



**MICHIGAN ENGINEERING
CENTER FOR NAVAL RESEARCH & EDUCATION**
UNIVERSITY OF MICHIGAN

2025 Annual Research Report



2025 ANNUAL RESEARCH REPORT

CENTER FOR NAVAL RESEARCH & EDUCATION

(cnre.engin.umich.edu)

UNIVERSITY OF MICHIGAN

Preface

The 2025 Annual Research Report describes the research conducted at the Center for Naval Research and Education (CNRE) at the University of Michigan in the past year. CNRE is supported by the Office of Naval Research under Grant N00014-24-1-2013 with the goal of conducting fundamental research into naval- and marine-relevant S&T problems, while training capable future generations. Our research program reflects the cutting-edge and multi-disciplinary nature of such problems and brings together researchers with a wide range of skills to work on them. This year's program includes challenging problems in the areas of complex hydrodynamics and flow control, electromagnetics, advanced materials, data-augmented modeling, biofouling, marine energy, shape optimization, machine learning and data assimilation, and physics-based modular construction for ship building. The year's cohort included 15 postdoctoral researchers and senior graduate students mentored by faculty from the Aerospace, Electrical Engineering and Computer Science, Materials Science, Mechanical, and Naval Architecture & Marine Engineering departments. *Engagement during education* is an important aspect of our research and I would like to thank the approximately 25 members of NSWC, Carderock, NRL and ARL for their collaboration and active interest in our activities. A brief summary of the assembled articles follows.

Lithium solid state batteries are central to electric shipping and sophisticated modeling methodologies for their physical behavior are explored. Surge wave energy converters are a means to harness wave energy, and a general framework for optimizing their performance that also considers subsystem interactions is studied. What will materials of the future be like? Advanced materials are explored by the use of kirigami to develop novel compliant mechanisms and surfaces with unique fluid-surface interaction properties. Can physical principles inform construction approaches? A novel Six-S-Principle Methodology is studied alongside Generative AI towards developing modular construction strategies for ship building.

Underwater vessels can puzzlingly, display highly asymmetric flow fields that influences their resulting loads. Global stability analysis is shown to reproduce this behavior and is used to explain the source of asymmetric behavior and its dependence on angle of attack. Efficient methods are developed to model the scattering of electromagnetic waves from the ocean surface. Biofouling is inevitable during marine operation and causes increased resistance, fuel consumption and maintenance costs. The mechanisms by which biofilms cause drag are experimentally explored using both live species and synthetic replicates, and a novel simulation approach is developed. An experimental study examines surface roughness as a means to control the flow around pitching hydrofoils. Cavitation can be very noisy and requires very large computational grids to directly capture the lowest pressures that drive inception. A novel subgrid model is developed that reliably predicts the inception on significantly smaller grids without directly capturing the lowest pressures thus making practical predictions feasible.

Can machine-learning assist in developing multidisciplinary optimization methodologies of the future? Deep reinforcement learning approaches are explored alongside gradient-based approaches to develop optimal hydrofoil shapes in the multi-phase regime. The simultaneous assimilation of experimental and simulation data opens up exciting possibilities in how future simulations are conducted and experimental measurement techniques are developed. A novel adjoint-based approach is developed for complex geometries and large-scale data. Also, a machine-learning based ap-

proach is used to predict the pressure field from coarse Lagrangian velocity data. Can high-fidelity experimental and simulation data enable the development of improved lower-fidelity models? A systematic methodology for developing data-driven turbulence models is studied.

The 2025 Annual Research Report also describes proceedings of the 2025 Undergraduate Summer Research Program that extended from May 27 to August 7, 2025. A total of 13 undergraduate students were mentored during summer by postdoctoral researchers, senior graduate students and faculty alongside 15 collaborators from NSWC, Carderock on problems related to the core CNRE research described above. The entire team was composed of members from Mechanical, Electrical, Aerospace, NA&ME, Computer Science and Math departments. In addition to biweekly meetings, Carderock engineers made presentations to the students and postdoctoral researchers as part of a *My Path To The Navy* series. The energetic undergraduate cohort was educated on cutting-edge naval and marine problems and conducted research using state of the art approaches. An exciting ten weeks culminated in a well-received final presentation workshop. Seven conference abstracts and three paper submissions have already resulted from their summer research.

This research report is CNRE's first such annual report. I would like to acknowledge Katie Klink for her help in designing the cover page of the report. A big thank you also goes to Mehedi H. Bappy, Divyanshu Gola, András Szabó, Gao Jun Wu and Xianzhang Xu for their considerable and cheerful help in proofing and compiling the report. Finally, I am grateful to all members of the CNRE team for their efforts that made the reported work possible!

KRISHNAN MAHESH
NOVEMBER 15, 2025

Contents

| | | |
|-----------|--|------------|
| 1 | Phasefield and Mesoscale Modeling of Anode Voiding in Lithium Solid State Batteries | 1 |
| | <i>Andrew Li & Venkat Viswanathan</i> | |
| 2 | Improvement of Hydrodynamic Properties using Kirigami Patterns & Organic Vapor Jet Deposited Coatings | 15 |
| | <i>Daniel Esterkin, Morgan Pen¹, Joshua Behringer², Eva Pontrelli, Lokesh Silwal, Ryan Harber¹ & Max Shtein</i> | |
| 3 | On the Optimal Performance of Oscillating Surge Wave Energy Converters | 29 |
| | <i>Binh D. Truong & Lei Zuo</i> | |
| 4 | 6S-Principle Methodology for Constructing a Grand Panel and Comparison with Generative AI Tools | 53 |
| | <i>Zhigang Wei & Pingsha Dong</i> | |
| 5 | Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis | 71 |
| | <i>András Szabó & Krishnan Mahesh</i> | |
| 6 | Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow | 93 |
| | <i>Lokesh Silwal, Rodrigo Vilumbrales-Garcia, & Anchal Sareen</i> | |
| 7 | Dynamic Modeling of Cavitation Inception in the Subgrid Scale of Large Eddy Simulation | 113 |
| | <i>Mehedi H. Bappy & Krishnan Mahesh</i> | |
| 8 | Understanding the Drag Production Mechanisms of Soft Biofilms using Live Species and Synthetic Replicates | 135 |
| | <i>Rodrigo Vilumbrales-Garcia & Harish Ganesh</i> | |
| 9 | Understanding the Drag of Soft Biofilms: A Multidisciplinary Experimental Approach | 145 |
| | <i>Mohammad A. Elsouht, Rodrigo Vilumbrales-Garcia, Vanessa Molina, Matthew First, Jason Lee & Harish Ganesh</i> | |
| 10 | Two-Way Coupled Simulation of Biofouling-Inspired Dynamic Roughness in Turbulent Channel Flow | 149 |
| | <i>Divyanshu Gola, Soham Prajapati & Krishnan Mahesh</i> | |

Contents

| | |
|---|------------|
| 11 Statistical Modeling of Electromagnetic Waves Scattering from Random Ocean Surfaces | 161 |
| <i>Behzad Yektakhah, Ryan Strohman, & Kamal Sarabandi</i> | |
| 12 Machine Learning-Aided Flow Field Reconstruction from Sparse and Noisy Measurements | 177 |
| <i>Xianzhang Xu, Daria Skalitzky & Krishnan Mahesh</i> | |
| 13 Optimizing Hydrofoil using Gradient-Based and Deep Reinforcement Learning Approach | 199 |
| <i>Mohamed A. S. Abdul-Kaiyoom & Joaquim R. R. A. Martins</i> | |
| 14 Adjoint-Based Assimilation of External Data with High-Fidelity Simulations of Complex Flows | 213 |
| <i>Gao Jun Wu, Sreevatsa Anantharamu & Krishnan Mahesh</i> | |
| 15 Development, Application, and Refinement of a New Data-Driven Turbulence Modeling Methodology | 229 |
| <i>Niloy Gupta & Karthik Duraisamy</i> | |
| 16 CNRE Undergraduate Summer Research Experience 2025 | 251 |
| 17 Publications and Presentations | 267 |
| 18 2024-2025 Roster | 271 |

Phasefield and Mesoscale Modeling of Anode Voiding in Lithium Solid State Batteries

ANDREW LI & VENKAT VISWANATHAN

Solid-state batteries have shown great potential as a next-generation technology for delivering higher energy density with greater safety than existing Li-ion battery technologies. However, their applicability is still severely hindered by interfacial instabilities such as Lithium dendrite formation and anode-electrolyte delamination through interfacial void formation. While there is a growing body of experimental and modeling work on this topic, the exact underlying mechanisms are still highly contested. Modeling works ranging from atomistic to mesoscale models disagree on the roles of flux focusing, point defect diffusion, and mechanical deformation, as well as the role of impurities and surface geometry. While much work has been done to characterize the mechanical properties of Lithium polycrystals and bulk diffusion behavior in Lithium single crystals, the contributions of Lithium self-diffusivity on void coalescence is still hotly contested. We believe that to resolve these fundamental issues, we must develop a more rigorous and flexible formulation of the governing equations at the mesoscale to capture the complex electro-chemo-mechanical behavior. In this report, we will provide an update on our development of a baseline phase field model which, investigates the electrochemical contributions more carefully through free energy considerations, reproduces key features of known voiding processes at the electrode-electrolyte interface, and allows for the incorporation of more complex geometries and contact angle effects down the line.

Keywords: phase field, solid state batteries, Lithium, voiding

1.1 Introduction

One of the reasons utilization of Lithium (Li) anodes in solid state batteries remain highly challenging is its poor cycling behavior due to Li-solid electrolyte interface degradation (Barai et al., 2024). In particular, there has been growing recognition of the role of void formation at the Li-solid electrolyte interface, which can result either in a highly resistive interface due to excessive contact loss/delamination, or create inhomogeneities/current hot spots which promote dendrite growth during cell cycling, as depicted in Figure 1.1.

To capture the effects of anode voiding, there is a growing number of multi- and meso-scale models which aim to capture the void initiation and growth behavior in Li anodes at the Li-solid electrolyte interface. At the continuum level, it is hypothesized that void formation could be dominated by various mechanisms, including vacancy coalescence (Barai et al., 2024), flux focusing (Yoon et al., 2025), creep deformation (Agier et al., 2022), and impurity buildup (Shishvan et al., 2024). Finally, there is active development of multi-scale models, where density functional theory (DFT), molecular dynamics (MD), and continuum models are linked together to provide better quantitative estimates (Feng et al., 2024).

Despite all these efforts, there is still a considerable gap between experimental and modeling efforts, with a wide range of predictions and disagreements on proposed mechanisms (Shishvan et al.,

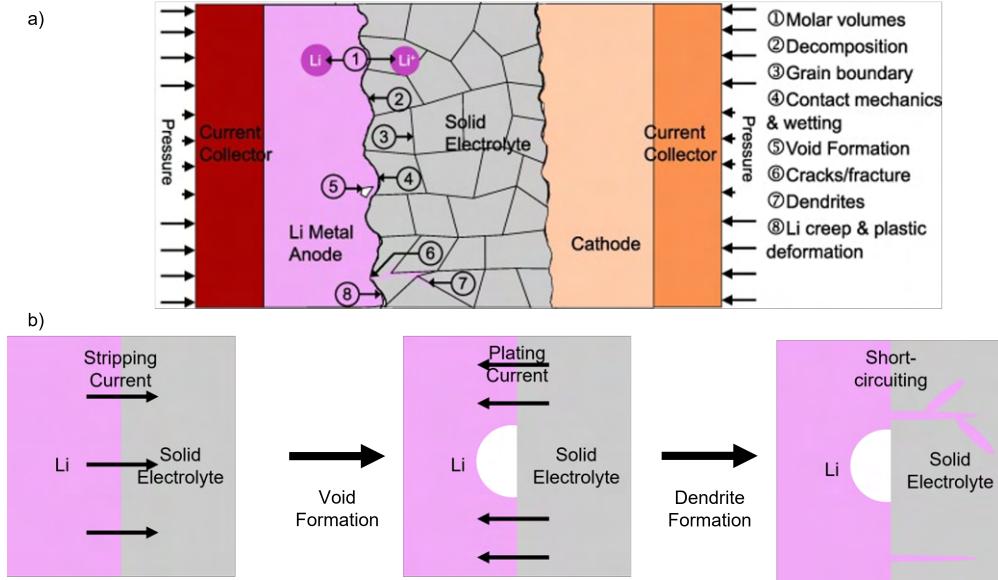


Figure 1.1: Schematic depiction of a Li solid state battery and voiding induced failure. a) Schematic of a full solid state battery with Li anode, solid electrolyte, cathode, and current collectors. Reproduced from Ahmad et al. (2022). b) Depiction of how voiding at the Li-solid electrolyte interface can promote dendrite growth. As the cell is stripped (Li removed from the anode region), an initially pristine Li-solid electrolyte interface begins to develop voids. The formation of voids result in an uneven current distribution during the plating process (Li being deposited onto the anode), which can result in dendrite formation at current hot spots.

2024). This likely arises from the complex Li anode microstructure evolution during cell production and operation (Shishvan et al., 2024; Milan and Pasta, 2022), making a direct translation of quantities from the nano- to meso-scale highly challenging. Authors have argued that flux focusing based on standard Butler-Volmer dynamics (Krauskopf et al., 2019) is insufficient to account for void formation (Agier et al., 2022), creep deformation around impurity particles (Shishvan et al., 2024; Shojaei et al., 2025) is unable to sustain stable voids and is inconsistent with pristine interfaces observed after cycling (Barai et al., 2024), and vacancy coalescence lacks a thermodynamic driving force after accounting for electrode thinning (Shishvan et al., 2023).

In contrast to Lithium deformation mechanics which, has been extensively characterized (LePage et al., 2019), models for void formation based on vacancy diffusion have been treated largely on a phenomenological basis (Barai et al., 2024), where the choice of the free energy is assumed to be sufficiently captured by a quartic form, i.e.,

$$F = \int A\phi^2(1-\phi)^2 + \frac{1}{2}\kappa(\nabla\phi)^2 dV, \quad (1.1)$$

where the first term, which is the bulk free energy contribution, contains a fitting parameter A that is not physically motivated, and the spinodal and binodal points are fixed by the polynomial expression, and the second term is the gradient term related to surface energy contributions. Combined with Butler-Volmer dynamics and creep deformation, these models have captured more intricate interactions of electro-chemo-mechanical models.

We instead argue and present some early results on how a more rigorous formulation of diffusion based contributions to Li anode voiding can be achieved by developing a phenomenological Cahn-Hilliard Stokes type model, similar to that of Barai et al. (2024), but with a more careful analysis of the free energy, allows us to estimate temperature dependent behaviors coming from the chemical contributions, and derive a reduced order model for contact loss time evolution un-

der clearly specified simplifying conditions that can be systematically generalized. In addition, we briefly discuss how void stabilization via mechanical feedback (Mosleh et al., 2025; Barai et al., 2024) can arise from a replenishment current contribution due to creep (Feng et al., 2024).

1.2 Free energy and spinodal points

Various authors (Barai et al., 2024) have argued that diffusion behavior dominates the void formation process, although arguments based on coalescence energy barrier have also been made against this claim (Agier et al., 2022; Shishvan et al., 2023, 2024). While the microscopic picture is likely vastly more complicated, such as in the case of voids in semiconductor connects (Yu et al., 2009), the objection (Shishvan et al., 2023) against a vacancy coalescence model ultimately relies on an ideal solution model that neglects vacancy-Li interactions, which cannot be sustained when stripping can result in highly non-equilibrium vacancy concentrations near the Li-electrolyte interface. Instead, equilibrium vacancy concentration typical of metals, which are on the order of parts per hundred thousand/million, suggests that the free energy of a Li-vacancy system is instead highly non-convex, and the high dimer and trimer formation energies noted in literature (Shishvan et al., 2023) reflects a free energy with strong vacancy concentration dependency, or equivalently, highly separated binodal points for a standard phase separating binary system (in terms of normalized concentration).

Therefore, we take a more coarse-grained approach by considering how to choose a free energy model that is self-consistent with the above-mentioned issues. In general, the total free energy of the system can be obtained as the sum of the bulk Li free energy and a gradient energy term estimated from the Li-void interfacial energy based on Molecular Dynamics (MD). Current phase field models (both conserved and nonconserved dynamics driven) utilize the quartic free energy formulation as Eq. (1.1), which benefits from improved numerical stability (Shojaei et al., 2025) and has an analytically well-known interface behavior. However, we argue that the voiding onset is better captured by a regular solution model (and generalizations via RK polynomials (Yao and Viswanathan, 2024)) where one minimum corresponds to the equilibrium vacancy concentration in Li metal, i.e.,

$$F_{reg}[\phi] = \int \left[RT(\phi \ln \phi + (1-\phi) \ln (1-\phi)) + \alpha \phi (1-\phi) + \frac{1}{2} \kappa (\nabla \phi)^2 \right] dV, \quad (1.2)$$

where α is the interaction strength and κ is the interfacial energy. As shown in Figure 1.2, this is usually not considered as the free energies appear to closely resemble each other with the same extrema. However, a key distinction between the regular solution formulation in Eq. (1.2) and quartic free energy in Eq. (1.1) is the position of the spinodal points, which is considerably closer to the binodal points for the case of a regular free energy model.

Recall the spinodals are given by

$$\{\phi_s \mid \frac{\delta^2 F}{\delta \phi^2}(\phi_s, \nabla \phi = 0) = 0\}. \quad (1.3)$$

Therefore, the quartic free energy function $f(\phi) = A(c - \phi_a)^2(\phi - \phi_b)^2$, where ϕ_a and ϕ_b are the binodal points, have

$$\phi_{qs} = \frac{1}{2} \left(\phi_a + \phi_b \pm \frac{1}{\sqrt{3}} |\phi_a - \phi_b| \right), \quad (1.4)$$

where the subscript qs denotes the spinodal of the quartic free energy, which does not account for any shift in the spinodal position due to the changes in temperature. The onset of voiding,

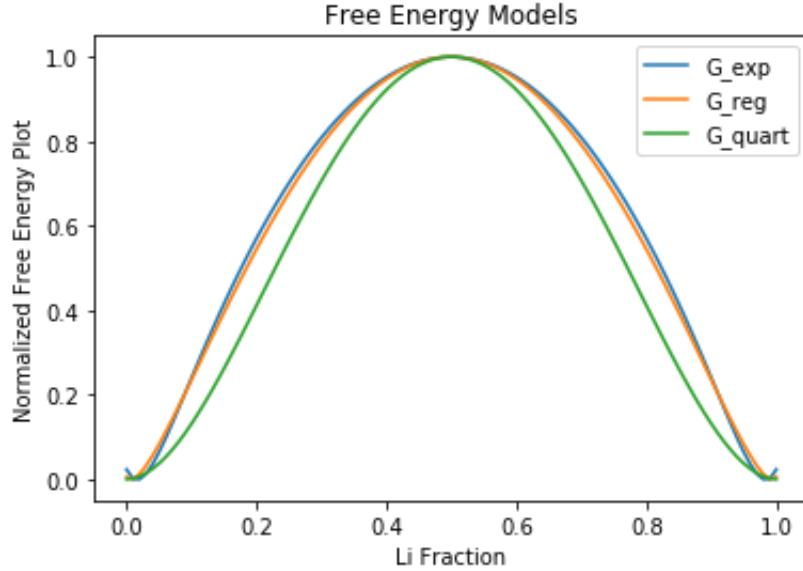


Figure 1.2: Comparison of three free energy models which have the same free energy barrier and binodal points near Li fractions at 0 and 1. G_{reg} is a regular solution model, G_{exp} uses an exponential function as a basis (Shojaei et al., 2025), and G_{quart} is the quartic free energy used in standard phase field models. Despite having (nearly) the same extrema, the spinodal points differ considerably among the three, which shows that a free energy model must be selected carefully for predicting the onset of interfacial voiding, even though simulations may be run with reasonable accuracy after the void onset/nucleation event using more stable free energy choices.

or equivalently the spinodal point is therefore predicted to be the same across all temperatures at $\phi_{qs} \approx 0.211$ and 0.788 . On the other hand, for the regular solution free energy model, where $f(\phi) = RT(\phi \ln \phi + (1-\phi) \ln (1-\phi)) + \alpha \phi (1-\phi)$, the spinodals are naturally given by the interaction strength α and temperature T ,

$$\phi_{rs} = \frac{1}{2} \left(1 \pm \sqrt{1 + \frac{2RT}{\alpha}} \right), \quad (1.5)$$

where the subscript rs denotes that these are the spinodal points in the case of the regular solution model. This shows that, in the limit of diffusion dominated voiding, assuming that reaching the spinodal point is what induces void formation at the interface, the critical current grows as a function of temperature, which can be used to verify if the measured open circuit voltage/critical capacity is consistent with a diffusion based model of this form in low pressure conditions where diffusion is expected to be dominant instead of mechanical creep.

Next, consider spinodal points for binodal points close to 0 and 1, which is the case for Li-vacancy phase diagram, where the binodal point close to 1 represents Li concentration with a very small vacancy equilibrium fraction (10^{-7}). While the binodal points do not change noticeably as Li and vacancies can be treated as nearly immiscible species, based on Eq. (1.5), the spinodal points strongly depend on temperature as

$$\frac{\partial \phi_{rs}}{\partial T} = \frac{R}{\alpha \sqrt{1 + \frac{2RT}{\alpha}}} = \pm \frac{R}{\alpha} \left(1 - \frac{2R}{\alpha} T + O(T^2) \right). \quad (1.6)$$

In other words, the spinodal points are predicted to drift apart nearly linearly in the temperature region of interest (near room temperature). Assuming voiding formation at the Li-electrolyte interface occurs when the spinodal point is crossed, an increase in temperature improves the critical

stripping current density as the spinodal point shifts towards lower Li concentration/higher vacancy concentration.

1.3 Basic formulation

We begin with the basic phase field model component in the Li anode region/bulk, where the Li phase is given by $\phi \in [0, 1]$, the normalized concentration. The void formation process, in the absence of mechanical deformation, is assumed to follow Cahn-Hilliard (CH) dynamics,

$$\frac{\partial \phi}{\partial t} = \nabla \cdot D(\phi) \nabla (\mu(\phi) - \kappa \nabla^2 \phi), \quad (1.7)$$

where $D(\phi)$ is the diffusivity, $\mu(\phi)$ is the chemical potential, and κ is the gradient energy coefficient. This formulation is chosen as the CH equation has been successfully utilized in modeling void formation due to electromigration (Mahadevan and Bradley, 1999; Bhat et al., 2002) and Kirkendall effects (Werner and Weinberg, 2017). Mechanical contributions can be included by adding an advection term such that the time evolution is given by the Cahn-Hilliard-Stokes (CH-Stokes) dynamics,

$$\frac{\partial \phi}{\partial t} = \nabla \cdot D(\phi) \nabla (\mu(\phi) - \kappa \nabla^2 \phi - \phi \mathbf{v}), \quad (1.8)$$

where \mathbf{v} is the local lattice velocity due to mechanical deformation (e.g. creep (Barai et al., 2024)), which is introduced to capture the creep-induced void healing/annihilation process, and a more quantitative model can be found in existing literature (Barai et al., 2024). In the initial simulations, we are treating it as a constant field for simplicity, as the void stabilization effects can already be captured in this manner, and benchmarks with more detailed mechanical coupling following Barai et al. (2024) is currently being developed.

