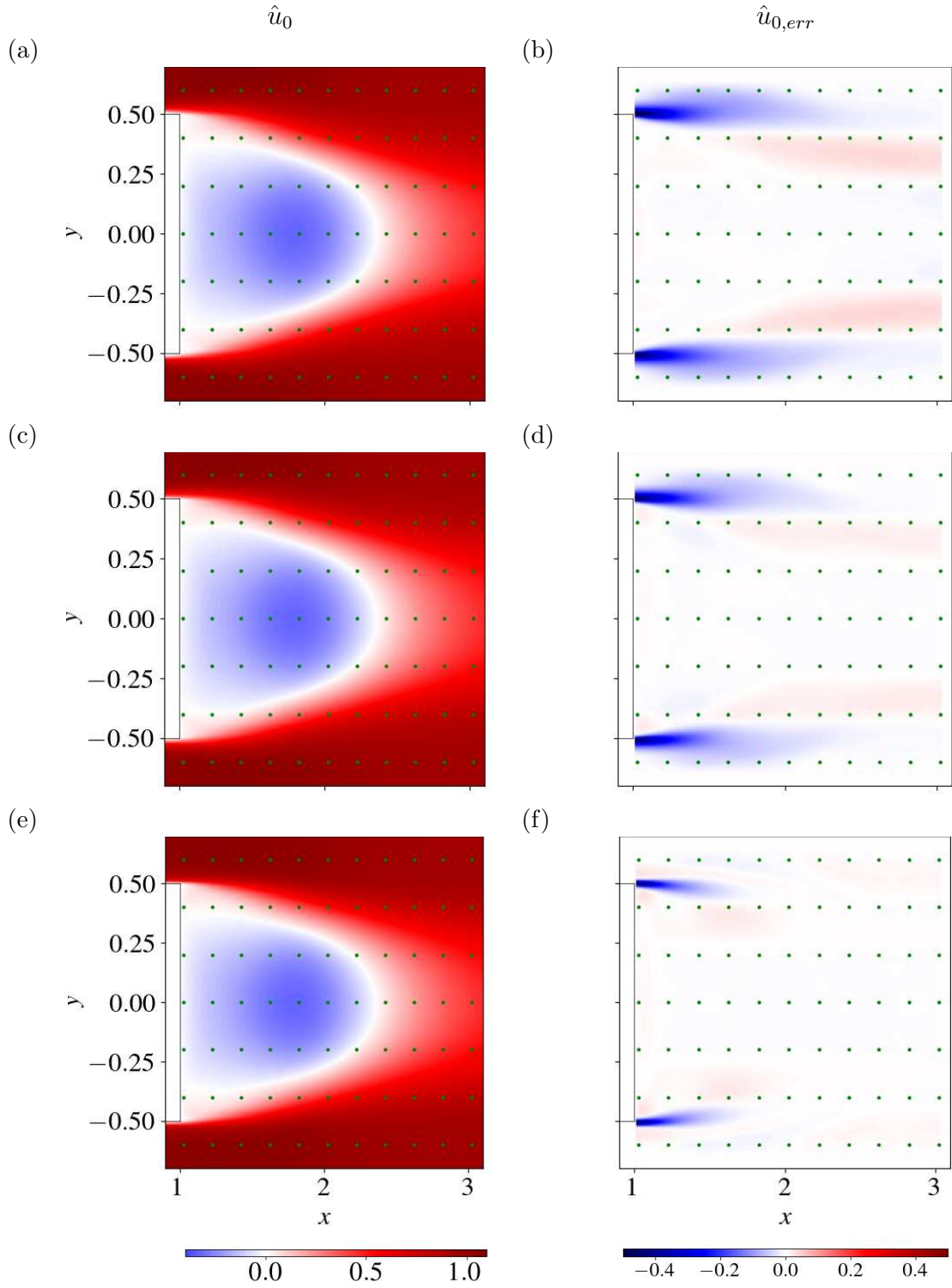


Figure D.1: PINN reconstruction using different methodologies for data spacing $\Delta L = 0.2$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-SA.

higher.