This basic CH-Stokes equation is then coupled to standard Butler-Volmer (BV) dynamics at the electrode-electrolyte interface and used to understand basic properties of electrochemically induced voiding.

1.4 Bulk versus surface diffusivity

An important aspect we have not yet discussed in the matter of Li (self-)diffusivity. While we have assumed the diffusivity to be constant in the previous cases, the diffusivity term $D(c)$ can play a crucial role in the determination of the void shape (Barai et al., 2024), as a larger bulk diffusivity contribution allows for a redistribution of vacancies in the anode bulk to reduce the bulk free energy, whereas a larger surface diffusivity allows for rapid vacancy redistribution at the anode interface to minimize the surface energy, and together determine the shape of the void. Moreover, sufficiently large surface diffusivity and low bulk diffusivity, for example, frequently seen during CVD of III-V semiconductors, can essentially be used to model surface rearrangements dominated by surface diffusion. In this case, we will assume that the surface diffusivity is much greater than the bulk diffusivity ($D_s \gg D_b$), and express the effective Li diffusivity in the anode as an interpolation of the bulk and surface diffusivity, i.e.,

$$D(c) = h(c)D_b + (1 - h(c))D_s, \quad (1.9)$$

where D_b and D_s are the bulk and surface diffusivities respectively, and $h(c)$ is an interpolation function, which, following Zhao et al. (2022), we model as

$$h(c) = c^3(6c^2 - 15c + 10), \quad (1.10)$$

although the choice is by no means not unique.

Finally, to incorporate adhesion energy between Li and solid electrolyte, we introduce, following Yu et al. (2012) and Warren et al. (2009), the contact angle effect, which is given at the boundary (with orientation \hat{n}) by the condition

$$\frac{\nabla c}{|\nabla c|} \cdot \hat{n} = -\cos\theta \quad (1.11)$$

at the triple junction between void, Li metal phase, and the electrolyte surface.

The slowdown in void formation due to creep-induced flux has already been systematically investigated in the previously mentioned works. Before we can rigorously prove that the creep flux, which requires a well-defined incompressible model, and the Cahn-Hilliard equation, where concentration variations exist, following (Feng et al., 2024), we will simplify the creep flux as a power law creep with a constant lattice drift velocity in the form of

$$\mathbf{v} \cdot \hat{e}_y = \dot{\varepsilon} = \sigma^m A_c \exp\left(-\frac{Q_c}{RT}\right), \quad (1.12)$$

where the applied strain rate is given by a power creep law. The reader interested in a more accurate mechanics modeling is referred to LePage et al. (2019), but should note that models coupling information from the deformation map and phase fields may exhibit various unphysical phenomena that must be carefully checked against.

Finally, note that the current is modeled with Butler-Volmer kinetics at the Li-electrolyte interface, where the Li-ion concentration in the electrolyte is assumed to be at steady state since most solid electrolytes of interest are chosen for high Li-ion conductivity and selectivity, and thus not modeled explicitly. This is a reasonable simplification for the case of voiding, but it may not be suitable for the case of dendrite growth, which we are not investigating yet.

1.5 Introduction to smoothed boundary method

As the smoothed boundary method (SBM) is not widely used outside of the phasefield community, we provide a brief description of its relevant formulation such that the next section is more tractable.

Consider a domain Ω . We can associate an indicator function I such that

$$I(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \in \Omega, \\ 0 & \mathbf{r} \notin \Omega. \end{cases} \quad (1.13)$$

The smoothed boundary method, as the name implies, smooths the boundary by replacing the indicator function with a smooth function ψ such that

$$\begin{aligned} (\psi(\mathbf{r}), \nabla \psi) &\approx (1, 0) & \mathbf{r} \in \Omega, \\ (\psi(\mathbf{r}), \nabla \psi) &\approx (0, 0) & \mathbf{r} \notin \Omega. \end{aligned} \quad (1.14)$$

In the 1D case, we can replace the indicator on the interval $[0, 1]$ with the smoothed domain order parameter ψ

$$\psi(x, x_0, L, \epsilon) = \frac{1}{2}(1 + \tanh(\frac{L - (x - x_0)}{\epsilon}))(1 + \tanh(\frac{x - x_0}{\epsilon})), \quad (1.15)$$

where x_0 and L determine the domain size, and ϵ determines the boundary width that can be set to a small value depending on the resolution of interest (e.g. $\epsilon \approx 0.001$ for a normalized domain size of 1).

Associated with this domain order parameter is the normal vector, which is given by

$$\hat{\mathbf{n}} = \lim_{\epsilon \rightarrow 0} \frac{\nabla \psi}{|\nabla \psi| + \epsilon}, \quad (1.16)$$

where an infinitesimal value $\epsilon \rightarrow 0$ term is introduced. The small value, which is strictly 0 for mathematical purposes, is set to be some floating number ($\epsilon_B \approx 10^{-16}$), and acts as a numerical stabilization term.

It is evident that many complex domains can be accurately modeled with this domain order parameter up to an interfacial skin region/width of ϵ_B , and several cases have been tested rigorously by Yu et al. (2012). This is the main limitation of this approach, as ϵ_B should be set as the smallest lengthscale in the system to avoid non-physical effects that come from the artificial interfacial layer.

Once the domain order parameter is defined, we can associate with it various types of boundary conditions (BC). This is achieved by leveraging two standard calculus identities, which we demonstrate for the Neumann and contact angle BC. The Dirichlet BC is similarly derived.

Finally, we note that in order to maintain numerical efficiency while balancing the need for finer mesh near boundaries, we utilize adaptive mesh refinement (ADR) as shown in Fig. 1.3.

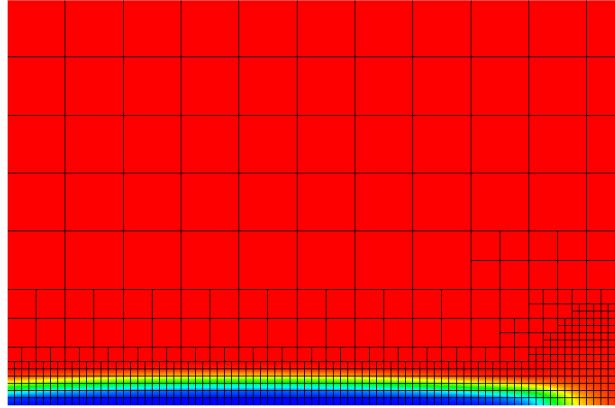


Figure 1.3: A zoomed in view of (part of) a simulation domain. Adaptive mesh refinement is applied automatically near regions with higher spatial variations of the simulation field, and the remaining domain can be efficiently simulated with just a coarse mesh.

1.6 Smoothed boundary based implementation

In this section, we provide a more detailed derivation that has been developed for the CH-Stokes equation with the inclusion of the contact angle effect to represent the interfacial adhesion energy between Li metal and different choices of electrolyte, as well as electrostatics for the reaction kinetics governed by Butler-Volmer kinetics. This derivation is based on the smoothed boundary method developed by Yu et al. (2012), which shows great promise for the simulation of more complex interactions down the line.

The first identity we make use of is that for a scalar field H ,

$$\psi \nabla^2 H = \nabla \cdot (\psi \nabla H) - \nabla \psi \cdot \nabla H. \quad (1.17)$$

The Neumann BC at a boundary with normal vector $\hat{\mathbf{n}}$ is given by

$$\mathbf{j} \cdot \hat{\mathbf{n}} = J_n, \quad (1.18)$$

where J_n is a scalar. Substituting Eq. (1.16) into Eq. (1.18) gives the smoothed boundary version of the Neumann BC as

$$J_n = \frac{\nabla \psi}{|\nabla \psi|} \cdot \mathbf{j}. \quad (1.19)$$

For the case of a general conserved dynamics for scalar field ϕ ,

$$\frac{\partial \phi}{\partial t} = -\nabla \cdot \mathbf{j}. \quad (1.20)$$

Multiplying the equation by the domain order parameter and rearranging the terms gives

$$\frac{\partial \phi}{\partial t} = -\frac{1}{\psi} \nabla \cdot \psi \mathbf{j} + \frac{|\nabla \psi|}{\psi} J_n. \quad (1.21)$$

Note that the derivation for the Neumann condition is true for both the case of non-phase-separating diffusion and phase-separating diffusion in the Cahn-Hilliard (CH) equation. In addition, adding a velocity term, as in the case of the Cahn-Hilliard-Stokes equations, does not affect this derivation either. However, for the CH equation, which is a 4th order PDE, an additional contact angle constraint can be specified. This is most evident when we express it as a coupled PDE through the mixed formulation.

$$\begin{aligned} \frac{\partial c}{\partial t} &= \nabla \cdot D \nabla \mu, \\ \mu &= f(c) - \kappa \nabla^2 c. \end{aligned} \quad (1.22)$$

A classic set of boundary conditions is the double Neumann BC, which is given by

$$\begin{aligned} \nabla c \cdot \hat{\mathbf{n}} &= 0, \\ D \nabla \mu \cdot \hat{\mathbf{n}} &= 0. \end{aligned} \quad (1.23)$$

The second equation in Eq. (1.23) is easy to understand, as it is the same no flux condition we see in the diffusion equation. The first equation corresponds to a contact angle condition, which is easy to implement as a Neumann boundary condition for the $\theta = 90^\circ$ case, but is in general more difficult to express, as exemplified by Warren et al. (2009), as a standard implementation requires the introduction of a delta function in the free energy functional. The SBM approach recovers the same expression with what effectively is a smoothed approximation to the delta function, which we detail as follows.

Consider an anode/matrix phase, with a single void/precipitate phase located along one of the boundaries. The triple junction (where the matrix, precipitate phase, and the boundary meet) is characterized by the wetting/contact angle θ , which is experimentally measurable and reflects the relative attractive/adhesive force of the two phases with respect to the boundary. Ignoring the boundary for now, the free energy of the system is given by

$$F = \int \left[f(\phi) + \frac{1}{2} \kappa |\nabla \phi|^2 \right] dV, \quad (1.24)$$

where the equilibrium condition requires that

$$\frac{\delta F}{\delta \phi} = 0. \quad (1.25)$$

In terms of the chemical potential, we therefore have

$$\frac{\partial f}{\partial \phi} - \kappa \nabla^2 \phi = 0. \quad (1.26)$$

Multiplying both sides by $\nabla \phi$, we have

$$\frac{\partial f}{\partial \phi} \nabla \phi = \kappa \nabla^2 \phi \nabla \phi, \quad (1.27)$$

or, equivalently,

$$\nabla \left(f - \frac{1}{2} \kappa (\nabla \phi)^2 \right) = 0. \quad (1.28)$$

Therefore, without loss of generality, we may express

$$|\nabla \phi| = \sqrt{2f/\kappa}, \quad (1.29)$$

which is the boundary on the magnitude of the gradient of ϕ . Recall next the contact angle condition is

$$\hat{\mathbf{n}} \cdot \frac{\nabla \phi}{|\nabla \phi|} = -\cos \theta, \quad (1.30)$$

where θ is the contact angle of the Li metal. Substituting Eqs. (1.16) and (1.29) into Eq. (1.30), we have

$$\nabla \psi \cdot \nabla \phi = -|\nabla \psi| \cos(\theta) \sqrt{\frac{2f}{\kappa}}. \quad (1.31)$$

Analogous to the case of the Neumann BC, multiplying the chemical potential by the domain order parameter ψ and rearranging terms gives

$$\mu = \frac{\partial f}{\partial \phi} - \frac{\kappa}{\psi} \nabla \cdot (\psi \nabla \phi) + \frac{\kappa}{\psi} \nabla \psi \cdot \nabla \phi, \quad (1.32)$$

which, substituting into Eq. (1.31), gives

$$\mu = \frac{\partial f}{\partial c} - \frac{\kappa}{\psi} \nabla \cdot (\psi \nabla c) - \frac{|\nabla \psi|}{\psi} \sqrt{2f\kappa} \cos \theta. \quad (1.33)$$

Finally, the electrical potential in the anode and electrolyte bulk are given by

$$\nabla \cdot \sigma \nabla \phi = 0, \quad (1.34)$$

with the reaction current at the anode-electrolyte interface as a boundary term. This can be expressed using SBM as a single equation

$$\nabla \cdot (\psi \kappa(c, \psi) \nabla \Phi) = -|\nabla \psi| J_n. \quad (1.35)$$

Similar derivations can also be found in Goel et al. (2022) and Hong and Viswanathan (2018). On the other hand, the reaction current J_n follows a simple Butler-Volmer expression in Eq. (1.36), where $\eta = \phi_s - \phi_e - U_{Li}$ is the overpotential, although more quantitative models including stress-induced electrochemical potential terms can be included as needed.

$$J_n = J_0 \left(\exp \left(-\alpha \frac{\eta}{RT} \right) - \exp \left((1-\alpha) \frac{\eta}{RT} \right) \right). \quad (1.36)$$

1.7 Numerical results

In this section, we provide some preliminary examples of simulation results for modeling void growth at the Li-solid electrolyte interface. This first set of simulations are run with the constant velocity term, and the Li anode domain shown in Fig. 1.4 and its boundary conditions are enforced using the smoothed boundary method as detailed in the earlier sections. A random point was chosen for a nonuniform current stripping to occur at the Li-electrolyte interface, which is the bottom edge of each of the simulated domains. As the stripping continues over time (from left to right), it can be observed that a red region, corresponding to the voiding region, appears at the interface, resulting in contact loss/delamination in both cases a) and b), corresponding to without and with a replenishment current from creep and a lattice shift of the Li anode.

Despite the simplicity of the simulation, we see that 2 key qualities are captured. 1. There is no significant flux from the bottom to the top of each domain, indicating in the case of vacancy coalescence induced voiding, after the critical current density is reached, the replenishment current from Li diffusion originating from the interior of the anode barely contributes to the dynamics. Instead, diffusion appears to play the role of surface rearrangement of atoms. 2. The replenishment current from either creep/lattice shift can significantly reduce the rate of void formation, even completely stopping them for sufficiently strong mechanical feedback, which can give rise to a critical/maximum void size in steady state.

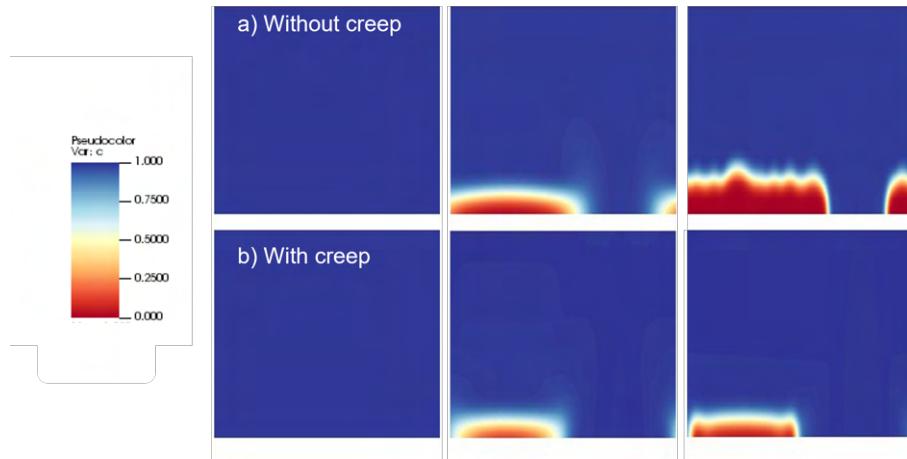


Figure 1.4: Comparison of phase field simulation results. Bottom boundary indicates the Li-electrolyte interface. In a) a constant stripping current density is applied, resulting in the continuous growth of a void. In b) A constant stripping current density, coupled with replenishment current from lattice velocity induced by the creep flow, results in a void that is stable in size and does not lead to full delamination.

In another set of simulations we investigated the effects of surface energy contributions of the Li-void region. For each of these simulations, the void surface area, which can be defined from the average concentration of Li at the anode-electrolyte interface, are plotted as a time series with some typical shapes shown in Fig. 1.5. Under constant stripping (galvanostatic) conditions, our simulations show that key characteristic results such as an initial period with relatively stable contact until the onset of voiding, a growth period with a sudden increase in void growth rate, and a gradual slowdown in void growth rate (which can eventually result in stable voids that stop growing), are all successfully captured.

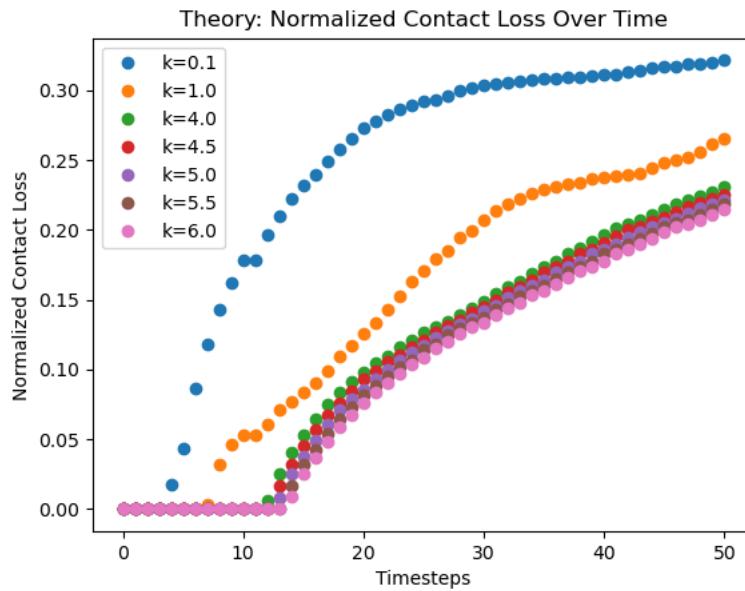


Figure 1.5: Normalized contact loss evolution of several simulations under various gradient energy coefficients k , which represent the Li-void surface energy. Sufficiently small gradient energy result in earlier onset of the voiding process, but increasing the gradient energy only slows down the voiding process marginally as the effect saturates. In all simulations the void growth process shows an initial stage where contact is well maintained with little to no void growth, an initial rapid increase, and a gradual slowdown which can taper off and eventually stop.

1.8 Conclusion and next steps

Having captured the basic properties of the voiding process at the mesoscale as discussed in the previous section, we are now working towards a multiscale model of the voiding process. On one hand, we are incorporating information from DFT and MD simulations to provide better estimates for diffusivity and free energy contributions to the Li-vacancy diffusion process, which play critical roles in the onset of the voiding process. A major consideration is the incorporation of grain boundary effects, where our experimentalist collaborators have demonstrated texturing (i.e., the various crystalline grains of the Li anode synchronization in a preferred orientation) due to the calendaring process involved in anode manufacturing, which we propose can affect the critical current density at which voiding occurs.

On the other hand, the morphological instabilities observed at the interface, which suggests that the voids can be very shallow, suggests a simplification at the mesoscale, where we can treat the void region as a 2D instead of 3D domain, and investigate intriguing temporal and spatial instabilities that can be observed experimentally in either coarse-grained measurements such as open circuit voltage as a function of stripping current and time (and modeled as highly out-of-equilibrium dynamical systems), and more fine grained direct observation through cryogenic computerized tomography scans (cryo-CT).

References

J. Agier, S. Shishvan, N. Fleck, and V. Deshpande. Void growth within Li electrodes in solid electrolyte cells. *Acta Materialia*, 240:118303, 2022.

References

Z. Ahmad, V. Venturi, S. Sripad, and V. Viswanathan. Chemomechanics: Friend or foe of the "AND problem" of solid-state batteries? *Current Opinion in Solid State and Materials Science*, 26(4):101002, 2022.

P. Barai, T. Fuchs, E. Trevisanello, F. H. Richter, J. Janek, and V. Srinivasan. Study of void formation at the lithium|solid electrolyte interface. *Chemistry of Materials*, 36(5):2245–2258, 2024.

D. N. Bhate, A. F. Bower, and A. Kumar. A phase field model for failure in interconnect lines due to coupled diffusion mechanisms. *Journal of the Mechanics and Physics of Solids*, 50(10):2057–2083, 2002.

M. Feng, X. Liu, S. J. Harris, B. W. Sheldon, and Y. Qi. A multiscale model to understand the interface chemistry, contacts, and dynamics during lithium stripping. *Journal of the Mechanics and Physics of Solids*, 193:105878, 2024.

V. Goel, Y. Lyu, S. DeWitt, D. Montiel, and K. Thornton. Simulating microgalvanic corrosion in alloys using the prisms phase-field framework. *MRS Communications*, 12(6):1050–1059, 2022.

Z. Hong and V. Viswanathan. Phase-field simulations of lithium dendrite growth with open-source software. *ACS Energy Letters*, 3(7):1737–1743, 2018.

T. Krauskopf, H. Hartmann, W. G. Zeier, and J. Janek. Toward a fundamental understanding of the Lithium metal anode in Solid-State Batteries—an electrochemo-mechanical study on the garnet-type solid electrolyte $\text{Li}_{6.25}\text{Al}_{0.25}\text{La}_3\text{Zr}_2\text{O}_{12}$. *ACS Applied Materials & Interfaces*, 11(15):14463–14477, 2019.

W. S. LePage, Y. Chen, E. Kazyak, K.-H. Chen, A. J. Sanchez, A. Poli, E. M. Arruda, M. D. Thouless, and N. P. Dasgupta. Lithium mechanics: Roles of strain rate and temperature and implications for lithium metal batteries. *Journal of The Electrochemical Society*, 166(2):A89, 2019.

M. Mahadevan and R. Bradley. Phase field model of surface electromigration in single crystal metal thin films. *Physica D: Nonlinear Phenomena*, 126(3):201–213, 1999.

E. Milan and M. Pasta. The role of grain boundaries in solid-state Li-metal batteries. *Materials Futures*, 2(1):013501, 2022.

S. Mosleh, E. Annevelink, V. Viswanathan, and L. Mahadevan. Controlling moving interfaces in solid-state batteries. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 481:20240785, 2025.

S. Shishvan, N. Fleck, R. McMeeking, and V. Deshpande. Vacancy diffusion and its consequences for void growth at the interface of a stripping metal electrode and solid electrolyte. *Electrochimica Acta*, 467:143081, 2023.

S. Shishvan, N. Fleck, R. McMeeking, and V. Deshpande. Failure mechanisms at the Li anode/solid electrolyte interface during Li stripping. *Mechanics of Materials*, 193:104990, 2024.

M. F. Shojaei, R. Gulati, and K. Garikipati. A continuum, computational study of morphogenesis in lithium intermetallic interfaces in solid state batteries. *Journal of the Mechanics and Physics of Solids*, 200:106073, 2025.

J. A. Warren, T. Pusztai, L. Környei, and L. Gránásy. Phase field approach to heterogeneous crystal nucleation in alloys. *Phys. Rev. B*, 79:014204, 2009.

M. Werner and K. Weinberg. Diffusion induced void nucleation in SnPb solder joints. *PAMM*, 17(1):573–574, 2017.

A. M. Yao and V. Viswanathan. Open-circuit voltage models should be thermodynamically consistent. *The Journal of Physical Chemistry Letters*, 15(4):1143–1151, 2024.

S. G. Yoon, B. S. Vishnugopi, D. L. Nelson, A. X. B. Yong, Y. Wang, S. E. Sandoval, T. A. Thomas, K. A. Cavallaro, P. Shevchenko, E. P. Alsaç, C. Wang, A. Singla, J. R. Greer, E. Ertekin, P. P. Mukherjee, and M. T. McDowell. Interface morphogenesis with a deformable secondary phase in solid-state lithium batteries. *Science*, 388(6751):1062–1068, 2025.

H.-C. Yu, D.-H. Yeon, X. Li, and K. Thornton. Continuum simulations of the formation of Kirkendall-effect-induced hollow cylinders in a binary substitutional alloy. *Acta Materialia*, 57(18):5348–5360, 2009.

References

H.-C. Yu, H.-Y. Chen, and K. Thornton. Extended smoothed boundary method for solving partial differential equations with general boundary conditions on complex boundaries. *Modelling and Simulation in Materials Science and Engineering*, 20(7):075008, 2012.

Y. Zhao, R. Wang, and E. Martínez-Pañeda. A phase field electro-chemo-mechanical formulation for predicting void evolution at the li-electrolyte interface in all-solid-state batteries. *Journal of the Mechanics and Physics of Solids*, 167:104999, 2022.

Improvement of Hydrodynamic Properties using Kirigami Patterns & Organic Vapor Jet Deposited Coatings

DANIEL ESTERKIN, MORGAN PEN¹, JOSHUA BEHRINGER², EVA PONTRELLI, LOKESH SILWAL, RYAN HARBER¹ & MAX SHTEIN

Traditionally, *kirigami* is a Japanese art of cutting paper. It affords a means of transforming a two-dimensional sheet into three-dimensional shapes with novel properties, which our team has leveraged to create novel compliant mechanisms and surfaces with unique fluid-surface interaction (FSI) properties. At present, the understanding of the relationship between cut topology and resulting properties is rudimentary / mostly phenomenological, and predictive models are challenging to develop and validate. This project focuses on creating libraries of kirigami structures, characterizing their FSI, and providing a basis for training AI models in subsequent prediction and optimization of kirigami-enabled, novel modes of FSI.

Keywords: kirigami, flow control, vapor jet coatings

2.1 Introduction

Kirigami is the Japanese art of cutting and folding paper to create different patterns that often contain three-dimensional (3D) features. Kirigami has become increasingly relevant in the fields of 3D microfabrication and nanofabrication because of its ability to create flexible, reconfigurable, and diverse designs by a relatively straightforward application of 2D patterning technologies and thin-film materials developed for other fields in recent decades (Chen et al., 2020; Jin and Yang, 2023). In this work the advantages of kirigami were leveraged to create thin conformal layers that attach to solid surfaces to control their fluid interactions. Beyond just changing how surfaces interact with fluids, the kirigami sheets can house monolithically integrated sensors that measure useful quantities from the fluid, such as pressure and shear forces, creating a single device that both influences surface fluid interaction and measures it.

2.2 Kirigami as a flow control mechanism

In this work, kirigami patterns were designed to be placed over rigid hydrofoils to improve their hydrodynamic properties. When a rigid body such as a sphere moves through a fluid medium, it creates a pattern of vortices behind it known as a von Karman vortex street shown in Figure 2.1. This vortex street consists of alternating vortices on both sides of the sphere. For an object like the sphere, these vortices decrease the wake flow momentum of the object, thus creating drag. For an object like a flapping fin the spin is in the opposite direction, and instead of creating drag

¹University of Michigan, Ann Arbor, MI

²Greenhills High School, Ann Arbor, MI

Improvement of Hydrodynamic Properties using Kirigami Patterns & Organic Vapor Jet Deposited Coatings

they create a velocity surplus by increasing the wake flow momentum (Figure 2.1) (Bianchi, 2023; Hosseini and Ashrafizadeh, 2015). We set out to explore the possibility of improving the velocity surplus via the reverse von Karman Effect by attaching kirigami patterns to a flapping fin, thereby creating more thrust than would be created by a smooth fin.

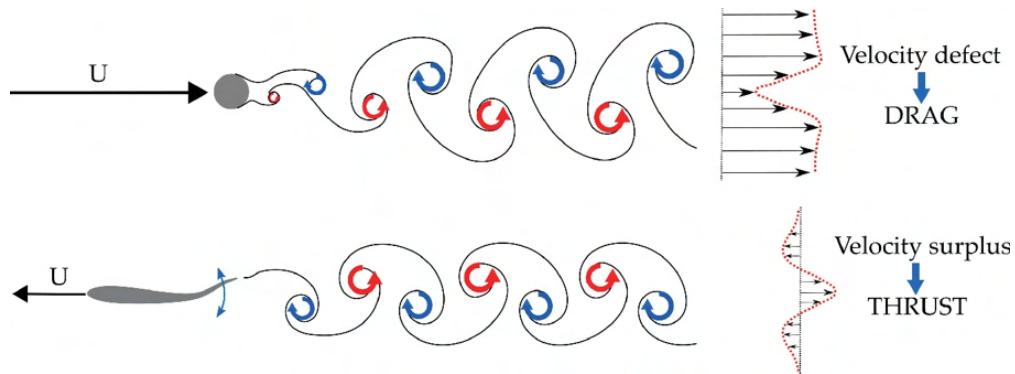


Figure 2.1: von Karman and Reverse von Karman Effect (Bianchi, 2023).

Choosing kirigami patterns

Inspirations for the designs of kirigami patterns that improve the reverse von Karman effect were taken from various aquatic lifeforms, some of which have been observed and/or theorized by prior studies to utilize these principles of action, and link them to certain surface topologies. One of the first designs created was made to mimic the pattern of fish scales (Figure 2.2).



Figure 2.2: Kirigami pattern consisting of repeated, curved cuts, each spanning 120 degree arcs.

In addition to fish scales, shark denticles are another natural structure studied for their hydrodynamic properties. Shark denticles owe their favorable hydrodynamic properties to small undulating surface protrusions called ripples (Ao et al., 2021) (Figure 2.3), which have been shown to both increase lift and decrease drag, with drag reductions reported to be up to 18.6% (Pu et al., 2016). Ripples do this by changing the turbulent boundary layer which describes how flow is either laminar or turbulent when getting close to a solid surface (Zhao et al., 2012; Landau and Lifshitz, 1987).

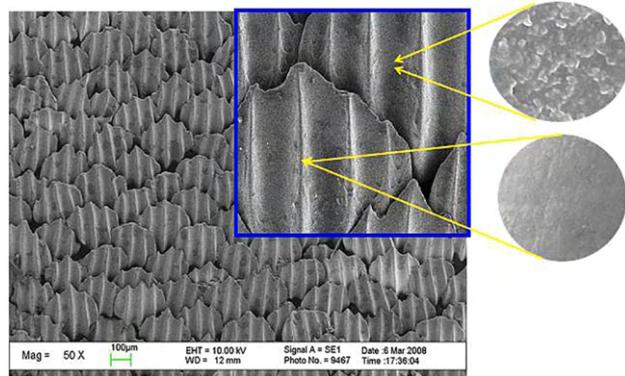


Figure 2.3: SEM image of shark denticles (Pu et al., 2016).

Various kirigami patterns

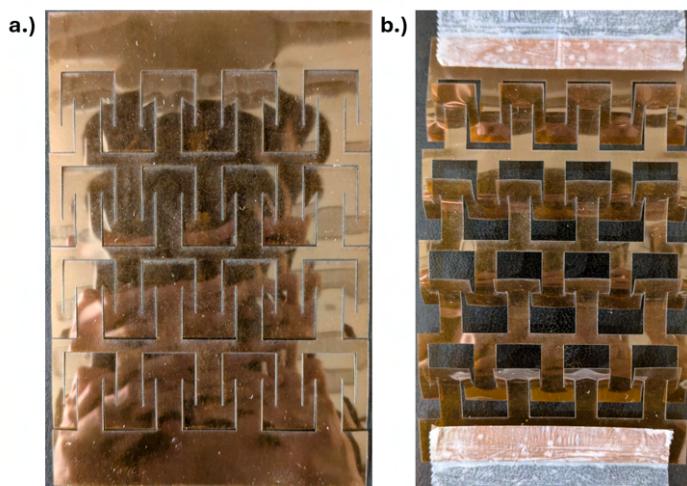


Figure 2.4: ‘Rectangular’ kirigami structure cut from Kapton polyimide film: (a) unstretched pattern, (b) stretched pattern.

Preliminary designs were cut from kapton polyimide film (Figures 2.4, 2.5). Later designs were cut from PET plastic film that better emulate the riblets observed in shark denticles (Figures 2.6, 2.7, 2.8). These structures generate a 3D profile that in principle can force flow into a turbulence pattern of counter-rotating longitudinal vortices, whose characteristic size and energy is jointly determined by the flow velocity, surface pattern, and mechanical compliance of the surface elements.

Methods

All kirigami designs were laser cut using a Universal Systems VLS3.50 CO₂ laser cutter system. Patterns were designed using Solidworks 2015 software and cut inside of a nitrogen environment. Films cut from PET plastic and kapton polyimide film were 100µm thick.

The kirigami patterns were tested by attaching them to both sides of a 3D printed hydrofoil with tape. This hydrofoil was then attached to a motor used to actuate it and placed inside a water tank. The water tank was filled with IR sediment, and motion inside the tank was recorded using a high speed camera which tracked the IR sediment’s movement. By tracking the movement of the

Improvement of Hydrodynamic Properties using Kirigami Patterns & Organic Vapor Jet Deposited Coatings

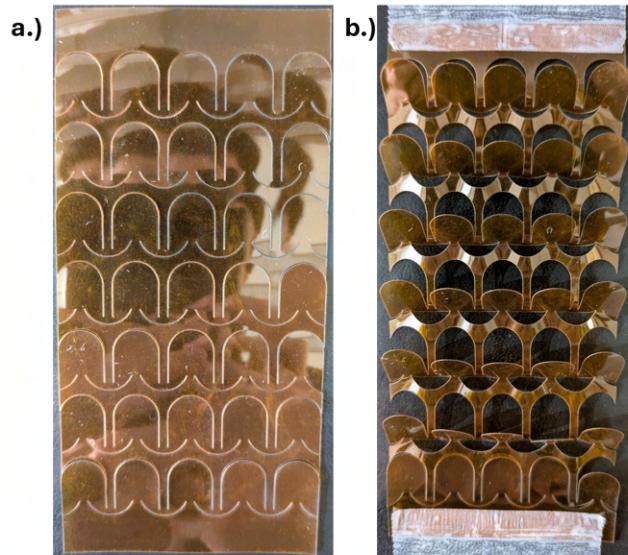


Figure 2.5: ‘Unequal Arcs’ kirigami structure cut from Kapton polyimide film: (a) unstretched pattern, (b) stretched pattern.

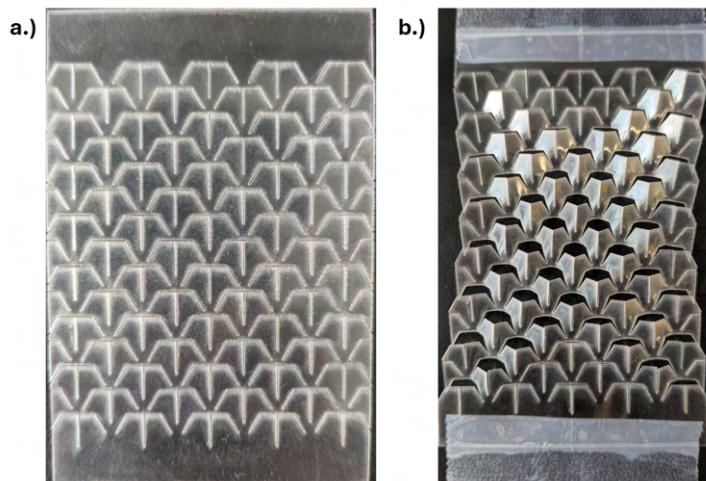


Figure 2.6: ‘Trapezoidal’ kirigami structure cut from PET plastic film: (a) unstretched pattern, (b) stretched pattern.

IR sediment, the vortices produced by the hydrofoil’s flapping can be measured in quiescent flow. The tests of all kirigami patterns were completed in collaboration with Prof. Anchal Sareen’s group at the University of Michigan, in particular with the help of Dr. Lokesh Silwal of Prof. Sareen’s group. Prof. Sareen’s group also provided the hydrofoil and constructed the accompanied testing apparatus. Figure 2.9 shows both the testing apparatus and the hydrofoil.

Results and discussion

Testing of the kirigami patterns inside of the quiescent flow chamber determined that the presence of a kirigami pattern on the hydrofoil positively affected its hydrodynamic properties. So far only the 120 degree circular arcs kirigami pattern has been tested. In these preliminary tests, it appears that for the same experimental parameters, the maximum velocity of the vortices was higher when

Improvement of Hydrodynamic Properties using Kirigami Patterns & Organic Vapor Jet Deposited Coatings

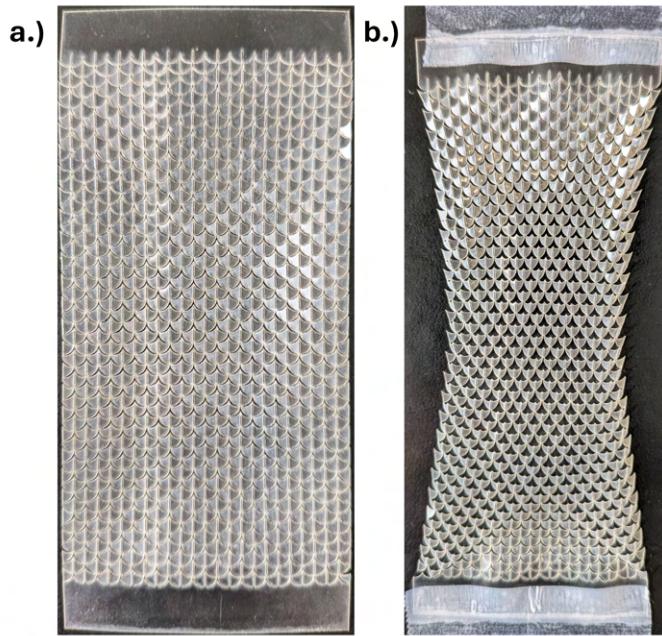


Figure 2.7: ‘Pointed Ridges’ kirigami structure cut from PET plastic film: (a) unstretched pattern, (b) stretched pattern.

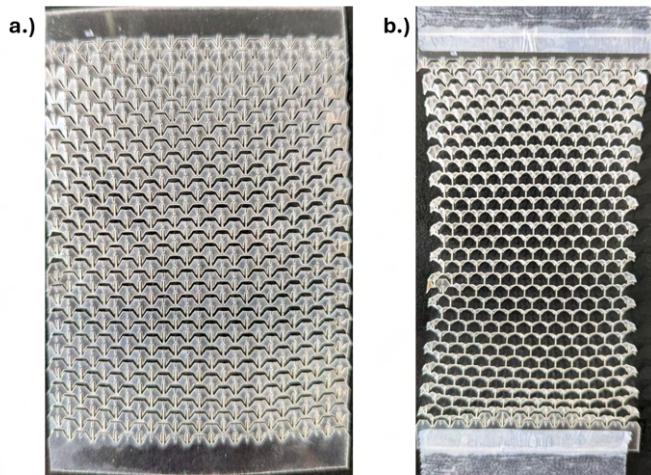


Figure 2.8: ‘Honeycomb Trapezoidal’ kirigami structure cut from PET plastic film: (a) unstretched pattern, (b) stretched pattern.

the kirigami pattern was attached, compared to that of the smooth foil. This difference can be seen in Figure 2.10. Figure 2.11 shows the qualitative differences in flow field with and without the kirigami pattern attached to the hydrofoil. Additional experiments on these and other patterns are planned to better quantify and understand the observed behavior, and to identify opportunities for functional optimization.

Improvement of Hydrodynamic Properties using Kirigami Patterns & Organic Vapor Jet Deposited Coatings

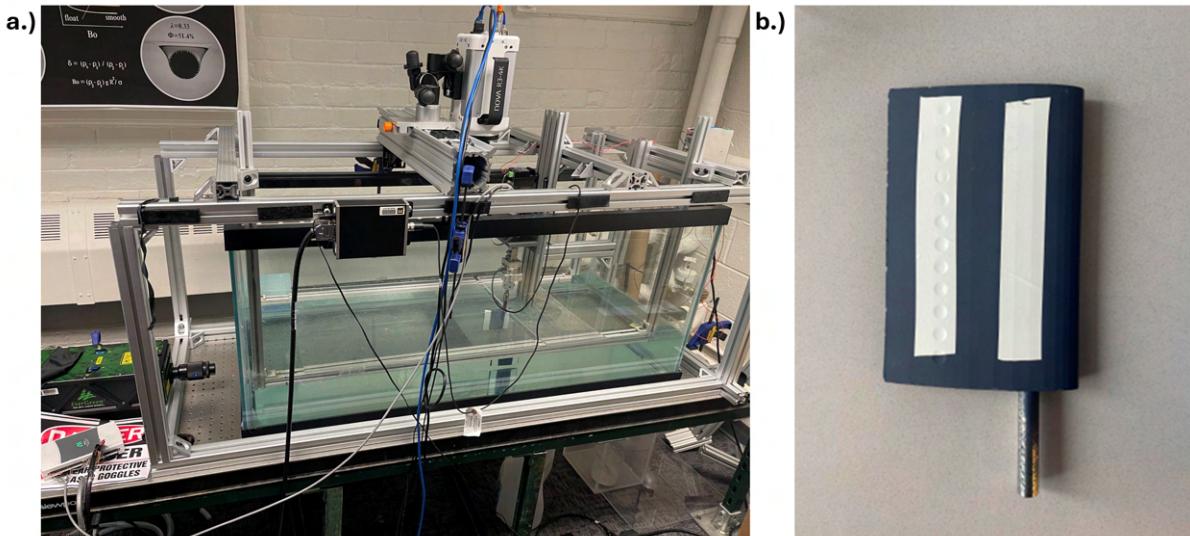


Figure 2.9: (a) Quiescent flow testing apparatus, (b) 3D printed hydrofoil.

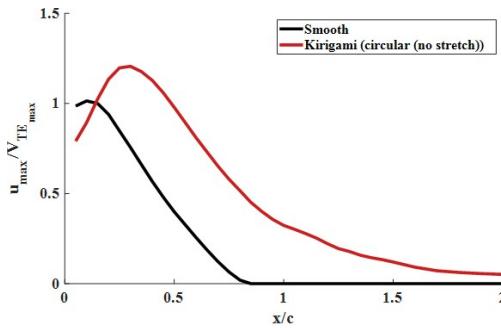


Figure 2.10: Measured velocities with kirigami pattern attached to hydrofoil and without.

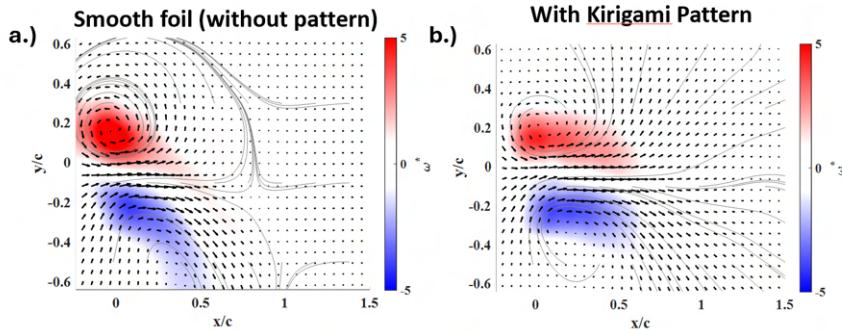


Figure 2.11: (a) Qualitative flow field without kirigami pattern attached. (b) Qualitative flow field with kirigami pattern attached.

2.3 Kirigami sensors

In our previous work, kirigami substrates have been used to house strain sensors for wearable sensors (Figure 2.12). Since strain measurements can often be transduced into pressure measurements (Sharma et al., 1987), pressure sensing mechanisms could be placed on kirigami substrates such as those used on the hydrofoil.

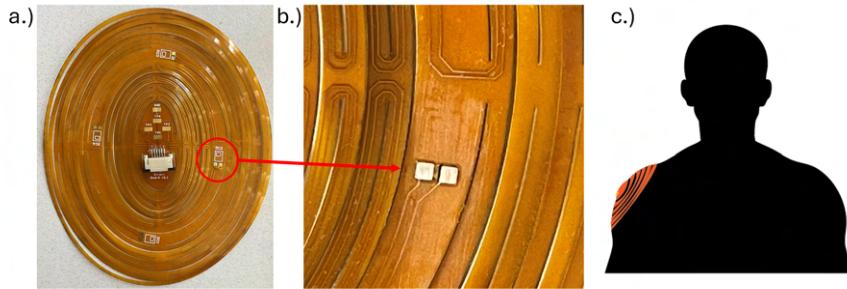


Figure 2.12: (a) Full kirigami sensor patch with strain gauge. (b) Close view of strain gauge placement location. (c) An example of a prior application of a sensor patch on a human shoulder, where movement of the arm can be reconstructed from the locally detected strain signals.

Most commercial strain gauges are made of ultra-thin traces of a metal alloy of copper and nickel called constantan (Millett et al., 1998). These strain gauges can be cumbersome and contain large solder joints, making them difficult to integrate in a truly monolithic fashion. In applications related to drag reduction and pressure sensing, these protrusions can affect surface-adjacent flow and therefore introduce artefacts into the measured pressure fluctuations, for example. For this reason, more conformal strain gauges would be desirable such as polymer strain gauges. To realize thin, monolithically integrated, and flexible strain sensors, we have recently developed polymer-based strain gauges, based on the ionically and electronically conducting polymer PEDOT:PSS, which can be mixed with dopants to improve its conductivity and durability.

Figure 2.13 shows a PEDOT:PSS strain gauge deposited on a kirigami substrate through atomized deposition on a kirigami substrate. Figure 2.14 shows the results of bending a PEDOT:PSS strain gauge by hand.