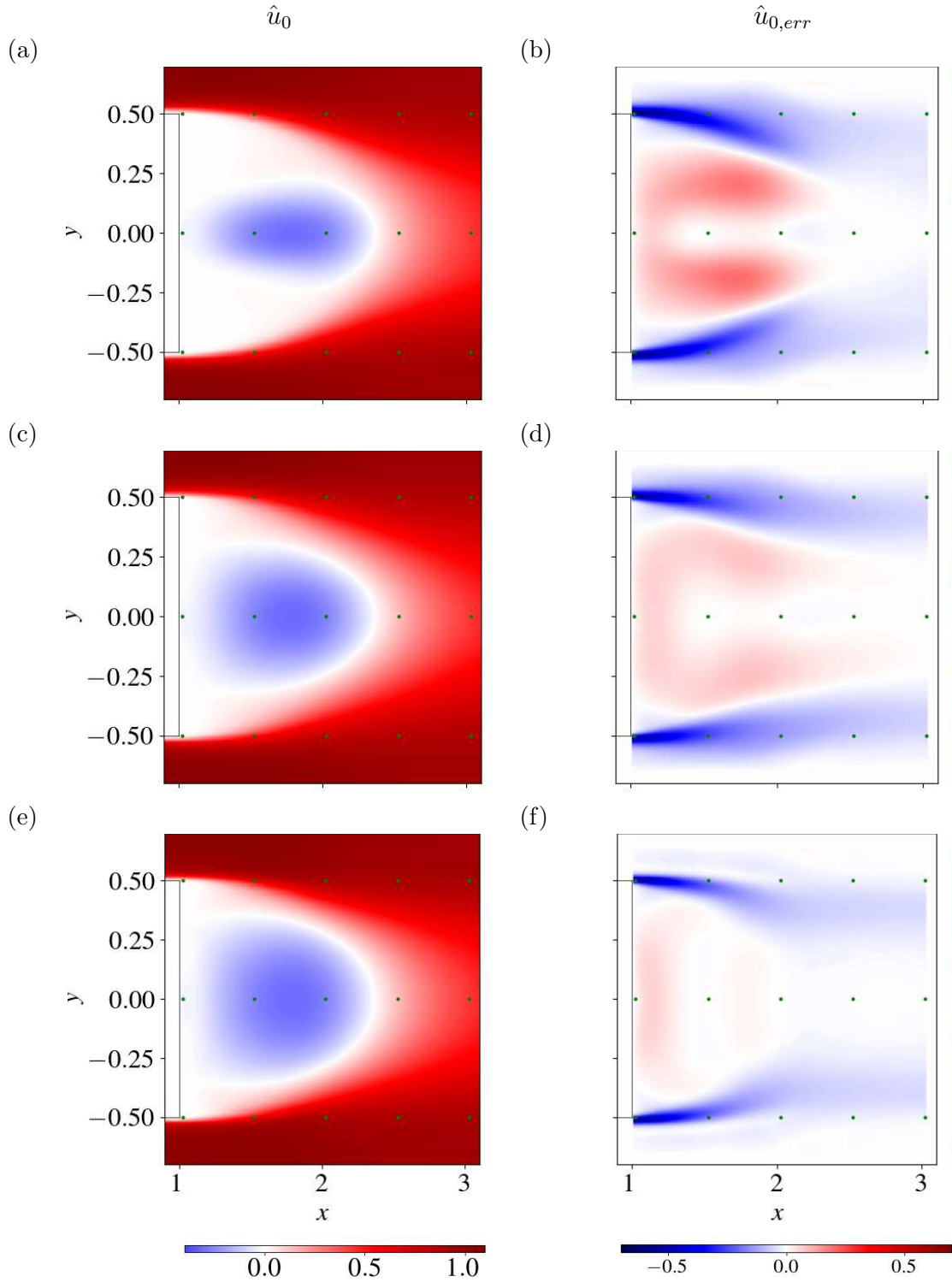


Figure D.2: PINN reconstruction using different methodologies for data spacing $\Delta L = 0.5$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-SA.

D.2 Flow reconstruction (with embedded transformation) of bluff body wake with $\Delta L = 0.5$ data spacing

In this section, the mean-flow reconstruction using the transformation embedded PINN-DA is presented for $\Delta L = 0.5$.