Our next steps include testing the viability of a PEDOT:PSS strain gauge in naval applications, where it can provide a means of measuring surface deformations and pressure fluctuations in the boundary layer, while also being integratable with the kirigami and other surfaces and control mechanisms.



Figure 2.13: PEDOT:PSS strain gauge deposited through atomized deposition on a kirigami substrate.

2.4 Organic vapor jet printing of hydrophobic surface coatings

Organic vapor jet printing (OVJP) was developed in the early 2000s for patterning organic semiconductors as a way of directly patterning the active layers and affords a higher material utilization efficiency over other deposition and patterning techniques (Shtein et al., 2003; Sun et al., 2005). Vapor jet printing has since been applied to active pharmaceutical ingredients to increase their surface

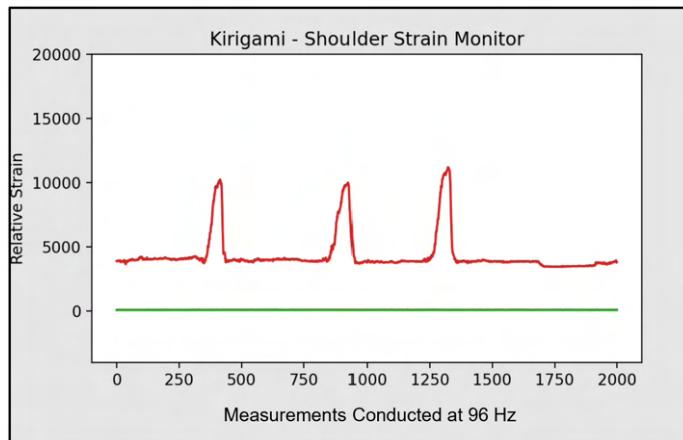


Figure 2.14: Bending a PEDOT:PSS strain gauge by hand. Horizontal axis is measurements conducted at 96 Hz. Vertical axis is modified voltage as a function of resistance, meaning an increase in the vertical axis corresponds to an increase in resistance due to bending.

area to volume ratio, thereby increasing their dissolution rate (Figure 2.15) (Shalev et al., 2017). Many of the molecular species being deposited can adopt highly irregular surface morphologies that effectively trap air and hydrophobic properties. Together with the ability to deposit such coatings in ambient conditions, this approach was attractive for testing its ability to reduce biofouling and drag of marine surfaces.

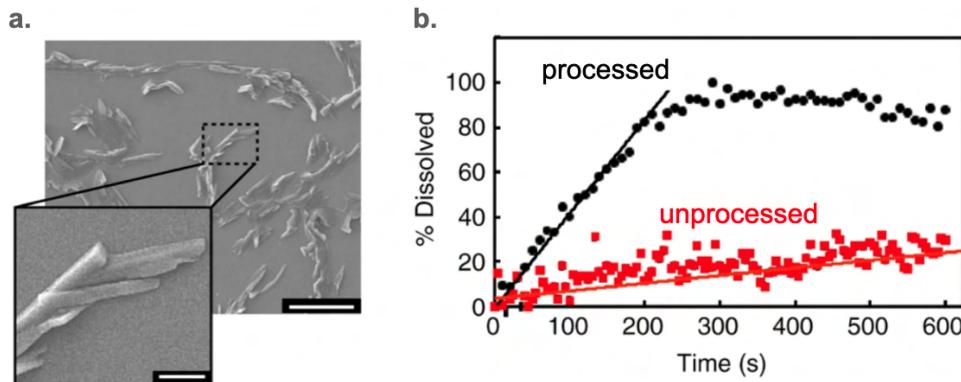


Figure 2.15: (a) Scanning electron micrograph of organic vapor jet printed tamoxifen (10 μ m scale bar, insert: 2 μ m scale bar) and (b) dissolution behavior of processed vs unprocessed tamoxifen. Reproduced from Shalev et al. (2017).

This coating technique involves heating an organic source material to produce vapor and using an inert gas to carry the vapor away from the source and direct it as a high velocity jet onto the target surface, where the organic material condenses (Figure 2.16).

Many small molecular organics are hydrophobic, and while OVJP can improve their wetting and/or dissolution properties in some cases, in others, it can produce morphologies with added geometric hydrophobicity. In this work, we attempt to coat hydrofoils with hydrophobic morphologies to reduce drag.

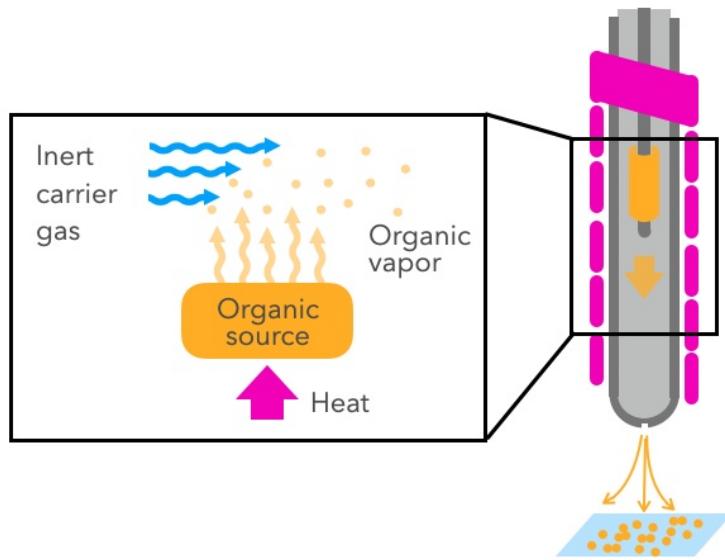


Figure 2.16: OVJP concept.



Figure 2.17: (a) Optical microscope image of organic vapor jet printed atovaquone with large (100s μ m needles). When the atovaquone deposit is submerged in water, (b) it holds an air pocket, and (c) is visibly dry after being removed from the water.

Experimental setup

Figure 2.18 gives a detailed look at the OVJP apparatus. The nozzle replaces a standard extruder block on a 3-D printer. The 3-D printer platform which allows motion in the x , y , and z directions. The temperature of the nozzle is controlled by a heating block, in which two heating cartridges are mounted. The thermocouple from the temperature controller is inserted, monitoring the evaporation temperature. Nitrogen carrier gas is admitted at a rate controlled by a calibrated mass flow controller. Inside the nozzle is the organic source mounted on the thermocouple housing.

Morphology control

Organic vapor jet deposited material is typically comprised of surface-bound micro or nano particles which usually have needle-like (Figure 2.19, panels a and b) or spherical (Figure 2.19, panels c and d) morphologies, depending on the material and deposition conditions.

Varying deposition parameters allows control of morphological features like needle size and degree of branching. Figure 2.20 shows atovaquone deposited with different particle sizes and amounts of branching.

Improvement of Hydrodynamic Properties using Kirigami Patterns & Organic Vapor Jet Deposited Coatings

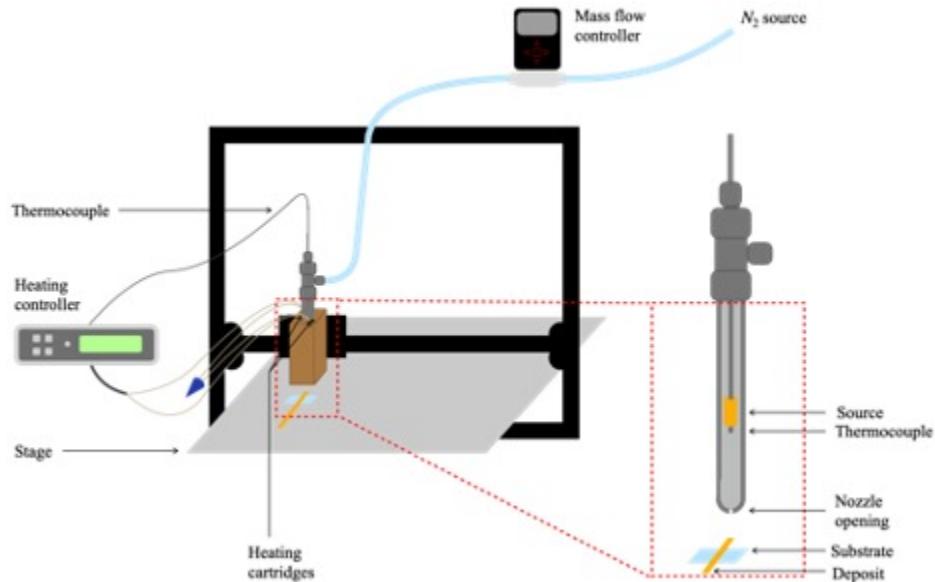


Figure 2.18: Schematic of organic vapor jet printing apparatus.

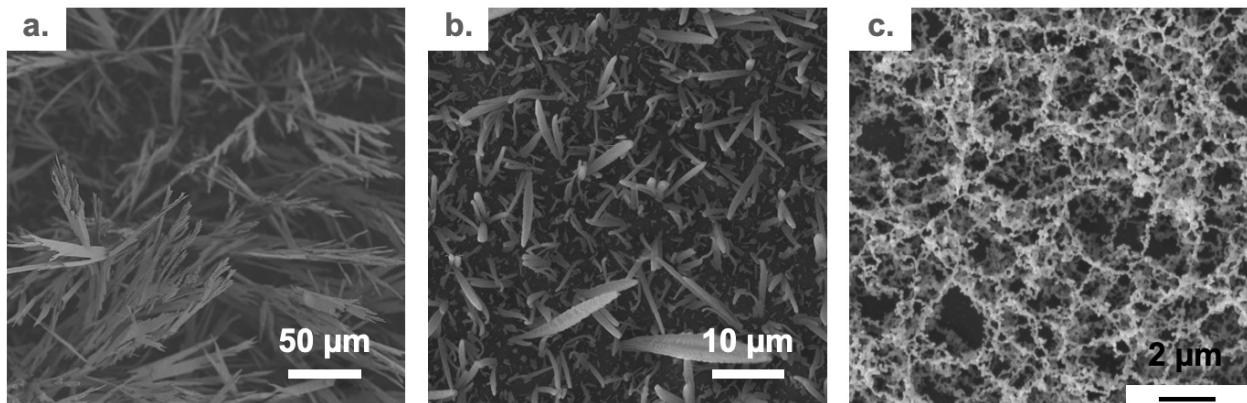


Figure 2.19: Scanning electron micrographs of organic vapor jet printed materials, including highly hydrophobic compounds, such as: (a) atovaquone with relatively high mass per area, (b) atovaquone deposited at a lower surface coverage, (c) griseofulvin.

Contact angle measurements

Vapor jet printed materials show significantly higher contact angles than flat samples of the same material (Figure 2.21) for both spherical and needle morphologies. Contact angle measurements were taken in Prof. Albert Liu's Lab in the Chemical Engineering Department at University of Michigan with assistance from Jae Wan Lee (Shtain Lab).

Hydrofoil coating

Small strips of indigo blue, approximately 1cm in width, were deposited on a 3D printed PLA polymer hydrofoil (Figure 2.22) to determine if organic vapor jet printed material would display poor wetting properties in situ, and if the films would be robust enough to undergo testing in the setup described in section 2.3.

While as-deposited films showed significant water beading compared with the bare PLA surface

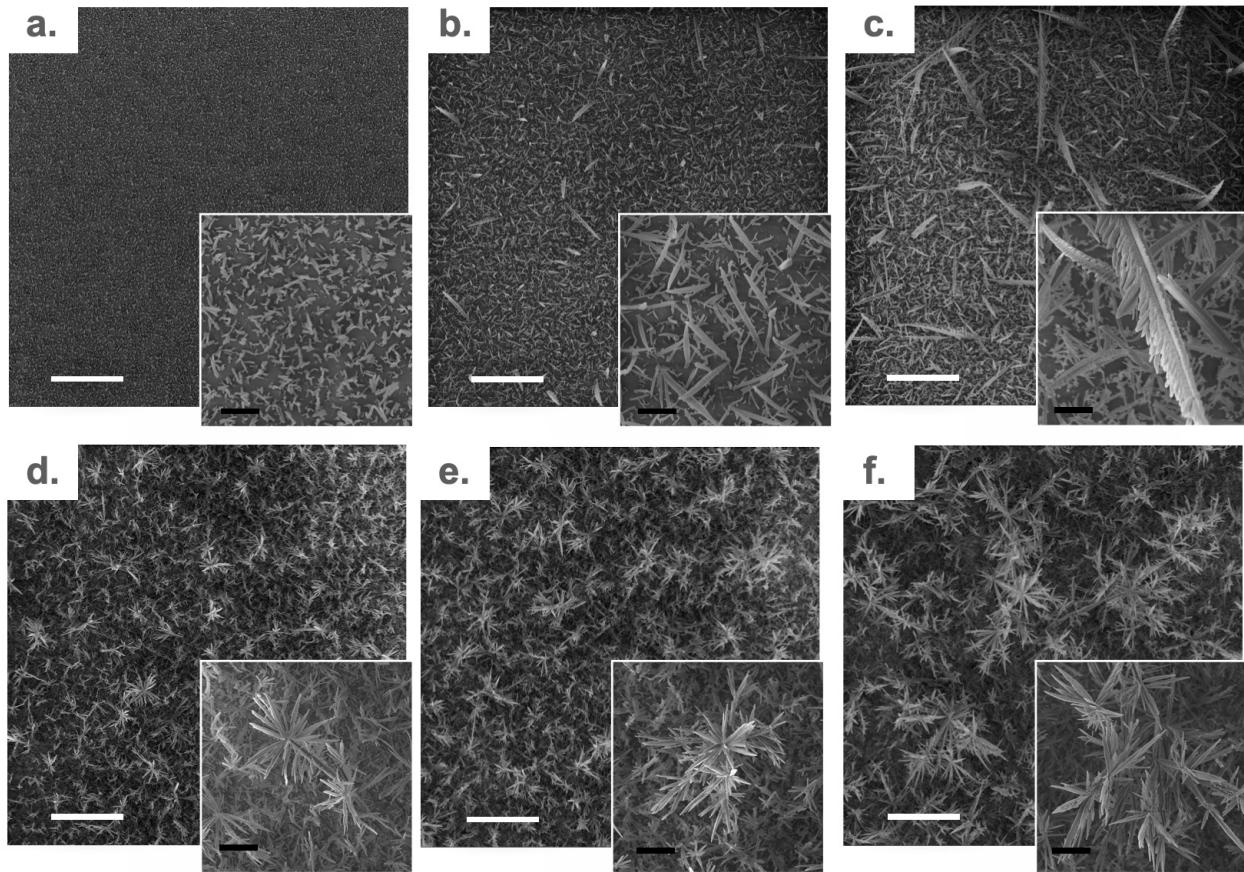


Figure 2.20: (a - c) Scanning electron micrographs of atovaquone deposited with decreasing (left to right) nozzle scan velocity, resulting in larger needle growth. White scale bars are 50 μm , black scale bars are 5 μm . (d - f) Scanning electron micrographs of atovaquone deposited with 1 (d), 3 (e), and 5 (f) passes of the nozzle over one location. Increasing the number of passes leads to increased branching. White scale bars are 50 μm , black scale bars are 5 μm .

(Figure 2.23, panels a and b vs d), a film which had been gently wiped with a piece of tissue showed only slight improvement from the bare hydrofoil surface (Figure 2.23. panel c vs panel d). Repeatedly adding and removing water drops from the as deposited strip in panel a of Figure 2.23 resulted in drop appearance similar to that in panel c of Figure 2.23 after only 3 or 4 drops had been added and allowed to roll away.

Freshly deposited films repelled water successfully, but particles washed away relatively quickly, leaving a less hydrophobic surface behind. As-deposited films appear to have sub-optimal adherence to withstand extended use at high flow velocity, and an adhesion promoting treatment and/or a protective polymer coating should be attempted to strengthen the organic vapor jet printed material and take advantage of the geometric hydrophobicity of the OVJP deposits. To this end, we may use vapor jet-deposited parylene and other materials, where a surface polymerization reaction produces a conformal and mechanically more robust coating.

2.5 Summary

Nature has evolved different strategies for controlling fluid-surface interactions that can benefit US Navy capabilities. In this project, we are exploring two such strategies for reducing drag and acoustic signatures: 1) easily manufactured surface textures that can trigger vorticity at certain

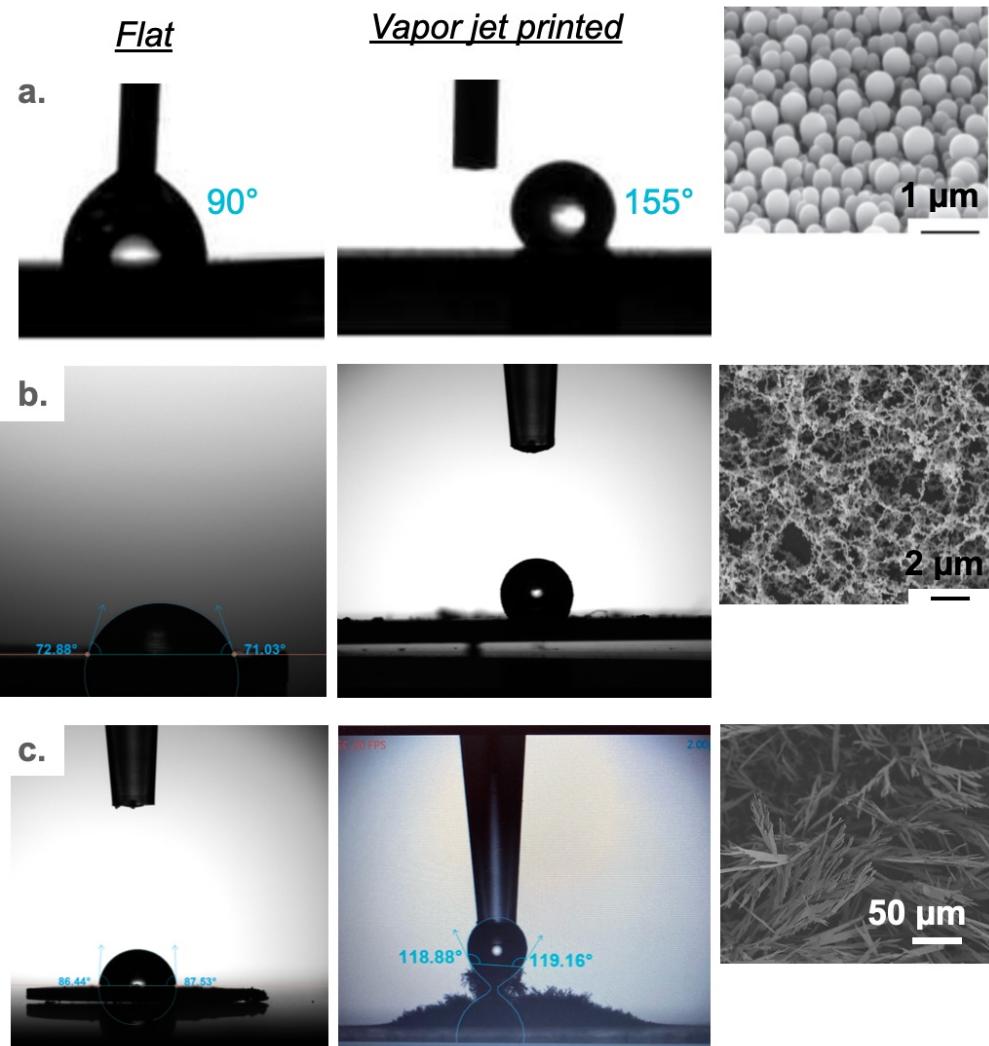


Figure 2.21: Contact angle measurements for flat (left) vs. organic vapor jet printed material (center) with scanning electron micrographs of the organic vapor jet printed material (right). The organic materials here comprised (a) Sub-phthalocyanine chloride (Shalev et al., 2017) (b) griseofulvin, and (c) atovaquone.

length and energy scales, 2) easily applied surface coatings that have texture and composition that effectively traps air and increases water contact angle.

How surface geometry influences FSI is currently only understood empirically, and prediction remains elusive. We are using kirigami, a Japanese art in which a pattern of cuts in a 2D sheet transforms the sheet into a 3D structure upon stretching. This allows for very fast, easy production of patterns with systematic variation in feature size, shape, spacing, etc. This in turn allows us to build a very large library of patterns with repeating 3-dimensional features, and test the samples to catalogue their fluid surface interactions (FSI), forming the basis of a directed evolution-style optimization approach. The catalogue of FSI, with the subset of desirable FSI, can then be used to train a machine learning (ML) / artificial intelligence (AI) model to later help with the inverse design problem, in which optimal geometries are suggested based on desired FSI. So far, we have created dozens of patterns for the texture library, characterized their 2D and 3D transformations, and began fluid dynamic testing. Initial results on some of the patterns have indicated substantial



Figure 2.22: Strips of indigo blue deposited on a 3D printed hydrofoil. Left and right are relatively thick coatings, center is a thinner coating, and the inset shows what the right strip looks like after gently wiping the deposit with a bit of tissue (to mimic light mechanical agitation).

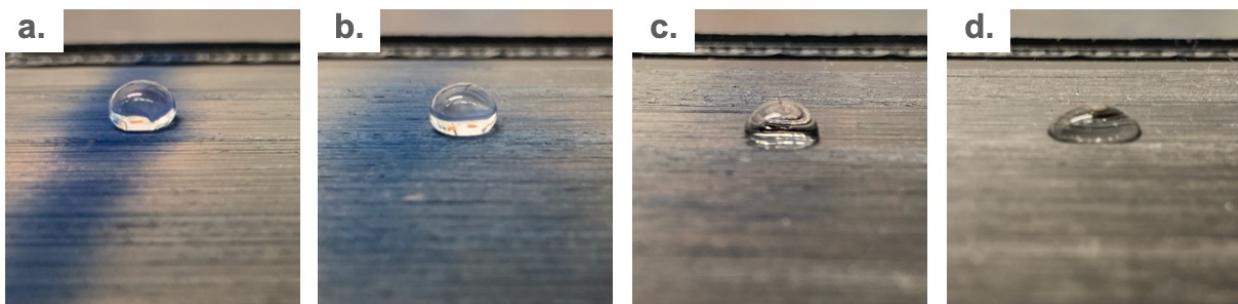


Figure 2.23: Drops of water on (a) a thick layer of indigo blue, (b) a thin layer of indigo blue, (c) a thick layer of indigo blue wiped with a tissue, and (d) the bare surface of the 3D printed hydrofoil.

reductions in drag and noise under some flow conditions.

The second goal is to use geometrically hydrophobic coatings to trap an air barrier resulting in reduced drag. This effect has been observed previously, except that practical means of coating large surfaces at low cost with appropriately textured coatings are still lacking. Our lab has been generating coatings by a dry, organic vapor jetting process, in which organic vapor is directed onto a surface by a high speed inert gas, producing dense mats of micron-scale particles. The coating conditions are modified to produce different morphologies, which in turn modify the contact angle. Thus far, we have shown control over deposited morphology at ambient conditions, and observed large increases in contact angle. We also began applying a protective barrier coating to increase the durability of these coatings while maintaining the gains in contact angle.

References

M. Ao, M. Wang, and F. Zhu. Investigation of the turbulent drag reduction mechanism of a kind of microstructure on riblet surface. *Micromachines*, 12(1):59, 2021.

G. Bianchi. A numerical tool for the analysis of bioinspired aquatic locomotion. *SpringerBriefs in Applied Sciences and Technology*, 2023.

S. Chen, J. Chen, X. Zhang, Z.-Y. Li, and J. Li. Kirigami/origami: Unfolding the new regime of advanced 3d microfabrication/nanofabrication with “folding”. *Light: Science & Applications*, 9(1), 2020.

References

A. A. Hosseini and A. Ashrafizadeh. Numerical simulation of the wake structure and thrust/lift generation of a pitching airfoil at low reynolds number via an immersed boundary method. *Journal of Aerospace Technology and Management*, 7(3):334–350, 2015.

L. Jin and S. Yang. Engineering kirigami frameworks toward real-world applications. *Advanced Materials*, 36(9), 2023.

L. D. Landau and E. M. Lifshitz. Boundary layers. *Fluid Mechanics*, pages 157–191, 1987.

J. Millett, N. Bourne, and Z. Rosenberg. The piezoresistance of constantan strain gauges under shock loading conditions. *Journal of Physics D: Applied Physics*, 31(9):1126–1130, 1998.

X. Pu, G. Li, and H. Huang. Preparation, anti-biofouling and drag-reduction properties of a biomimetic shark skin surface. *Biology Open*, 5(4):389–396, 2016.

O. Shalev, S. Raghavan, J. M. Mazzara, N. Senabulya, P. D. Sinko, E. Fleck, C. Rockwell, N. Simopoulos, C. M. Jones, A. Schwendeman, and et al. Printing of small molecular medicines from the vapor phase. *Nature Communications*, 8(1), 2017.

J. K. N. Sharma, K. K. Jain, and G. F. Molinar. Use of strain gauge pressure transducer as working pressure standard up to 500 mpa. *Sensors and Actuators*, 11(3):275–282, 1987.

M. Shtain, P. Peumans, J. B. Benziger, and S. R. Forrest. Micropatterning of small molecular weight organic semiconductor thin films using organic vapor phase deposition. *Journal of Applied Physics*, 93(7):4005–4016, 2003.

Y. Sun, M. Shtain, and S. R. Forrest. Direct patterning of organic light-emitting devices by organic-vapor jet printing. *Applied Physics Letters*, 86(11), 2005.

D.-Y. Zhao, Z.-P. Huang, M.-J. Wang, T. Wang, and Y. Jin. Vacuum casting replication of micro-riblets on shark skin for drag-reducing applications. *Journal of Materials Processing Technology*, 212(1):198–202, 2012.

On the Optimal Performance of Oscillating Surge Wave Energy Converters

BINH D. TRUONG & LEI ZUO

The sequential process is a conventional procedure in designing an ocean wave energy converter, from mechanical systems to electrical loading. However, many studies do not consider the reciprocal influences between the two domains. Therefore, the potential of a wave energy converter has not been properly explored. In this paper, the performance of an oscillating surge wave energy converter driven by regular wave excitation is investigated, in which the dynamics of subsystems and their couplings are fully taken into account. We analytically determine the maximum possible power that can be harvested for a specific geometry of the wave capture structure. We further show that these geometric dimensions and the rest of the system can be optimized within a unified framework to maximize the output power under given ocean wave characteristics. The upper bound on power is expressed as a function of an effective figure of merit for the power take-off of the wave energy converter, which combines electromagnetic transducer coupling and parasitic losses. We also find that for complex systems with multiple energy conversion stages, such as wave energy converters, the gradient descent method is more appropriate for optimizing the objective variables compared to the impedance matching principle. In these cases, the latter approach, although widely employed in the literature, does not yield the global maximum power delivered to the load. We derive analytical solutions to system parameters, including the complex load, mechanical transmission, and generator moment of inertia. These findings provide a general framework for optimizing the performance of an oscillating surge wave energy converter and similar architectures.

Keywords: ocean wave energy converter, power optimization, gradient descent method, impedance matching technique

3.1 Introduction

The vast availability of wave energy, approximately 2.11 TW of average power in the entire globe (Gunn and Stock-Williams, 2012; Reguero et al., 2015), or an estimated theoretical wave energy potential of nearly 29,500 TWh/year (Khojasteh et al., 2023), demonstrates its potential as a clean and abundant resource that could provide 60% to 100% of global electricity demand. Various wave energy converter architectures have been developed to capture the kinetic and potential energy of ocean waves and convert them into useful electrical energy (López et al., 2013; Guo and Ringwood, 2021; Zhang et al., 2021; Gallutia et al., 2022). Among these architectures, oscillating surge wave energy converters have gained attention for their ability to efficiently harness wave energy from the surge motion of ocean waves in nearshore regions (Sell et al., 2018; Ghasemipour et al., 2022; Ahmed et al., 2024b; Mi et al., 2024). Differing from point absorbers, whose dimensions and efficiency are fundamentally constrained by the incident wave wavelengths, oscillating surge wave energy converters utilize a hinged flap mechanism that pivots about a base anchored to the seabed

or a floating platform, enabling greater scalability and improved adaptability to a range of wave conditions.

In general, the design of a wave energy converter involves complex procedures (Barbarelli et al., 2021; Harms et al., 2022; Paduano et al., 2024). It includes computing the fluid–structure interactions based on hydrodynamic principles, optimizing the geometry of the wave energy capture device, maximizing the efficiency of the power take–off system, and identifying optimal control strategies. These processes have been iterative and sequential, focusing separately on individual objectives (Ruellan et al., 2010; Trueworthy and DuPont, 2020; Giorcelli et al., 2022). However, this approach often leads to suboptimal solutions because it neglects the coupling effects and interdependencies among subsystems. In contrast, a co–design method simultaneously considers the dynamics of the entire system to maximize overall energy conversion (Coe et al., 2020; Giannini et al., 2022; Peña-Sánchez et al., 2022; Ströfer et al., 2023; Grasberger et al., 2024; Yang et al., 2024, 2025). It ensures that design rules for one subsystem align with the objectives of others, thereby enhancing system performance. While several studies have explored various aspects of co–design in wave energy converters, comprehensive methodologies and generalized principles remain limited in the literature. This limitation arises from the fact that these studies were concerned with specific numerical examples of wave converters, which could have been chosen differently and potentially led to different conclusions.

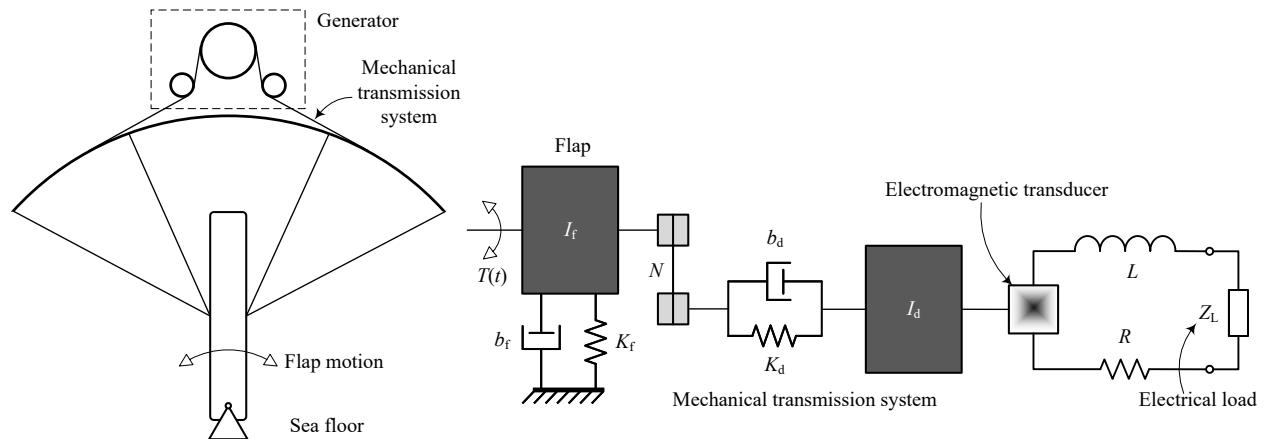


Figure 3.1: Concept of an oscillating surge energy converter using a flap as the wave capture structure.

In the context of wave energy harvesting, ocean waves are usually categorized into two types, regular and irregular (Li and Lin, 2012; Stansby et al., 2015; Young, 2017). Regular waves are idealized representations often used in controlled environments, characterized by harmonic patterns with defined wave height and period. On the contrary, irregular waves more accurately represent ocean conditions, which exhibit randomness and variability in height, period, and direction due to the complex interactions of multiple waves and other environmental factors. Despite the predominance of irregular waves in the ocean, exploring regular waves remains highly valuable in the design and optimization of wave energy converters (Peng et al., 2020; Hollm et al., 2022; Liu et al., 2023; Barua and Salauddin Rasel, 2024). These investigations provide a simplified framework that enables engineers to understand fundamental interactions between ocean waves and wave energy harvesters, optimize system parameters, and establish standard metrics for performance evaluation. The results obtained from considering power harnessed under regular waves are insightful for both refining mathematical models and experimental setups for the entire system. This approach thus facilitates the development of robust and efficient wave energy converters that can operate reliably in the complex conditions encountered in real–world applications.

The primary objective of this paper is to determine the maximum possible power that can be harvested by an oscillating surge wave energy converter under a given torque strength induced by regular ocean waves. This optimal power, here referred to as the physical power bound, can be utilized as a theoretical benchmark to evaluate the effectiveness of an energy converter, offering insight into the system performance in comparison with an ideal. In the following sections, different power optimization techniques are explored, aiming to (i) identify the most suitable method for a complex architecture such as a wave energy converter, (ii) provide a comprehensive understanding of the conditions under which system parameters need to be satisfied to achieve the output power bound, and (iii) determine whether reaching this bound is practically feasible. Generalized principles on design rules for each subsystem of a wave energy converter are also discussed, with a focus on optimizing the extracted power.

3.2 Modeling of oscillating surge wave energy converters

Structure and principle of operations

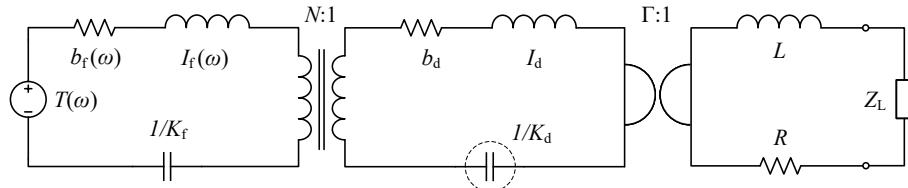


Figure 3.2: Equivalent circuit model for an oscillating surge energy converter.

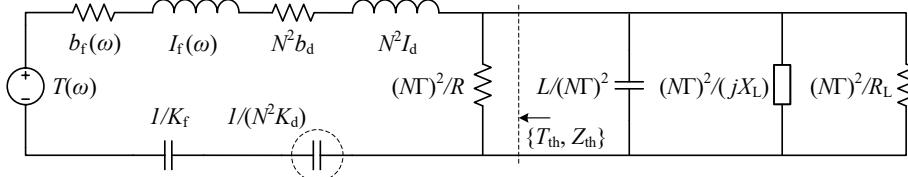


Figure 3.3: Alternative equivalent circuit model based on reflected impedance theory.

A typical configuration of an oscillating surge wave energy converter is shown in Figure 3.1 (Mi et al., 2024). It consists of a structure employed to capture energy of the ocean waves, usually referred to as a flap, a mechanical transmission system for velocity transfer and amplification, and a linear electromagnetic generator functioning as a power take-off device. The generated electrical energy is eventually delivered to the utility grid or loads. A simple and widely used geometry of the flap is rectangular, which oscillates about the bottom edge anchored to the sea floor during operation. For the wave energy converter under consideration, its dynamics can be described using one mechanical coordinate (single degree of freedom). The flap is characterized by an effective moment of inertia $I_f(\omega)$, an overall damping $b_f(\omega)$, and a linear hydrostatic stiffness K_f , where ω is the angular frequency of the wave. $I_f(\omega)$ is the sum of the physical moment of inertia I and the added moment of inertia $I_a(\omega)$, i.e., $I_f(\omega) = I + I_a(\omega)$. $b_f(\omega)$ may include both linear viscous damping and radiation damping terms (Ahmed et al., 2024b). In general, the added moment of inertia $I_a(\omega)$ and the radiation damping coefficient $b_f(\omega)$ are frequency-dependent. These parameters can be numerically obtained through boundary element method (BEM) solvers such as NEMOH (Kurnia and Ducrozet, 2023), an open-source tool, Capytaine (Ancellin and

Dias, 2019), a Python-based package, or WAMIT, a commercial software. The modeling of fluid–structure interaction between the flap and ocean waves is complex, requiring solving the coupled equations governing fluid flow and structural dynamics. This objective is beyond the scope of the paper but can be found in other works, such as (Ahmed et al., 2024b; Penalba et al., 2017; Winter and Motley, 2020; Ahmed et al., 2024a).

The flap is coupled to the electric generator by a mechanical transmission system composed of a belt mechanism with a linear transduction factor (also known as the gear ratio) of N , assuming no slippage occurs. The mechanical part of the drivetrain and generator is equivalently represented by a moment of inertia I_d , a damping coefficient b_d , and a mechanical stiffness K_d that is included for a general theoretical investigation. In the electrical port, a pickup coil with an inductance of L is placed in proximity to permanent magnets to form an electromagnetic transducer. The inevitable conductive losses of the coil are accounted for by a resistance R in series with the coil inductance. The coupling between the mechanical and electrical domains is modeled as a linear transducer, with an electromagnetic (electrodynamic) transduction factor of Γ . Under the excitation of the waves, the flap oscillates in a rotational trajectory. Its kinetic energy is mechanically transferred to the generator shaft and subsequently converted to electricity through Faraday’s law of induction.

In the final stage of the energy conversion process, the power produced by the wave energy converter is directed to a storage element or the alternating current (AC) grid through electronic interfaces or a power management and control unit. In order to simplify the analysis without loss of generality, a general complex load of the form $Z_L = jX_L + R_L$ is considered, where $\{X_L, R_L\}$ are real quantities and $R_L > 0$. While Z_L does not exactly represent a realistic interface circuit, it can be considered as the equivalent input impedance of the circuit to be supplied, which is useful for identifying the upper limit on the performance of the ocean wave energy converter. The analysis in this paper applies to a linear wave energy converter modeled by deterministic equations and driven by regular (monochromatic) waves with a precisely known angular frequency. As a consequence, the moment of inertia I_f and mechanical damping coefficient b_f are assumed to be constant at the given operating frequency, and the external torque takes the form of $T = T_0 \cos(\omega t)$. The performance evaluation of a wave energy conversion system with non-linear properties and a wave converter driven by irregular waves is left open for future study. In the latter scenario, the waves can be approximated by non-white or colored noise, characterized by a frequency-dependent power spectral density. This case may require a hybrid modeling approach where the drive torque and relevant parameters such as $I_f(\omega)$ and $b_f(\omega)$ are described by stochastic processes.

Equivalent circuit models

An equivalent circuit model is formulated based on impedance analogies, also known as Maxwell’s analogy. The mechanical domain is represented in the $e - V$ convention where the torque is an across variable and the velocity is a through variable. The mechanical transmission system (often referred to as drivetrain) and the electromagnetic generator are modeled as a linear two-port transformer and gyrator, respectively. This circuit model is depicted in Figure 3.2. Given the well-established theories, mathematical models, and analytical techniques in electrical systems compared to other energy domains, unifying the entire energy conversion system into a single electrical domain can simplify the analysis procedure. This approach also provides a common framework for efficiently analyzing different components and subsystems. An alternative circuit model is further developed utilizing the reflected impedance theory, in which the transducer constitutive relationships are embedded in the effective circuit components. For instance, as shown in Figure 3.3, the generator inductance L is transformed into a capacitance with a value of $L/(N^2\Gamma^2)$. As a representative example, assuming that the load encompasses a capacitance C connected in series with the resis-

tance R_L , their reflected impedances are an inductance of $CN^2\Gamma^2$ and a resistance of $N^2\Gamma^2/R_L$, respectively, which are now in parallel with each other and with the capacitance $L/(N^2\Gamma^2)$. This purely electrical circuit constituted solely of passive elements is convenient for the analysis in the subsequent sections.

In the mechanical transmission model, the transduction factor N can be chosen as either negative or positive to establish consistency between the torque and angular velocity sign conventions and the assumption that power flow is defined as positive into the transformer at both ports (Tilmans, 1996). However, the circuit model in Figure 3.3 shows that the effective impedance parameters are always quadratic functions of N . Therefore, the sign of N does not affect the dynamics of the entire system or the output power, and from now on, we assume that $N > 0$. For the linearized electromagnetic transducer, two sign conventions are commonly adopted (Kuo, 1967; Rowell and Wormley, 1997). With the definition that power is positive into each port of the transducer, the motor torque constant and the motor back electromotive force (back EMF) voltage constant are equal in absolute value but opposite in sign, i.e., $\pm|\Gamma|$. Alternatively, a sign convention where electrical power into the transducer is positive and mechanical power out is positive can be used, making these two constants identical. Similar to the drivetrain modeling, the electrical elements in Figure 3.3 are quadratic functions of Γ , which eliminates the difference between the two sign conventions and the significance of the sign of Γ . This allows a single mathematical structure of the system dynamics to cover both, and the same form of output power can be obtained regardless of the sign convention in use. In the models formulated in this paper, the electrodynamic transduction factor Γ is treated as a positive parameter.

3.3 Theoretical power bound

In this section, we seek to find the maximum possible power that can be delivered to an electrical load using the Thévenin equivalent circuit and network theory. Here, we adopt the perspective of the power available from the source to assess a strict upper bound on the system performance. Practical limitations and the feasibility of the outcomes are not considered at this stage. For instance, resonance conditions may not be achievable due to various physical and implementation constraints. However, the theoretical limit derived can serve as a reference for evaluating the effectiveness of a given wave energy converter.

In Appendix 3.7, we have shown that the maximum power transmitted through a lossless network is equal to the power available from the source that supplies it. Without loss of generality, the internal inductance L of the generator is a lossless element and can be incorporated in an effective load, $Z'_L = j\omega L + Z_L$. Thus, the Thévenin equivalent circuit corresponding to the model in Figure 3.3 is developed in Figure 3.4. The Thévenin equivalent torque and impedance are defined as follows,

$$T_{th} = T_0 \frac{N^2\Gamma^2}{R} \left/ \left(b_f + N^2 b_d + \frac{N^2\Gamma^2}{R} + j\omega I_f + j\omega N^2 I_d + \frac{K_f}{j\omega} + \frac{N^2 K_d}{j\omega} \right) \right., \quad (3.1)$$

$$Z_{th} = \frac{N^2\Gamma^2}{R} \left(b_f + N^2 b_d + j\omega I_f + j\omega N^2 I_d + \frac{K_f}{j\omega} + \frac{N^2 K_d}{j\omega} \right) \left/ \left(b_f + N^2 b_d + \frac{N^2\Gamma^2}{R} + j\omega I_f + j\omega N^2 I_d + \frac{K_f}{j\omega} + \frac{N^2 K_d}{j\omega} \right) \right.. \quad (3.2)$$

Based on the network theory, the maximum possible power that can be delivered to an electrical

load, which is referred to as the physical power bound, can be determined by

$$P_B = \max \left\{ P_{\text{avs}} = \frac{|T_{\text{th}}|^2}{8\Re\{Z_{\text{th}}\}} \right\}. \quad (3.3)$$

Substituting (3.1) and (3.2) into the expression of P_{avs} in (3.3), we arrive at

$$P_{\text{avs}} = T_0^2 N^2 \Gamma^2 \left/ \left\{ (b_f + N^2 b_d)(N^2 \Gamma^2 + R N^2 b_d + R b_f) + R \left[\frac{K_f + N^2 K_d}{\omega} - \omega(I_f + N^2 I_d) \right]^2 \right\} \right. \quad (3.4)$$

Neglecting practical considerations, we assume that multiple system parameters can be adapted, including inertia, stiffness, and transduction factor. The second term in the denominator of P_{avs} is always non-negative, and must be equal to zero to maximize P_{avs} , yielding

$$K_f + N^2 K_d = \omega^2(I_f + N^2 I_d), \quad (3.5)$$

which implies that the wave energy converter operates under an effective resonance condition. The optimal mechanical transmission factor N is determined by the equation $\partial P_{\text{avs}} / \partial N = 0$, which results in

$$N^4 = \frac{R b_f^2}{\Gamma^2 b_d + R b_d^2} = \left(\frac{\mathcal{M}_1}{\mathcal{M}_2} \right)^2 \frac{1}{\mathcal{M}_1 + 1} \quad (3.6)$$

where the non-dimensional parameters are defined as

$$\mathcal{M}_1 = \frac{\Gamma^2}{R b_d}, \quad \mathcal{M}_2 = \frac{\Gamma^2}{R b_f}.$$

The upper bound of the output power is then given by

$$P_B = P_0 \frac{\sqrt{\mathcal{M}_1 + 1} - 1}{\sqrt{\mathcal{M}_1 + 1} + 1} \quad (3.7)$$

where we define the critical power

$$P_0 = \frac{T_0^2}{8 b_f}. \quad (3.8)$$

As shown in Figure 3.5, the power P_B is always less than P_0 for all positive \mathcal{M}_1 , and P_B is a strictly monotonic increasing function of \mathcal{M}_1 . P_0 can be regarded as the power limit of a theoretically ideal system, attained when one or more of the following conditions are met, resulting in a very large value of \mathcal{M}_1 : The mechanical and electrical loss factors b_d and R are negligibly small, and the electromagnetic transduction factor Γ is high. \mathcal{M}_1 can be interpreted as a figure of merit for the power take-off of a wave energy converter, which should be as large as possible to maximize the output power. At the limit, $P_B \rightarrow P_0$ when $\mathcal{M}_1 \rightarrow \infty$, and approximately $P_B \approx P_0$ if $\mathcal{M}_1 \gg 1$. \mathcal{M}_2 is an intermediate parameter that, along with \mathcal{M}_1 , identifies the optimal value of the mechanical transmission factor N . Once the wave energy converter is fabricated and its intrinsic parameters are identified, no electronic circuit interface or control unit can extract more power than P_B .