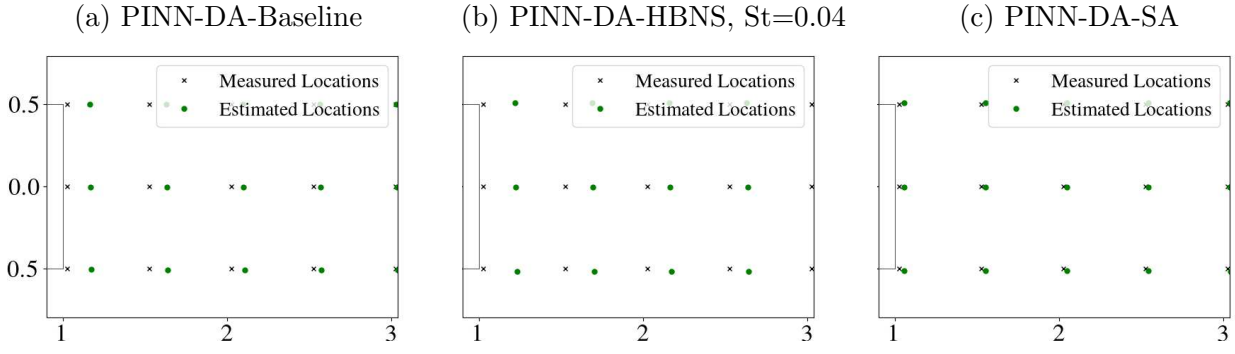


Figure D.3: Nominal and transformed measurement coordinates showing the effect of using embedded transformation (restricted to the xy plane) during training of the PINNs when data spacing is $\Delta L = 0.5$.

Figure D.3, showing the estimated measurement transformation at $\Delta L = 0.5$, highlights how the transformation embedded PINN-DAs are sensitive to data resolution. Although the same trends in measurement transformation are observed (shifting and translation) as for $\Delta L = 0.2$ (Figure 8.9), the magnitude of transformations are significantly reduced. The reduction in stretching of the measurement data, in particular, has large implications on reconstruction accuracy. The reconstruction accuracy of the shear layer is highly sensitive to misalignment of data, and thus the reduction in stretching at coarser data resolution reduces the improvement in reconstruction accuracy from use of transformation embedded PINN-DA. As expected, PINN-DA-SA (and also PINN-DA-HBNS-SA) behaves most consistently at the coarser resolution, as a result of additional physics constraints from the modelled forcing component. The estimated transformations at the coarser resolution demonstrates how, just as with the flow reconstruction accuracy itself, the transformation embedded PINN-DA approach is also sensitive to coarseness of data resolution.

The error profiles at $\Delta L = 0.5$, presented in Figure D.4, show similar trends to the profiles at the finer spacing (Figure 8.7). In the profile nearest to the wall, the maximum error reduces from approximately 60% to 40% of free-stream velocity. This then decays further downstream as the effect of different PINN-DA formulations become more apparent. These error profiles demonstrate that whilst the small measurement transformations learnt by the transformation embedded PINN-DA do reduce the peak error, their reduced magnitude results in still relatively high peak errors.