It is essential to emphasize that, given the characteristics of the waves such as wave height and frequency, the maximum scavenge power is defined by the input excitation torque T_0 , the overall damping coefficient of the flap configuration b_f , and the effective figure of merit \mathcal{M}_1 . Specifically, the power output P_0 is proportional to T_0^2 and inversely proportional to b_f . Since T_0 and b_f are

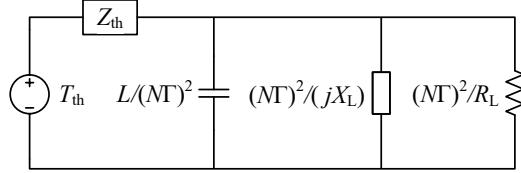


Figure 3.4: Thévenin equivalent circuit.

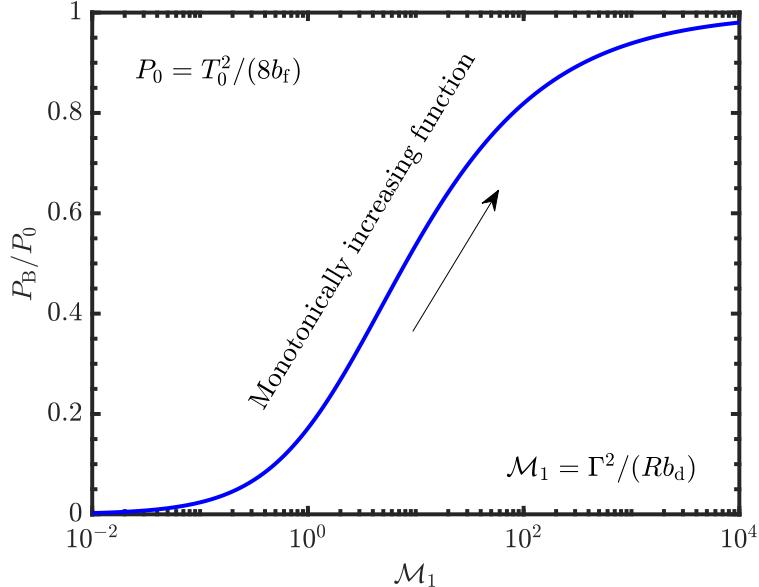


Figure 3.5: Normalized upper bound on power P_B/P_0 as a function of the effective figure of merit \mathcal{M}_1 for power take-off of a wave energy converter, as determined by equation (3.7). Given P_0 , P_B always increases with increasing \mathcal{M}_1 .

both dependent on geometry and frequency, the ultimate goal in designing a flap structure is to optimize its shape and dimensions to achieve the highest P_0 for a given wave frequency, rather than simply minimizing b_f or maximizing T_0 alone. There are several methods to enhance the effective figure of merit \mathcal{M}_1 for power take-off, such as decreasing the electrically parasitic loss R , lowering the mechanical loss b_d , and increasing the electrodynamic transduction factor Γ . Minimizing the power loss in a belt transmission system, characterized by the lumped mechanical damping constant b_d , has been extensively studied for decades (Shim and Kim, 2009; Silva et al., 2018; Zhu et al., 2021). Strategies to reduce b_d include proper tensioning and alignment as well as employing high-quality materials and low-friction coatings. For an electromagnetic generator, in the absence of size constraints, the number of turns of the pickup coil is proportional to both the resistance R and the electrodynamic transduction factor Γ (Wheeler, 1928; Cheng et al., 2007). However, $\mathcal{M}_1 \propto \Gamma^2$ and inversely proportional to R . Therefore, increasing the number of turns could significantly increase \mathcal{M}_1 . Since the ratio $\mathcal{M}_1/\mathcal{M}_2 = b_f/b_d$ is independent of Γ and R , a larger \mathcal{M}_1 achieved through this approach can also be beneficial for reducing the optimum N when necessary. Under space constraints, increasing the number of turns while keeping the same wire diameter is not feasible. Alternative methods to increase Γ include using stronger permanent magnets or incorporating magnetic flux concentrators.

While primarily focusing on the power available from the source, the Thévenin-equivalence technique can also provide insights into the properties of the load. The optimal complex load can be obtained by satisfying the conjugate impedance matching condition, where the load impedance

equals the complex conjugate of the Thévenin source impedance, $Z'_L = Z_{\text{th}}^*$. This is motivated by the fact that the optimum complex load impedance can be implemented either through an active pulse width modulation (PWM) rectifier or a passive rectifier coupled with a buck-boost converter (Bowden et al., 2015; Badel and Lefevre, 2016; Tse and Chung, 2020; Costanzo et al., 2021). As a consequence of Equation (3.5), the imaginary parts of Thévenin equivalent impedance Z_{th} are canceled out, therefore, $\Im\{Z'_L\} = 0$. In order to eliminate the imaginary impedance introduced by the inductance in the effective load Z'_L , a capacitance can be connected at the output port, with $X_L = 1/(-\omega C) = -\omega L$, leading to

$$C = 1/(\omega^2 L). \quad (3.9)$$

The load resistance is then adjusted to match the real part of Z_{th} , which is now identical to Z_{th} . This matching results in

$$R_L = R\sqrt{\mathcal{M}_1 + 1}. \quad (3.10)$$

Once (3.5), (3.6), (3.9), and (3.10) are satisfied, the power induced in the load R_L reaches its physical bound P_B . This section focuses on the specific aim of finding the global optimum of a wave energy converter operating in regular waves of a single frequency. However, the conditions required for this optimal performance may not be achievable in practice due to operational or engineering constraints. How system performance behaves away from the optimum under various scenarios remains an open question, which is to be explored in subsequent sections.

3.4 Power optimization for resonant wave energy converter

We investigate various techniques for optimizing the power delivered to a load, aiming to comprehensively understand the optimal conditions of the system parameters required to achieve the maximum possible output power. Our focus in this section shifts to the load side, with the complex load impedance Z_L considered as the primary variable in the optimization scheme. However, the potential contributions of other parameters, such as the mechanical transduction factor N , to the maximization of harvested power are also examined. We concentrate on the special case of resonance, as it has been shown that the maximum output power of an energy harvester is attained when the resonance frequency of the system closely matches or is identical to the frequency of the driving excitation (Mitcheson et al., 2008; Truong, 2022). The performance of a more general system, whose resonance frequency may be far from the frequency of the waves, is evaluated in section 3.5.

A general framework

Our first aim is to provide a framework to analyze the performance of an ocean wave energy converter in a general scenario. The lumped element circuit model depicted in Figure 3.2 is utilized for a clearer physical interpretation of the actual dynamic quantities, such as velocities of the mechanical parts and voltage across the load, and how they are derived from the model. The complex velocity amplitudes of the flap and the generator shaft are given by

$$\Theta = \frac{T_0}{b_f + Z_{\text{in}}}, \Psi = N\Theta \quad (3.11)$$

where the input impedance is

$$Z_{\text{in}} = j\omega I_f + \frac{K_f}{j\omega} + N^2 \left(b_d + j\omega I_d + \frac{K_d}{j\omega} + \frac{\Gamma^2}{R + j\omega L + jX_L + R_L} \right). \quad (3.12)$$

In this formulation, the excitation torque T and the damping coefficient b_f can be considered to mathematically constitute an effective source of power, in which the largest possible power it can provide the rest of the system is P_0 . Z_{in} then represents the impedance seen from this power source. The voltage across the load resistance R_L and the corresponding output power are

$$V_L = \frac{\Psi\Gamma}{R + j\omega L + jX_L + R_L} R_L, \quad (3.13)$$

$$P_L = \frac{1}{2} \frac{|V_L|^2}{R_L}. \quad (3.14)$$

The set of equations [(3.11)–(3.14)] offers a general procedure applicable to any set of physical parameters for studying the system performance under various conditions. In the next two subsections, we explore simple cases operating at resonance, where resonators are formed in both the mechanical and electrical domains to suppress the imaginary parts of impedance.

Optimal load and mechanical transmission ratio at resonance using gradient descent method

In theory, a resonance structure can be achieved by designing the moment of inertia and mechanical spring stiffness in the mechanical domain and connecting an external capacitance in the electrical output port of the electromagnetic generator such that the following conditions are satisfied,

$$K_f = \omega^2 I_f, K_d = \omega^2 I_d, X_L + \omega L = 0. \quad (3.15)$$

At resonance, the output power in (3.14) is simplified to

$$P_L = \frac{1}{2} \frac{R_L}{(R + R_L)^2} \frac{(T_0 N \Gamma)^2}{\left[(b_f + N^2 b_d) + \frac{N^2 \Gamma^2}{R + R_L} \right]^2}. \quad (3.16)$$

The stationary points of the power with respect to R_L and N are determined by $\partial P_L / \partial R_L = 0$ and $\partial P_L / \partial N = 0$, which are equivalent to

$$(\Gamma^2 + (R - R_L)b_d)N^2 + b_f(R - R_L) = 0, \quad (3.17)$$

$$(\Gamma^2 + (R + R_L)b_d)N^2 - b_f(R + R_L) = 0. \quad (3.18)$$

In analogy to \mathcal{M}_1 and \mathcal{M}_2 , we define the following dimensionless parameters

$$\mathcal{M}_3 = \frac{\Gamma^2}{R(N^2 b_f + b_d)} = \frac{\mathcal{M}_2}{N^2 + \mathcal{M}_2/\mathcal{M}_1},$$

$$\mathcal{M}_4 = \frac{\Gamma^2}{R(b_f + N^2 b_d)} = \frac{\mathcal{M}_1}{N^2 + \mathcal{M}_1/\mathcal{M}_2}.$$

Equation (3.17) results in the optimal load as a function of the mechanical transmission factor

$$R_L = R + \frac{N^2 \Gamma^2}{b_f + N^2 b_d} = R(1 + N^2 \mathcal{M}_4), \quad (3.19)$$

and the output power becomes

$$P_L = P_0 \frac{N^2 \mathcal{M}_4}{1 + N^2 \mathcal{M}_2(1 + 1/\mathcal{M}_1)}. \quad (3.20)$$

Simultaneously solving equations (3.17) and (3.18) yields identical solutions for the optimal mechanical transmission factor and load resistance as in (3.6) and (3.10), respectively. Substituting the optimal solutions back into (3.16) leads to an output power equal to the physical bound P_B in (3.7).

This analysis shows that, regardless of the initial pathway chosen, whether starting with the power available from the Thévenin equivalent source in (3.3) or the output power expression in (3.16), the global optimal solutions for the load and mechanical transduction factor at resonance, as well as the upper bound on generated power, remain the same. In addition, it seems to indicate the effectiveness of the impedance matching method in finding the optimal load compared to the gradient descent technique. The former, as presented at the end of section 3.3, may be more straightforward in some circumstances, especially for complex architectures with multiple stages of energy conversion and numerous system parameters. However, it is less clear whether these arguments always hold. Therefore, both approaches are employed as tools for further studies in this paper, and the results are evaluated in detail.

An important observation here is that the first two equalities in (3.15) form a special case of the condition (3.5). It can be seen as one specific solution to (3.5) that is independent of the mechanical transmission factor N . Any physical set of parameters $\{K_f, I_f, K_d, I_d\}$ that does not satisfy (3.15), but fulfills (3.5) when combined with N , will also suppress the imaginary parts of the input impedance Z_{in} , and all the results in this section are recovered. Therefore, the local resonance frequencies of the flap and drivetrain are not necessarily equal to the wave frequency to maximize the harvested power. For example, in an extreme scenario when the mechanical stiffness of the drivetrain K_d is set equal to zero, the moment of inertia I_d can be designed such that $I_d = (K_f/\omega^2 - I_f)/N^2$, thereby (3.5) is achieved. In this case, the local resonance frequency of the flap should not match the wave frequency as it leads to $I_d = 0$, an unphysical value.

It should be noted that conditions (3.5) and (3.15) are associated with (3.6), where the optimum of N is expressed as a function of dimensionless quantities \mathcal{M}_1 and \mathcal{M}_2 . An independent optimization procedure utilizing the bi-conjugate impedance matching method leads to a scenario where the equality $P_L = P_B$ is achieved at $N = 1$, but it requires that $\mathcal{M}_1 = (b_f/b_d)^2 - 1$. The detailed analysis is presented in Appendix 3.7. Applying the same logic to equation (3.6), if N is chosen beforehand, the effective figure of merit for power take-off must be $\mathcal{M}_1 = (b_f/(N^2 b_d))^2 - 1$, so that P_L is equal to P_B . However, this expression is valid if and only if $b_f > N^2 b_d$, which is not always guaranteed. Tuning the mechanical damping coefficients to make \mathcal{M}_1 physical and to maximize this figure of merit is challenging. Therefore, choosing to optimize the mechanical transmission factor N seems to be the more common approach.

Since the impedance matching method is widely used in the literature, one might question the possible outcomes if the optimal load resistance R_L is determined by this theorem, while the influence of the mechanical transmission factor N on the harvested power is explored over the entire physical range of $N \in (0, +\infty)$ through the gradient descent technique. Additionally, it is unclear whether N must be strictly computed by (3.6) or if other possibilities exist that could yield an output power identical to or close to the power limit P_B . These concerns are addressed in the following investigations.

Impedance matching to the load and optimal mechanical transmission ratio at resonance

Our starting point is the conjugate impedance matching of a load for maximizing the power delivered to it from a generator. The condition of this classical circuit theory concept is

$$\begin{aligned} R_L &= \Re\{Z_{\text{out}}\}, \\ X_L &= -\Im\{Z_{\text{out}}\} \end{aligned} \quad (3.21)$$

where the general form of the output impedance seen from the load is given by

$$Z_{\text{out}} = R + j\omega L + \frac{\Gamma^2}{b_d + j\omega I_d + \frac{K_d}{j\omega} + N^2 \left(b_f + j\omega I_f + \frac{K_f}{j\omega} \right)}.$$

These analyses are performed under the assumption that conditions (3.5) and (3.9) are satisfied. The imaginary part of the fractional term in the output impedance is suppressed. jX_L becomes the impedance of a capacitance in series with R_L to cancel out with the impedance of the coil inductance. The optimal load resistance and the corresponding power are thus derived as

$$R_L = R(1 + \mathcal{M}_3), \quad (3.22)$$

$$P_L = P_0 \frac{4(\mathcal{M}_3 + 1)}{\mathcal{M}_2 \left[(\mathcal{M}_3 + 2)/(N\mathcal{M}_4) + N \right]^2}. \quad (3.23)$$

We note that, the load resistance in (3.22) is a function of the mechanical transmission factor N .

The optimal solutions to N , if they exists, is obtained by differentiating P_L in (3.23) with respect to variable N and setting equal to zero, resulting in

$$X^4 + C_3 X^3 + C_2 X^2 + C_1 X + C_0 = 0 \quad (3.24)$$

where $X = N^2 > 0$, and the polynomial coefficients are

$$\begin{aligned} C_3 &= 2 \left[\left(\frac{\mathcal{M}_2}{\mathcal{M}_1} - \frac{\mathcal{M}_1}{\mathcal{M}_2} \right) \frac{1}{\mathcal{M}_1 + 2} + \mathcal{M}_2 \left(\frac{1}{\mathcal{M}_1} + 1 \right) \right], \\ C_2 &= 3 \left[(\mathcal{M}_2/\mathcal{M}_1)^2 + \mathcal{M}_2^2/\mathcal{M}_1 - 1 \right], \\ C_1 &= 2 \frac{\mathcal{M}_2}{\mathcal{M}_1} \left\{ \left[\left(\frac{\mathcal{M}_2}{\mathcal{M}_1} \right)^2 - 1 \right] \frac{(\mathcal{M}_1 + 1)^2}{\mathcal{M}_1 + 2} - 1 \right\}, \\ C_0 &= - \left(\mathcal{M}_2/\mathcal{M}_1 \right)^2 (1 + \mathcal{M}_1). \end{aligned}$$

Although the quartic equation (3.24) can be solved analytically, these explicit closed-form solutions are generally cumbersome and difficult to interpret. Therefore, a numerical method, such as the *vpasolve* function in MATLAB, is more appropriate in this case.

Figure 3.6 shows the maximum output power under the resonance conditions (either (3.5) or (3.15)), computed utilizing the gradient descent approach and impedance matching theorem, and expressed as a function of the mechanical transmission factor. The example set of system parameters in use $\{b_f, b_d, \Gamma, R\}$ is listed in Table 3.1. It can be seen that the former method results in an output power that is always equal to or higher than that of the latter for any value of N . Equality occurs only at $N = 1$, as discussed further in section 3.6. Hence, the impedance matching technique can

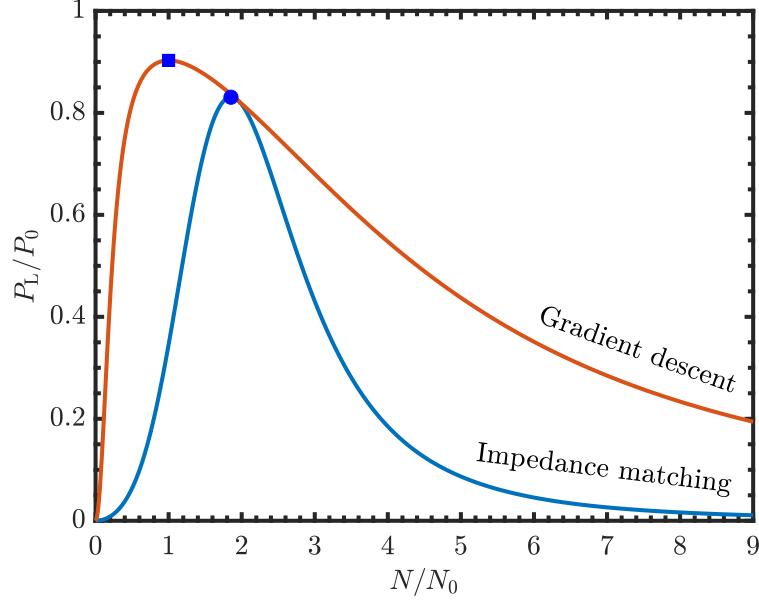


Figure 3.6: Normalized maximum output power P_L/P_0 at resonance as a function of normalized mechanical transmission factor N/N_0 , obtained from equations (3.20) and (3.23) using gradient descent and impedance matching methods. The optimal load resistance is determined by (3.19) and (3.22), and the optimal mechanical transmission factor N_0 is the solution to (3.6). Blue solid circle and square represent the power maxima corresponding to the optimal values of N given by (3.6) and (3.24), respectively.

be considered suboptimal compared to the gradient descent method. In theory, the number of positive solutions of equation (3.24), a polynomial equation of the fourth degree, depends on the relationship between \mathcal{M}_1 and \mathcal{M}_2 . This indicates that there may exist other or multiple solutions for N and $R_L(N)$ such that the output power can closely approach the physical power bound P_B . However, in the specific example presented, a unique solution is observed.

The results derived in this section offer solutions for designing an optimal wave energy converter without considering practical constraints. However, solving equation (3.5) for I_d and K_d may yield physically unrealistic values, depending on the mechanical properties of the flap and the frequency of the regular waves. Thus, some authors may be more interested in the performance of a system whose characteristics do not fulfill this condition, and how it compares to the global optimum. This aspect is addressed in the next section.

3.5 General power optimization and system effectiveness

Wave energy converter effectiveness

We now consider more general scenarios in which resonators in the mechanical domain may or may not be formed, and only the load can be arbitrarily adjusted by power electronic interface circuits. To assess the performance of a wave energy converter, we introduce the system effectiveness, which can be defined as the ratio between the harvester output power and the largest achievable power of the optimal system,

$$E_H = \frac{P_L}{P_B}. \quad (3.25)$$

This metric indicates how well a wave energy converter performs compared to an ideal configuration. The output power P_L is computed by the process [(3.11)–(3.14)], while P_B is given by (3.7).

In this section, we further evaluate the gradient descent and impedance matching methods to determine which is preferable for optimizing the performance of the general complex system under consideration.

Gradient descent approach

In order to reduce the complexity of derivation, we introduce the notations

$$\begin{aligned} b_e &= b_f + N^2 b_d, \\ X_e &= X_f + N^2 X_d = (\omega I_f - K_f/\omega) + N^2 (\omega I_d - K_d/\omega), \\ \Gamma_e &= N\Gamma. \end{aligned}$$

The stationary points of the output power are obtained by setting its derivatives in relation to the real and imaginary parts of the load equal to zero, $\partial P_L / \partial X_L = 0$ and $\partial P_L / \partial R_L = 0$, leading to the equations

$$\begin{aligned} (X + X_L)(X_e^2 + b_e^2) - \Gamma_e^2 X_e &= 0, \\ (R^2 - R_L^2)(X_e^2 + b_e^2) + \left[(X + X_L)b_e \right]^2 + 2Rb_e\Gamma_e^2 + \left[\Gamma_e^2 - (X + X_L)X_e \right]^2 &= 0, \end{aligned}$$

respectively. Hence the optimal solution for the load is

$$\begin{aligned} X_L &= -X + \frac{\Gamma_e^2 X_e}{X_e^2 + b_e^2}, \\ R_L &= R + \frac{\Gamma_e^2 b_e}{X_e^2 + b_e^2}. \end{aligned} \tag{3.26}$$

Substituting these expressions back into (3.14), we arrive at

$$P_L = \frac{\Gamma_e^2 T_0^2}{8 \left[\Gamma_e^2 b_e + R(X_e^2 + b_e^2) \right]}. \tag{3.27}$$

Considering X_e a variable, the derivative of the output power in (3.27) in terms of X_e is

$$\frac{\partial P_L}{\partial X_e} = -\frac{\Gamma_e^2 T_0^2 R X_e}{4 \left[\Gamma_e^2 b_e + R(X_e^2 + b_e^2) \right]^2}.$$

We have that $\partial P_L / \partial X_e = 0$ if and only if $X_e = 0$, $\partial P_L / \partial X_e > 0$ for $X_e < 0$, and $\partial P_L / \partial X_e < 0$ for $X_e > 0$. In the limits, $P_L \rightarrow 0$ when $X_e \rightarrow \pm\infty$. The output power P_L monotonically increases and decreases with negative and positive X_e , respectively, and attains a unique maximum when X_e is zero. Therefore, if it is not feasible to optimize the wave energy converter such that X_e is canceled out, minimizing the quantity $|X_e|^2$ is always desirable to maximize the output power.