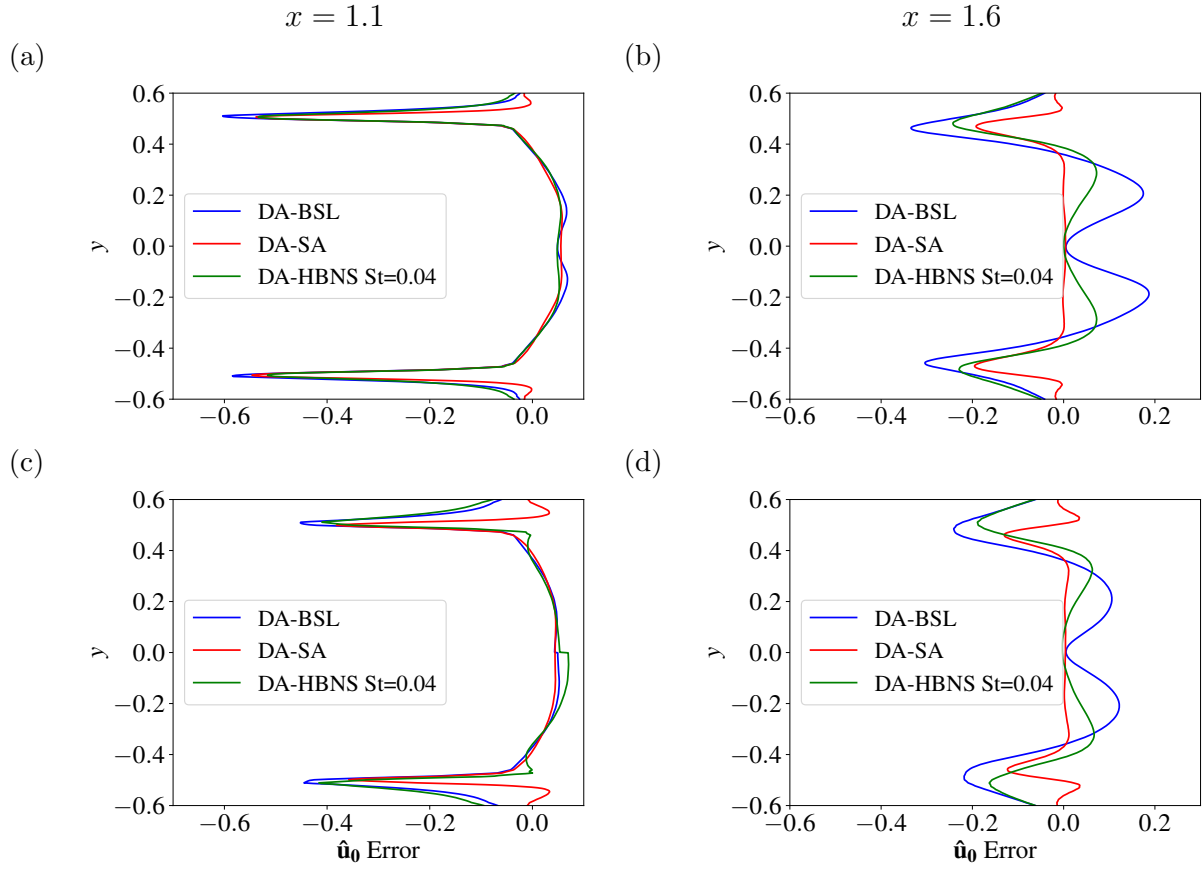


Figure D.4: Reconstructed streamwise mean velocity error profiles for different PINN-DA formulations, evaluated at two different streamwise locations (sampled at $\Delta L = 0.5$. (Top) without transformation embedded PINN-DA. (Bottom) with transformation embedded PINN-DA.

Figure D.5 shows the mean-flow reconstruction using three formulations of transformation embedded PINN-DAs (DA-Baseline, DA-HBNS resolved at $St = 0.04$ and DA-SA) from measurements with $\Delta L = 0.5$ spacing. One can observe similar trends to the reconstruction at finer resolution (Figures 8.11 and 8.14). PINN-DA-Baseline performs worst with high reconstruction error in the shear layer (both from the limitations of the PINN-DA-Baseline formulation in addition to the more limited measurement transformations) as well as in the recirculation. Use of PINN-DA-HBNS (resolved at $St = 0.04$), whilst still having similar reconstruction error in the shear layer, does improve prediction of the recirculation, with a reduction in error magnitude. Finally, PINN-DA-SA performs best, with better capture of the shear layer (as a result of the modelled forcing component, although the peak error magnitude shows the effect of the limited measurement transformation), in addition to better capture of the flow recirculation downstream. This result shows that the behaviour between PINN-DA formulations is consis-