We note that the parameters b_e , X_e , and Γ_e are dependent on the mechanical transmission factor N . The optimal value of N can be found by setting $\partial P_L / \partial N = 0$, where the power P_L is given by (3.27), resulting in

$$N^4 = \frac{R(X_f^2 + b_f^2)}{\Gamma^2 b_d + R(X_d^2 + b_d^2)} = \frac{(X_f/b_d)^2 + (\mathcal{M}_2/\mathcal{M}_1)^2}{(X_d/b_d)^2 + \mathcal{M}_1 + 1}. \tag{3.28}$$

Inserting the mechanical transmission factor from (3.28) into (3.27), we obtain the maximum output power with the simultaneous optimization of the variables X_L, R_L , and N

$$P_L = P_0 \left\{ \mathcal{M}_1 - 2 \left[\sqrt{\left(\left(\frac{X_d}{b_d} \right)^2 + 1 \right) \left(\left(\frac{X_f}{b_f} \right)^2 + \mathcal{M}_1 + 1 \right)} - \frac{X_d}{b_d} \frac{X_f}{b_f} - 1 \right] \right\} \\ / \left[\mathcal{M}_1 + 4 \frac{X_d}{b_d} \frac{X_f}{b_f} - 4 \left(\frac{X_f}{b_f} \right)^2 - \frac{4}{\mathcal{M}_1} \left(\frac{X_d}{b_d} - \frac{X_f}{b_f} \right)^2 \right]. \quad (3.29)$$

Direct consequences of these investigations are the findings presented in section 3.4, which focus on resonance when $X_d = X_f = 0$. In a more general case, if $X_e = 0$, following the same gradient descent optimization procedure yields identical results for the optimal N, X_L, R_L , and output power as shown in (3.6), (3.9), (3.10), and (3.7), respectively.

Assuming that in a realistic system, the condition $X_e = 0$ may not be achievable. The minimization problem of $|X_e|^2$ with the optimal N in (3.28) can be formulated as follows,

$$\min \left\{ |X_e|^2 = \left[X_f + X_d \sqrt{\frac{(X_f/b_d)^2 + (\mathcal{M}_2/\mathcal{M}_1)^2}{(X_d/b_d)^2 + \mathcal{M}_1 + 1}} \right]^2 \right\}.$$

Since X_f results from optimizing the flap geometry such that the power P_0 is maximized, and mechanical loss factors such as b_d and b_f are always desired to be as minimal as possible, the common choice of variable to minimize $|X_e|^2$ is X_d . Without any constraints, X_d can be determined by the equation $d|X_e|^2/dX_d = 0$, leading to the solution

$$X_d = -X_f \frac{\mathcal{M}_2}{\mathcal{M}_1} \sqrt{\mathcal{M}_1 + 1}. \quad (3.30)$$

Denoting the right-hand side of (3.30) as X_d^* , we also find that $d|X_e|^2/dX_d > 0$ for all $X_d > X_d^*$ and $d|X_e|^2/dX_d < 0$ for all $X_d < X_d^*$. In other words, $|X_e|^2$ monotonically decreases for X_d on the left side of the minimum point X_d^* and monotonically increases for X_d on the right side. Therefore, if X_d is subject to a constraint of $X_d \in [X_{d,\ell}, X_{d,u}]$, where $X_{d,\ell}$ and $X_{d,u}$ are the lower and upper bounds to X_d ($X_{d,\ell} < X_{d,u}$), then the optimal X_d is chosen depending on the relationships among these bounds and X_d^* . In particular, $X_d = X_{d,u}$ if $X_{d,u} < X_d^*$, $X_d = X_d^*$ if $X_d^* \in [X_{d,\ell}, X_{d,u}]$, and $X_d = X_{d,\ell}$ if $X_d^* < X_{d,\ell}$. Note that the practical bounds on X_d depend on the wave frequency ω and constraints on the moment of inertia and spring constant of the drivetrain, I_d and K_d . This also underscores an important role of K_d in maximizing the system performance, and that assuming or setting $K_d = 0$ in a numerical optimization scheme is not justified.

Impedance matching method

Applying the condition of impedance matching to the load in (3.21), its real and imaginary parts are derived as

$$R_L = R \left(1 + \frac{\mathcal{M}_3 \mathcal{M}_f^2}{\mathcal{M}_3^2 + \mathcal{M}_f^2} \right), \quad (3.31)$$

$$X_L = -X + \frac{\Gamma^2 (N^2 X_f + X_d)}{(N^2 X_f + X_d)^2 + (N^2 b_f + b_d)^2} \quad (3.32)$$

where $X = \omega L, X_f = \omega I_f - K_f/\omega, X_d = \omega I_d - K_d/\omega$, and the dimensionless parameters are defined as

$$\mathcal{M}_f = \frac{\Gamma^2}{R(N^2 X_f + X_d)}, \quad \mathcal{M}_d = \frac{\Gamma^2}{R(X_f + N^2 X_d)}.$$

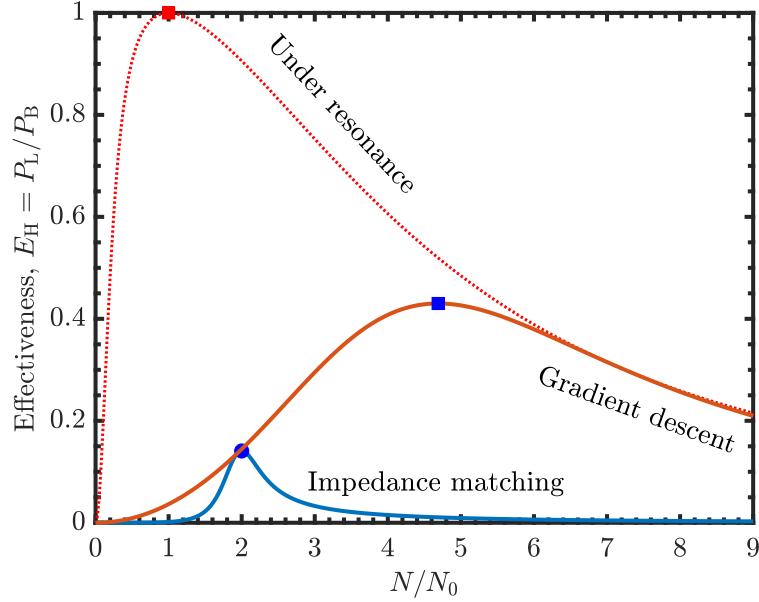


Figure 3.7: Maximum effectiveness of the wave energy converter as a function of the normalized mechanical transmission factor N/N_0 , computed from equations (3.27) and (3.33) utilizing gradient descent and impedance matching principles. The optimal complex loads are expressed by (3.26) and (3.31), and the optimal mechanical transmission factor at resonance N_0 is specified by (3.6). The blue solid circle and square represent the highest effectiveness at the optimal values of N , determined analytically by Equation (3.28) and numerically by the Genetic Algorithm based on the power expression (3.33), respectively.

The output power can then be written as

$$P_L = P_0 \frac{4N^2(\lambda - 1)(\lambda^2 + \gamma^2)/\mathcal{M}_2}{\left[(\lambda^2 + \gamma^2)/\mathcal{M}_4 + \lambda N^2\right]^2 + \left[(\lambda^2 + \gamma^2)/\mathcal{M}_d - \gamma N^2\right]^2} \quad (3.33)$$

where

$$\lambda = 2 + \frac{\mathcal{M}_3 \mathcal{M}_f^2}{\mathcal{M}_3^2 + \mathcal{M}_f^2}, \quad \gamma = \frac{\mathcal{M}_3^2 \mathcal{M}_f}{\mathcal{M}_3^2 + \mathcal{M}_f^2}.$$

Taking the derivative of the power in (3.33) with respect to N leads to a highly complex equation that is very difficult to solve exactly. However, the optimal mechanical transmission factor N can be determined numerically using the Genetic Algorithm, a global optimization toolbox in MATLAB, under the inequality constraint $N \in (0, +\infty)$.

Figure 3.7 illustrates an example of the maximum effectiveness of an oscillating surge wave energy converter, based on the complete set of system parameters detailed in Table 3.1. The optimal mechanical transmission factor at off-resonance is considerably higher than at resonance. Since neither condition (3.5) nor (3.15) is satisfied, the system effectiveness is well below the theoretical optimum (represented by the dashed line in the Figure). In this case, the difference between the results derived from the gradient descent and impedance matching approaches becomes even more pronounced compared to the discrepancy shown in Figure 3.6. Furthermore, it is essential to emphasize that as the mechanical transmission factor increases, the effectiveness of the considered system gradually converges with that of the optimal case. This property explains why the numerical optimization presented in (Grasberger et al., 2024) tends to yield a high value of N as the variation in output power between the true-optimal design and a non-optimal solution in this region is nearly

| Parameters | Value |
|--|-----------------------|
| Flap structure (Ahmed et al., 2024b) | |
| Effective mechanical damping coefficient, b_f | 10.34, Nms |
| Overall moment of inertia, I_f | 59.10, kgm^2 |
| Mechanical stiffness, K_f | 290, Nm/rad |
| Mechanical transmission system (drivetrain) (Grasberger et al., 2024) | |
| Effective mechanical damping coefficient, b_d | 2, Nms |
| Moment of inertia, I_d | 20, kgm^2 |
| Mechanical stiffness, K_d | 0, Nm/rad |
| Electromagnetic generator (Sergienko et al., 2021; Grasberger et al., 2024) | |
| Coil inductance, L | 1.4, mH |
| Parasitic resistance, R | 38, $\text{m}\Omega$ |
| Electromagnetic transduction factor, Γ | 5.397, Nm/A |
| Wave frequency, f | 0.15, Hz |

Table 3.1: Example of system parameters

negligible. This finding underscores the importance of deriving analytical solutions for the optimal design of an oscillating surge wave energy converter, as they can offer insights that numerical algorithms typically cannot provide.

Up to this point, questions regarding how to determine the optimal set of system parameters (such as the mechanical transmission factor, the moment of inertia and spring stiffness of the drivetrain, and the complex load) and which method is most appropriate have been thoroughly investigated for both ideal and general cases. The modeling techniques and optimization procedures presented in this paper can also be generalized and adapted for other similar structures, including buoy point absorber wave energy converter and ocean current energy harvester.

3.6 Discussions

Although the impedance matching theorem is convenient as it provides direct relationships between the real and imaginary parts of the optimal load and the output impedance of the system, it does not always yield the same results as the gradient descent method. Furthermore, two-side impedance matching is not guaranteed to always have a physical solution. In some particular scenarios, identical outcomes can be obtained regardless of the approach used (Truong et al., 2019). For instance, the investigations in sections 3.3, 3.4, and Appendix 3.7 yield the same global maximum point of the output power. However, while the corresponding optimal conditions are identical for the first two analyses, the last one is only a specific case of them. In other circumstances, the impedance matching technique can be considered sub-optimal compared to the gradient descent approach, such as the results in sections 3.4 and 3.5 relative to those in sections 3.4 and 3.5,

respectively. We prove that the output power achieved with the optimal load determined by the impedance matching theorem is always less than or equal to that with the optimal load derived from the gradient descent method. Therefore, in general, the gradient descent algorithm is preferable to determine the optimal parameters for the wave energy converter. Detailed proofs are provided in Appendix 3.7.

Equation (3.5) (or equivalently, $X_e = 0$) can be solved with at least two independent variables, I_d and K_d . Thus, there are infinite possibilities to optimize the performance of a wave energy converter such that the output power can achieve its physical bound, from both physics and mathematics points of view. However, due to potential practical limitations, the final parameter selections are made based on their specific constraints, which may lead to a maximum output power less than its upper bound. While the analysis presented in this work has accounted for the dynamics of the entire system as a unified unit, the analytical findings show that each subsystem can be optimized separately using its corresponding optimal criteria. This could reduce computational complexity compared to numerically optimizing the system as a multiple-variable problem. Moreover, achieving global optimality with numerical solutions in a consistent and reliable manner is challenging.

A general procedure to optimize the wave energy converter can be briefly summarized as follows. The geometry of the flap plays a vital role in capturing the energy carried by ocean waves. It should be designed such that the effective damping coefficient b_f is minimized and the external torque amplitude T_0 acting on the flap is maximized, thereby increasing the power P_0 . The parameters I_f and K_f are not considered as variables but rather consequences of this optimization scheme. The electrical components of the electromagnetic transducer are adjusted to achieve a sufficiently large effective figure of merit \mathcal{M}_1 so that P_B closely approaches P_0 . The moment of inertia I_d and spring stiffness K_d of the drivetrain are then optimized accordingly to minimize the quantity $|X_e|^2$. Finally, a power electronic circuit and control unit is employed to synthesize the optimal effective load impedance to maximize the output power.

In the ideal case where $X_e = 0$, i.e., the mechanical characteristics of the drivetrain I_d and K_d are configured to satisfy (3.30), the optimal transmission factor N and the imaginary and real parts of the load Z_L are given by (3.6), (3.9), and (3.10), respectively. The output power in this case reaches its upper limit P_B , as defined in (3.7). When subject to practical constraints that cause X_e to differ from zero, the optimal X_d is at one of its constraints to minimize $|X_e|^2$, and the pair $\{I_d, K_d\}$ is selected accordingly. The optimum of N is calculated using (3.28), and the optimal complex load is expressed by (3.26). The maximum achievable power in this scenario is described by (3.29), which is always lower than P_B . The performance of the system can be evaluated by the effectiveness (3.25). The analytical solutions presented in this paper partially provide general design rules for an optimal wave energy converter.

3.7 Conclusions

We have investigated the performance of an oscillating surge wave energy converter under resonance and for a general scenario. We have determined the physical power bound and identified the sufficient conditions to achieve this maximum power. We have introduced an effective figure of merit for power take-off of a wave energy converter, a comprehensive metric for evaluating the overall performance and effectiveness of the system. A rigorous analysis shows that the gradient descent approach, in comparison with the impedance matching method, is preferable for finding the global maximum output power and determining the desired real and imaginary characteristics of the electrical load. This observation could change how results established in the literature by impedance matching theory are interpreted and provide significant insights into selecting the

proper procedure to optimize a complex energy harvesting system. Furthermore, optimal values of other parameters, such as the mechanical transmission factor and generator mechanical properties, are explored for both unconstrained and constrained regimes, with analytical solutions derived accordingly. The findings offer valuable principles for designing an efficient linear oscillating surge wave energy converter under harmonic excitation of regular waves, and the presented framework can be extended to optimize various structures, including point absorber wave energy converters.

Appendix A: Power transmitted through a lossless network

We consider a harmonic voltage source $V_S = V_0 \cos(\omega t)$ with its internal impedance $Z_S = R_S + jX_S$, connected with a lossless network represented by an imaginary impedance jY . This configuration is shown in Figure 3.8. We now prove that the maximum power transmitted through this network is identical to the power available from the source. The voltage and impedance of the Thévenin equivalent circuit in the frequency domain are given by

$$V_{\text{th},S} = \frac{V_0}{R_S + j(X_S + Y)}, \quad (3.34)$$

$$Z_{\text{th},S} = \frac{jY(R_S + jX_S)}{R_S + j(X_S + Y)}. \quad (3.35)$$

The maximum power transmitted from port 1 to port 2 is the power available from the Thévenin source, which is derived as

$$P_{\text{avs,th}} = \frac{1}{8} \frac{|V_{\text{th},S}|^2}{\Re\{Z_{\text{th},S}\}} = \frac{1}{8} \frac{V_0^2}{R_S} = \frac{1}{8} \frac{|V_S|^2}{\Re\{Z_S\}}, \quad (3.36)$$

which is the power available from the voltage source V_S , or the highest power that the source can deliver to the rest of the circuit.

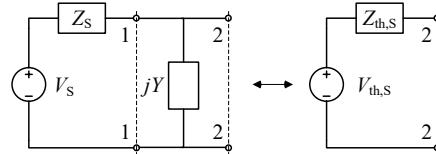


Figure 3.8: Thévenin equivalent circuit representation for the lossless network under consideration.

Appendix B: Bi-conjugate impedance matching at resonance

A classical method to maximize the power delivered to the load of a complex system is the bi-conjugate matching theorem. In particular, the input impedance is matched with the conjugate impedance of the effective power source, and the output impedance is equal to the conjugate impedance of the complex load,

$$Z_{\text{in}} = Z_S^*, \quad (3.37)$$

$$Z_{\text{out}} = Z_L^*. \quad (3.38)$$

The effective power source is specified by $\{T, b_f\}$, where the source impedance is purely real, $Z_S = b_f$. Under resonance, i.e., when the conditions in (3.15) are fulfilled, the required impedance matching

is simplified as

$$b_f = N^2 \left(b_d + \frac{\Gamma^2}{R + R_L} \right), \quad (3.39)$$

$$R_L = R + \frac{\Gamma^2}{N^2 b_f + b_d}. \quad (3.40)$$

The physical (positive) solution to R_L and N are

$$R_L = R \sqrt{\frac{\mathcal{M}_1(\mathcal{M}_1+2)\mathcal{M}_2^2}{\mathcal{M}_1^2 + \mathcal{M}_2^2} + 1}, \quad (3.41)$$

$$N^2 = \sqrt{(\alpha/\beta)^2 + 1} + \alpha/\beta \quad (3.42)$$

where

$$\alpha = \mathcal{M}_1/\mathcal{M}_2 - \mathcal{M}_2/\mathcal{M}_1 - \mathcal{M}_2, \quad (3.43)$$

$$\beta = \mathcal{M}_1 + 2. \quad (3.44)$$

Substituting (3.41) and (3.42) into (3.16) yields the output power

$$P_L = P_0 \frac{\mathcal{M}_1 r_L}{(r_L + 1)(\mathcal{M}_1 + r_L + 1)} \quad (3.45)$$

where r_L denotes the optimal load ratio defined as $r_L = R_L/R$ from (3.41), which is a function of both \mathcal{M}_1 and \mathcal{M}_2 . We now look for the optimal values of \mathcal{M}_1 and \mathcal{M}_2 that make the power in (3.45) stationary. Differentiating and setting equal to zero, we get

$$\mathcal{M}_2 = \mathcal{M}_1 / \sqrt{\mathcal{M}_1 + 1}, \quad (3.46)$$

while $\partial P_L / \partial \mathcal{M}_1 > 0$ for all $\{\mathcal{M}_1, \mathcal{M}_2\} > 0$. Inserting (3.46) back into the expressions of the optimal load, mechanical transmission factor, and output power, we obtain

$$R_L = R \sqrt{\mathcal{M}_1 + 1}, \quad (3.47)$$

$$N = 1, \quad (3.48)$$

$$P_L = P_0 \frac{\left(\sqrt{\mathcal{M}_1 + 1} - 1 \right)^2}{\mathcal{M}_1} = P_B. \quad (3.49)$$

Equations (3.46) and (3.48) can be viewed as a special case of a more general optimal condition expressed in (3.6). It is also shown that maximizing the output power through the process of bi-conjugate impedance matching leads to identical results in power and optimal load resistance in (3.7) and (3.10), respectively.

We can write equation (3.46) with \mathcal{M}_1 as a dependent variable as follows

$$\mathcal{M}_1 = (b_f/b_d)^2 - 1. \quad (3.50)$$

This solution is only valid for $b_f > b_d$, which is not always the case. Assuming that the condition $b_f > b_d$ is met and \mathcal{M}_1 is physical, the desired value of \mathcal{M}_1 can be attained in practice by selecting permanent magnets with appropriate magnetic strength or by deliberately designing magnetic flux loops such that $\Gamma^2/R = (b_f^2 - b_d^2)/b_d$. However, implementing \mathcal{M}_1 based on the mechanical damping coefficients b_f and b_d is not preferable, as expression (3.50) may not yield a large \mathcal{M}_1 , which in turn decreases P_B .

Appendix C: Comparison of output power with optimal load from gradient descent method and impedance matching theorem

We first consider the optimal power derived in section 3.4 for the ideal scenario. We show that the power expressed in (3.20) is always larger than or equal to that in (3.23). By subtracting the two equations from both sides correspondingly, the normalized power difference can be written as $\Delta P_r/P_0 = \mathcal{N}_r/\mathcal{D}_r$. All the terms in the denominator \mathcal{D}_r are positive, therefore, it is also positive. The detailed expression is quite complicated and is not shown here. The numerator is given by $\mathcal{N}_r = [\mathcal{M}_1^2 \mathcal{M}_2 (N^4 - 1)]^2$, which is non-negative for all N . As a consequence, ΔP_r is always non-negative, indicating that the power achieved by the gradient descent method is always larger than or equal to that by the impedance matching theorem. The two power values are equal only for $N = 1$, as can be seen in Figure 3.6.

For the more general case in section 3.5, we examine the power expressions in (3.27) and (3.33), which result from the gradient descent and impedance matching methods, respectively. Following the same procedure, we can write the power difference in the form of $\Delta P_g = \mathcal{N}_g/\mathcal{D}_g$. Similarly, we find that $\mathcal{D}_g > 0$ and $\mathcal{N}_g = \Gamma^6 N^2 T_0^2 (N^4 - 1)^2 (X_f^2 + b_f^2) \geq 0$ for all N . Equality occurs only for $N = 1$, as demonstrated in the example in Figure 3.7. These findings prove that the gradient descent approach is preferable for optimizing system parameters to maximize the power delivered to the load or utility grid.

References

- A. Ahmed, J. Mi, J. Huang, R. Datla, K. Connington, L. Zuo, and M. R. Hajj. Power capture and power take-off load of a self-balanced dual-flap oscillating surge wave energy converter. *Energy*, 291:130431, 2024a.
- A. Ahmed, L. Yang, J. Huang, A. Shalaby, R. Datla, L. Zuo, and M. Hajj. Performance characterization and modeling of an oscillating surge wave energy converter. *Nonlinear Dynamics*, 112(6):4007–4025, 2024b.
- M. Ancellin and F. Dias. Capytaine: a Python-based linear potential flow solver. *Journal of Open Source Software*, 4(36):1341, 2019.
- A. Badel and E. Lefeuvre. Nonlinear conditioning circuits for piezoelectric energy harvesters. In E. Blokhina, A. El Aroudi, E. Alarcon, and D. Galayko, editors, *Nonlinearity in Energy Harvesting Systems: Micro- and Nanoscale Applications*, pages 321–359. Springer International Publishing, 2016.
- S. Barbarelli, M. Amelio, T. Castiglione, G. Florio, and N. M. Scornaienchi. Design and analysis of a new wave energy converter based on a point absorber and a hydraulic system harvesting energy from waves near the shore in calm seas. *International Journal of Energy Research*, 45(1):661–690, 2021.
- A. Barua and M. Salauddin Rasel. Advances and challenges in ocean wave energy harvesting. *Sustainable Energy Technologies and Assessments*, 61:103599, 2024.
- J. A. Bowden, S. G. Burrow, A. Cammarano, L. R. Clare, and P. D. Mitcheson. Switched-mode load impedance synthesis to parametrically tune electromagnetic vibration energy harvesters. *IEEE/ASME Transactions on Mechatronics*, 20(2):603–610, 2015.
- S. Cheng, N. Wang, and D. P. Arnold. Modeling of magnetic vibrational energy harvesters using equivalent circuit representations. *Journal of Micromechanics and Microengineering*, 17(11):2328, 2007.
- R. G. Coe, G. Bacelli, S. Olson, V. S. Neary, and M. B. R. Topper. Initial conceptual demonstration of control co-design for WEC optimization. *Journal of Ocean Engineering and Marine Energy*, 6:441–449, 2020.
- L. Costanzo, T. Lin, W. Lin, A. L. Schiavo, M. Vitelli, and L. Zuo. Power electronic interface with an adaptive mppt technique for train suspension energy harvesters. *IEEE Transactions on Industrial Electronics*, 68(9):8219–8230, 2021.