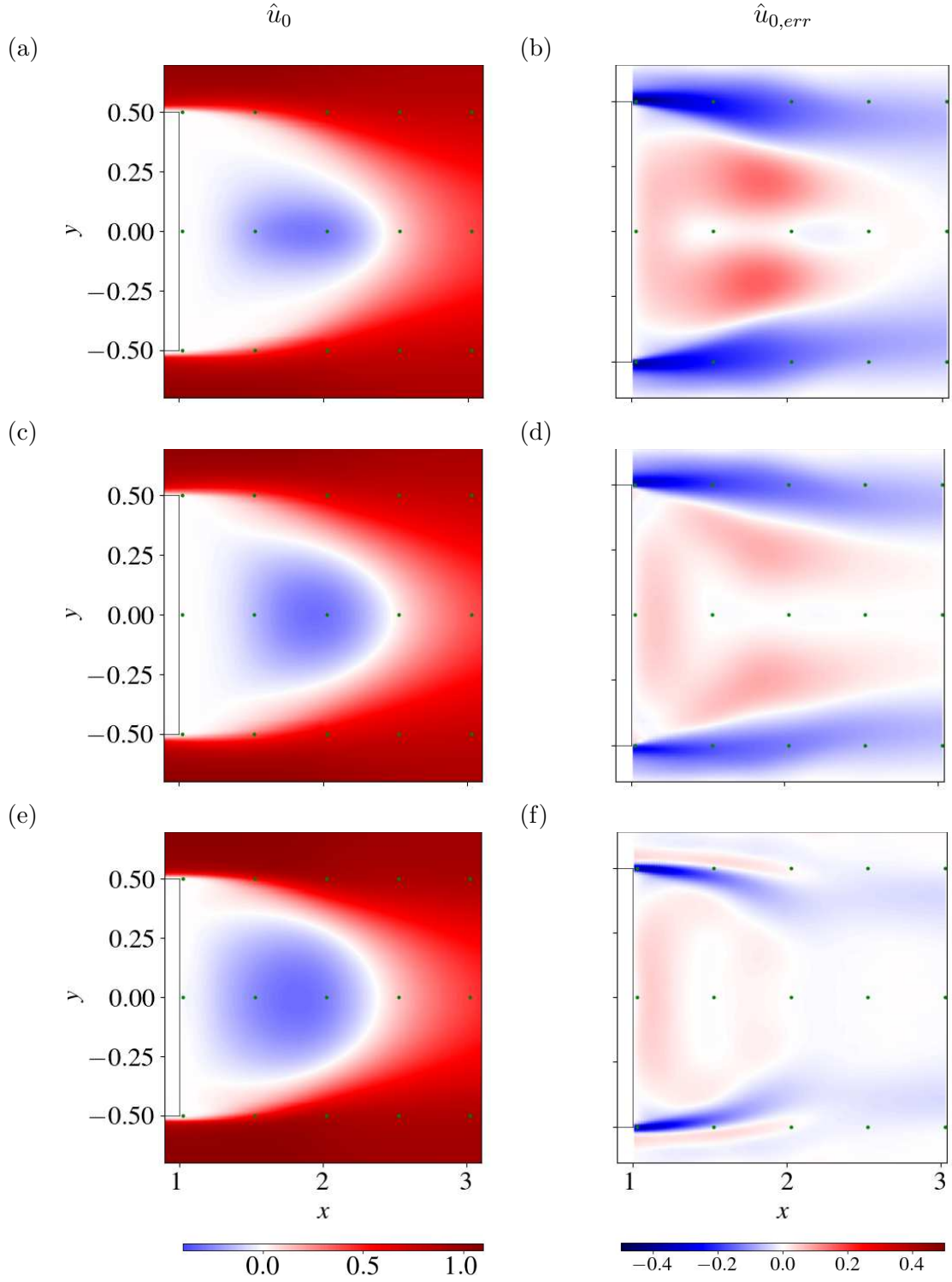


Figure D.5: PINN reconstruction using different methodologies for data spacing $\Delta L = 0.5$, when embedded transformation method is being used. (a,b) PINN-DA-Baseline, (c,d) PINN-DA-HBNS resolved at $St=0.04$, (e,f) PINN-DA-SA.

tent across data resolutions with the turbulence model augmented formulation (PINN-DA-SA) reconstructing the mean flow most accurately, followed by the resolution of some low-frequency scales (PINN-DA-HBNS) and finally the basic RANS formulation (PINN-DA-Baseline).

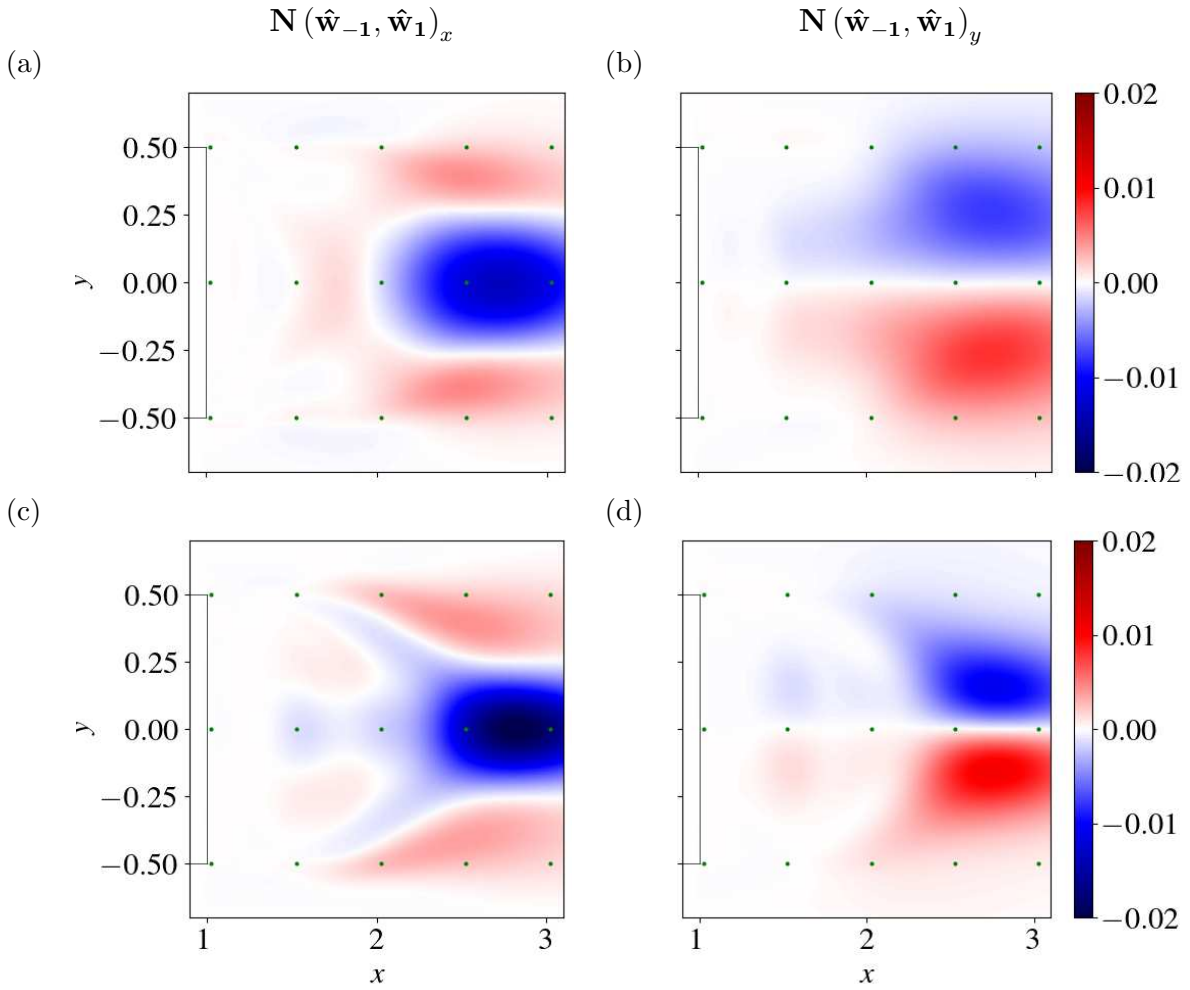


Figure D.6: PINN reconstruction of resolved forcing at $St = 0.22$ for data spacing $\Delta L = 0.5$. (a,b) HBNS-SA with mean eddy viscosity only as in Chapter 8, (c,d) HBNS-SA with mean and fluctuating eddy viscosity.

D.3 PINN-DA-HBNS-SA with fluctuating eddy viscosity field

In this section, the effect of considering an unsteady eddy-viscosity formulation is trialed (similar to the approach in von Saldern et al. (2023)). This uses the formulation presented in Eqs. (3.25) and (3.26).

Figure D.6 shows the reconstructed resolved forcing component for PINN-DA-HBNS-SA (resolved at $St = 0.22$) using fluctuating eddy viscosity from velocity measurements spaced at $\Delta L = 0.5$. One can observe a clear difference in the prediction of resolved forcing when assuming only a mean eddy-viscosity component as in Chapter 8 (Figure D.6(a,b)) and when the

fluctuating low-frequency harmonic eddy viscosity is also included (Figure D.6(c,d)). Comparing these coarse measurement reconstructions with the equivalent fine measurement resolved forcing components, presented in Figure 8.12(c,d), shows that use of a fluctuating eddy-viscosity field improves the prediction of the resolved forcing. This result shows that implementation of the fluctuating, harmonic eddy viscosity improves the reconstruction accuracy of the harmonic velocity components ($\hat{\mathbf{u}}_1$), especially at coarser data resolution, than when only a mean eddy viscosity is assumed.

Although the prediction of the harmonic velocity components, $\hat{\mathbf{u}}_1$, are improved, this does not result in a corresponding improvement in reconstruction accuracy. This is attributed to the behaviour seen in Figure 8.16. Despite the improvement in prediction of the low-frequency resolved components, the relatively small contribution of the resolved forcing component means its effect on flow reconstruction accuracy is limited. However, provided the effect of low-frequency resolved scales can be maximised in future work, this approach may be an effective method of improving reconstruction accuracy of PINN-DA-HBNS-SA, especially at coarse measurement resolutions.