References

D. Gallutia, M. T. Fard, M. G. Soto, and J. He. Recent advances in wave energy conversion systems: From wave theory to devices and control strategies. *Ocean Engineering*, 252:111105, 2022.

N. Ghasemipour, P. Izanlou, and M. H. Jahangir. Feasibility study on utilizing oscillating wave surge converters (OWSCs) in nearshore regions, case study: Along the southeastern coast of Iran in Oman sea. *Journal of Cleaner Production*, 367:133090, 2022.

G. Giannini, P. Rosa-Santos, V. Ramos, and F. Taveira-Pinto. Wave energy converters design combining hydrodynamic performance and structural assessment. *Energy*, 249:123641, 2022.

F. Giorcelli, S. A. Sirigu, E. Pasta, D. G. Gioia, M. Bonfanti, and G. Mattiazzo. *Wave Energy Converter Optimal Design Under Parameter Uncertainty*, page V008T09A085. 2022.

J. Grasberger, L. Yang, G. Bacelli, and L. Zuo. Control co-design and optimization of oscillating-surge wave energy converter. *Renewable Energy*, 225:120234, 2024.

K. Gunn and C. Stock-Williams. Quantifying the global wave power resource. *Renewable Energy*, 44:296–304, 2012.

B. Guo and J. V. Ringwood. A review of wave energy technology from a research and commercial perspective. *IET Renewable Power Generation*, 15(14):3065–3090, 2021.

J. Harms, M. Hollm, L. Dostal, T. A. Kern, and R. Seifried. Design and optimization of a wave energy converter for drifting sensor platforms in realistic ocean waves. *Applied Energy*, 321:119303, 2022.

M. Hollm, L. Dostal, D. Yurchenko, and R. Seifried. Performance increase of wave energy harvesting of a guided point absorber. *The European Physical Journal Special Topics*, 231(8):1465–1473, 2022.

D. Khojasteh, A. Shamsipour, L. Huang, S. Tavakoli, M. Haghani, F. Flocard, M. Farzadkhoo, G. Iglesias, M. Hemer, M. Lewis, S. Neill, M. M. Bernitsas, and W. Glamore. A large-scale review of wave and tidal energy research over the last 20 years. *Ocean Engineering*, 282:114995, 2023.

B. Kuo. *Linear Networks and Systems*. McGraw-Hill, 1967.

R. Kurnia and G. Ducrozet. Nemoh: Open-source boundary element solver for computation of first- and second-order hydrodynamic loads in the frequency domain. *Computer Physics Communications*, 292:108885, 2023.

Y. Li and M. Lin. Regular and irregular wave impacts on floating body. *Ocean Engineering*, 42:93–101, 2012.

Z. Liu, X. Zhang, and C. Xu. Hydrodynamic and energy-harvesting performance of a BBDB-OWC device in irregular waves: An experimental study. *Applied Energy*, 350:121737, 2023.

I. López, J. Andreu, S. Ceballos, I. Martínez de Alegría, and I. Kortabarria. Review of wave energy technologies and the necessary power-equipment. *Renewable and Sustainable Energy Reviews*, 27:413–434, 2013.

J. Mi, J. Huang, X. Li, A. Ahmed, L. Yang, U. Chung, R. Datla, M. Hajj, and L. Zuo. Oscillating surge wave energy converter using a novel above-water power takeoff with belt-arc speed amplification. *Ocean Engineering*, 310:118503, 2024.

P. D. Mitcheson, E. M. Yeatman, G. K. Rao, A. S. Holmes, and T. C. Green. Energy harvesting from human and machine motion for wireless electronic devices. *Proceedings of the IEEE*, 96(9):1457–1486, 2008.

B. Paduano, L. Parrinello, F. Niosi, O. Dell’Edera, S. A. Sirigu, N. Faedo, and G. Mattiazzo. Towards standardised design of wave energy converters: A high-fidelity modelling approach. *Renewable Energy*, 224:120141, 2024.

Y. Peña-Sánchez, D. García-Violini, and J. V. Ringwood. Control co-design of power take-off parameters for wave energy systems. *IFAC-PapersOnLine*, 55(27):311–316, 2022.

M. Penalba, G. Giorgi, and J. V. Ringwood. Mathematical modelling of wave energy converters: A review of nonlinear approaches. *Renewable and Sustainable Energy Reviews*, 78:1188–1207, 2017.

H. Peng, W. Qiu, W. Meng, et al. Experimental studies and time-domain simulation of a hinged-type wave energy converter in regular waves. *Marine Systems & Ocean Technology*, 15:1–15, 2020.

References

B. Reguero, I. Losada, and F. Méndez. A global wave power resource and its seasonal, interannual and long-term variability. *Applied Energy*, 148:366–380, 2015.

D. Rowell and D. Wormley. *System Dynamics: An Introduction*. Prentice Hall, 1997.

M. Ruellan, H. BenAhmed, B. Multon, C. Josset, A. Babarit, and A. Clement. Design methodology for a SEAREV wave energy converter. *IEEE Transactions on Energy Conversion*, 25(3):760–767, 2010.

N. P. Sell, A. R. Plummer, and A. J. Hillis. A self-zeroing position controller for oscillating surge wave energy converters with strong asymmetry. *Journal of Ocean Engineering and Marine Energy*, 4:137–151, 2018.

N. Y. Sergienko, L. S. P. da Silva, B. Ding, and B. S. Cazzolato. Importance of drivetrain optimisation to maximise electrical power from wave energy converters. *IET Renewable Power Generation*, 15(14):3232–3242, 2021.

H.-J. Shim and J.-K. Kim. Cause of failure and optimization of a V-belt pulley considering fatigue life uncertainty in automotive applications. *Engineering Failure Analysis*, 16(6):1955–1963, 2009.

C. A. Silva, M.-A. Andrianoely, L. Manin, S. Ayasamy, C. Santini, E. Besnier, and D. Remond. Optimization of power losses in poly-V belt transmissions via genetic algorithm and dynamic programming. *Mechanism and Machine Theory*, 128:169–190, 2018.

P. Stansby, E. Carpintero Moreno, T. Stallard, and A. Maggi. Three-float broad-band resonant line absorber with surge for wave energy conversion. *Renewable Energy*, 78:132–140, 2015.

C. A. M. Ströfer, D. T. Gaebele, R. G. Coe, and G. Bacelli. Control co-design of power take-off systems for wave energy converters using WecOptTool. *IEEE Transactions on Sustainable Energy*, 14(4):2157–2167, 2023.

H. A. C. Tilmans. Equivalent circuit representation of electromechanical transducers: I. Lumped-parameter systems. *Journal of Micromechanics and Microengineering*, 6:157–176, 1996.

A. Trueworthy and B. DuPont. The wave energy converter design process: Methods applied in industry and shortcomings of current practices. *Journal of Marine Science and Engineering*, 8(11):932, 2020.

B. D. Truong. Power optimization of a magnetoelectric wireless power transfer system with volume constraint. *Sensors and Actuators A: Physical*, 341:113226, 2022.

B. D. Truong, S. Williams, and S. Roundy. Experimentally validated model and analytical investigations on power optimization for piezoelectric-based wireless power transfer systems. *Journal of Intelligent Material Systems and Structures*, 30(16):2464–2477, 2019.

K. H. Tse and H. S. Chung. MPPT for electromagnetic energy harvesters having nonnegligible output reactance operating under slow-varying conditions. *IEEE Transactions on Power Electronics*, 35(7):7110–7122, 2020.

H. Wheeler. Simple inductance formulas for radio coils. *Proceedings of the Institute of Radio Engineers*, 16(10):1398–1400, 1928.

A. O. Winter and M. R. Motley. Development of a fluid-structure interaction model of an oscillating wave surge converter using OpenFOAM. *Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering (OMAE)*, 9:V009T09A034, 2020.

L. Yang, J. Huang, J. Mi, M. Hajj, G. Bacelli, and L. Zuo. Control-inspired design and power optimization of an active mechanical motion rectifier based power takeoff for wave energy converters. *Renewable Energy*, 228:120614, 2024.

L. Yang, J. Huang, S. J. Spencer, X. Li, J. Mi, G. Bacelli, M. Hajj, and L. Zuo. Electrical power potential of a wave energy converter using an active mechanical motion rectifier based power take-off. *Renewable Energy*, 252:123477, 2025.

I. Young. Regular, irregular waves and the wave spectrum. In *Encyclopedia of Maritime and Offshore Engineering*. John Wiley & Sons, Ltd, 2017.

Y. Zhang, Y. Zhao, W. Sun, and J. Li. Ocean wave energy converters: Technical principle, device realization, and performance evaluation. *Renewable and Sustainable Energy Reviews*, 141:110764, 2021.

References

H. Zhu, W. Zhu, and W. Fan. Dynamic modeling, simulation and experiment of power transmission belt drives: A systematic review. *Journal of Sound and Vibration*, 491:115759, 2021.

6S-Principle Methodology for Constructing a Grand Panel and Comparison with Generative AI Tools

ZHIGANG WEI & PINGSHA DONG

US naval shipbuilding faces major challenges in ramping up production to meet unprecedented national security demands. One major challenge is the one-of-a-kind structural design and low volume production, both of which make it difficult to introduce state-of-the-art automation construction technologies and statistical accuracy control. The 6S Principles is a method consisting of six fundamental principles distilled from decades of research and teaching experience in the construction of marine structures. Their applications and implications are far beyond specific domains, as demonstrated previously, and could be universal and beneficial for broader applications. The aim of this study is to explore novel AI implementation of the 6S Principles for smart modular design and construction for enabling shipbuilding automation. To do so, a grand panel as a final level of the modularity definition is considered here. Advanced generative AI algorithms are then demonstrated for following the 6S Principle-based reasoning for identifying modularity at each level of the construction process leading to the final interim product which is the grant panel. The resulting modularity definitions and construction sequencing inferred from generative AI tools through initial training sessions are compared with mechanics principle-based solutions. The results are rather promising. It is found that, with some strategically defined data patterns as input, the 6S Principles and the AI generated modular construction procedures can converge to very similar results (e.g., 90% correct), indicating the potential for implementing a 6S Principle-Guided Neural Network (6SPG-NN) for more complex ship structure components.

Keywords: 6S-Principle methodology, grand panel, welding sequence, AI

4.1 Introduction

US naval shipbuilding faces major challenges in ramping up production to meet unprecedented national security demands. One of the major challenges is the one-of-a-kind structural design and low volume production, both of which make it difficult to introduce state-of-the-art automation construction technologies and statistical accuracy control. Modular ship construction, i.e., a ship is built in large pre-outfitted modules that are later joined together, and modern modular open systems architectures (MOSA) offer significant advantages in speed, quality control, cost predictability, and flexibility. When successfully managed, modular construction can shorten build schedules and deliver higher overall quality. However, modularity also brings a set of unique challenges throughout the design, production, and assembly phases. Overcoming these challenges demands careful upfront planning, mature digital-engineering tools, and a highly coordinated supply chain, which are often challenging in themselves in practice.

One potential solution, in view of some of the most recent research findings at the University of Michigan, is to introduce a smart modularity definition by combining the Group Technology (GT) approach and mechanics-based principles at interim product levels. First, the goal of introducing GT-based interim product families is to reduce the number of distinct interim products and maximize the population within each product family so that mass-production technologies and modern statistical quality control methods can be implemented to the maximum extent possible. Second, to meet today's manufacturing challenges, particularly in dealing with the construction of lightweight structures, the above conventional interim product definitions must consider important thermomechanical interaction effects on dimensional variations, distortion, and manufacturing quality. With the consideration of GT and the mechanics-based understanding of how thermomechanical interactions among materials, manufacturing processes, and interim product forms, a set of principles with an axiomatic framework has been firmly established recently and successfully demonstrated by numerous shop-floor construction examples. These fundamental principles extracted from the domain knowledge have been dubbed the 6S Principles that can cover the necessary conditions to minimize unevenly distributed residual stress/strain caused distortion, without relying on detailed thermomechanical analysis, at least for some typical simplified construction scenarios.

Based on what has already been established for the 6S Principles, the objective of this report is to demonstrate the 6S Principles and its implementation into a simple but important example of a ship grand panel. The results obtained from the 6S Principles will be compared to the construction procedures suggested by current generative AI tools, which are operated on Big Data, and the results generated can be treated as the current best practice if the prompt is relevant and accurate. If successful, further efforts will be made to develop a rigorous mathematical framework for generalizing the 6S Principles and AI implementation for demonstrating its effectiveness in achieving smart modularity. This report is organized as follows: the 6S Principles are briefly described and discussed in Section 4.2. A grand panel with pre-defined plate patterns is used as an example to demonstrate the panel construction procedure with the 6S Principles and the solutions are compared with those obtained from the generative AI tools with different levels of prompt in Section 4.3, and the results are discussed in Section 4.4 and future research directions are outlined.

4.2 6S Principles

The major drivers in naval structure construction are cost, quality, and time to completion. As a result, automation is the key, as shown in Figure 4.1. Automation requires sufficient repetitiveness in modular forms or at sub-module or component levels, also referred to as interim products. In shipbuilding, this is perhaps the most challenging task. With repetitiveness in interim products, some proven mass-production manufacturing technologies and production processes can be introduced for achieving modular manufacturing and modular construction, leading to the final structural form. Otherwise, naval structure construction would remain a labor-intensive industry. As illustrated in Figure 4.1, smart modular design in the context of Smart Manufacturing, as all industry sectors strive to achieve under the notion of Industry 4.0 today, becomes critically important to the naval structure construction for overcoming some of the challenges. The principles of interim product definitions ("6S") for achieving "smart modular construction" are Group Technology-based interim product definitions, seeking repetitiveness based on similarity in forms, shapes, or geometry, and similarity in processes. Mechanics-based interim product requirements for accuracy control based on symmetry in forms and processes, stability in geometry, and supportability.

In summary, the smart modularity definitions can be described as the 6S Principles as follows.

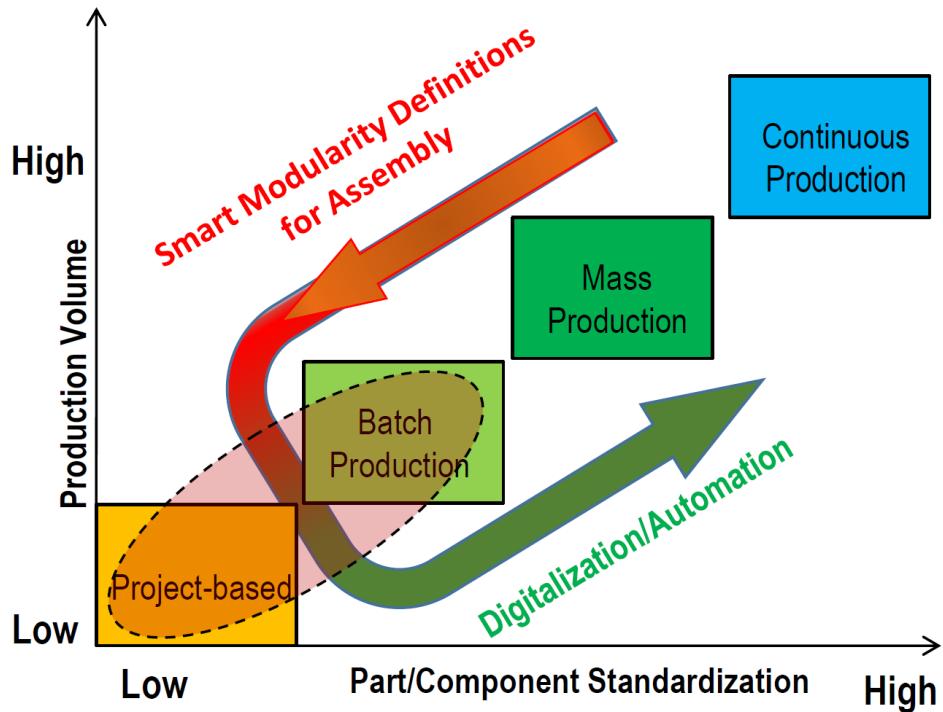


Figure 4.1: Smart modularity definitions for production volume and part standardization.

Principle 1 Similarity in interim product forms

Principle 2 Similarity in processes

Principle 3 Symmetry in interim product forms

Principle 4 Symmetry in processes

Principle 5 Stability in interim product shape during manufacturing and assembly

Principle 6 Supportability in positioning, clamping, locating, etc.

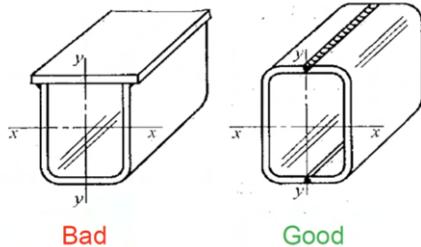
The first two principles are taken from the concept of Group Technology developed in Japan in the 1980s and transferred to the US shipbuilding industry in the late 1990s. The last four are mechanics-based principles developed by UM researchers over the last decade or so. If there is any lack of the above attributes, efforts must be taken on how to introduce them through a “divide” and “combine” process to either redefine an interim product or through alternative process sequencing and introducing temporary support and reinforcement for achieving the above criteria. The 6S Principles seem to be complete based on all the demonstrations conducted so far, and have been found to be critically important at various stages of the construction of naval structures.

Figure 4.2 shows several examples demonstrating the importance of 6S Principles in structure construction and why some of the interim product attributes should be considered. One simple interim product example is a fabricated rectangular box beam component, which can be constructed through two interim product forms (plate strip and C-shaped section, and two C-shaped sections on the left in Figure 4.2 for welding assembly into a box beam complex component. The first

one can be made by performing fillet-welding of a “C” section (interim product form) onto a plate strip (another interim product), while the second one can be made by butt-seam welding of two “C” sections (interim product). From a dimensional accuracy control point of view and the interim products involved (i.e., “C” sections), the second design is clearly a better interim product form since it only deals with one single “C” sections and enables symmetrically positioned butt seam welds with respect to two neutral planes (i.e., $x-x$ and $y-y$ planes), which would significantly reduce welding-induced distortions in the form of “longitudinal bow”. In contrast, welding assembly of the box component in the first design will not only have to deal with two distinct interim product forms (plate strip and C section), but also inevitably introduce longitudinal bow distortion with respect to the $x-x$ axis due to longitudinal shrinkage of the two fillet-welded zones. An additional advantage of the second design is that it maximizes repetitiveness in the form of “C” sections in the manufacture of the box components. The right structures shown on the right in Figure 4.2 have similar issues, especially a lack of symmetry in geometry, structural stability, and supportability.

Enhanced Interim Product Definitions (6S Principles)

- Group Technology concept:
 - Similarity in interim product form
 - Similarity in process
- Mechanics-based principles
 - Symmetry
 - Geometry
 - Process
 - Stability in shape
 - Supportability



Bad in meeting all 6S's, especially last 4

Figure 4.2: Examples demonstrating the importance of 6S Principles in structure construction.

4.3 Modularity definition and build sequence – a test case: a grand panel

A grand panel starting from standardized plate ($w \times l$)

A “grand panel” refers to the modern shipbuilding process, where large and prefabricated sections of a ship are constructed separately and then joined together, reflecting the concept of modular construction. Compared to traditional shipbuilding methods, which involve building the ship from

the keel up, the modularity and grand panel-based approach increase quality and efficiency, and reduce construction time and cost through improved dedication, simplification, repeatability, automation, and logistics. A schematic of a typical grand panel is shown in Figure 4.3, where the grand panel ($W \times L$) is made as an interim product by butt-seam welding of a standardized base plate size ($w \times l$). $W = 8w$ and $L = 4l$ are respectively the width and length of the grand panel; w and l are respectively the width and length of the standard base plate. The constructed grand panels ($W \times L$) can then be used for the construction of the deck and bulkheads by cutting them to size. To offset plates in a symmetrical manner across the width direction, a half plate is cut from a full plate to make the following patterns starting from the bottom all the way to the top of the grand panel: HFFFH, FFFF, HFFFH, FFFF, HFFFH, FFFF, HFFFH, FFFF, with F representing a full plate while H represents a half plate.

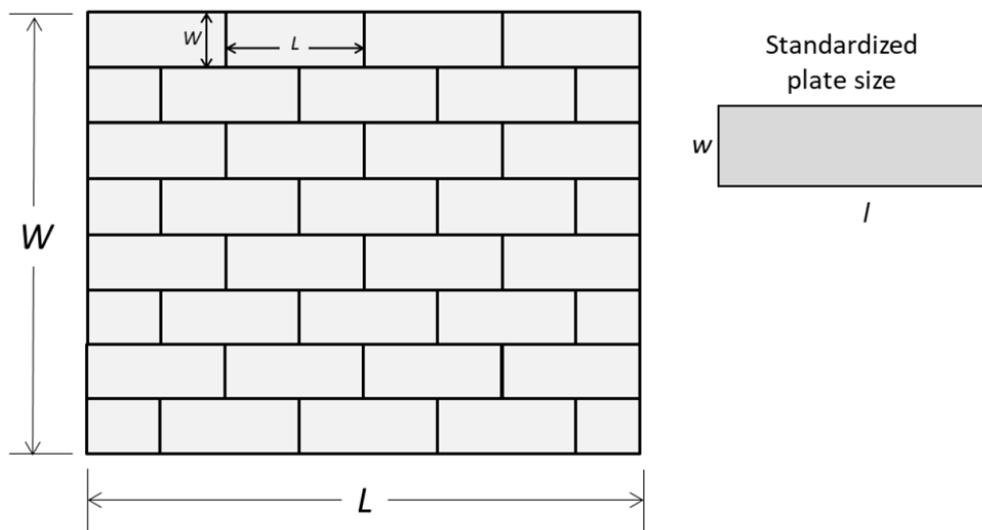


Figure 4.3: A grand panel ($W \times L$) with 8 rows in the width direction and alternating HFFFH and FFFF patterns along the length direction. F represents a standardized full base plate, and H represents a half base plate.

In Figure 4.3, the base plates of the predefined grand panel are stacked in staggered patterns in the width direction to interlock the base plates, distribute loads, and prevent potential cracks from forming along relatively weak straight vertical welding lines under potentially vertical bending dominant loads that result in normal stress in the length direction. By sticking to this brick stacking pattern, the “weaving” of the plates adds significant strength and stability to the grand panel, with load being transferred across multiple plates and through the stronger plates instead of concentrating in a single welding line. The horizontal welding lines are allowed to be straight under the assumption that the normal stress along the width direction is minor as compared with that along the length direction. It should be noted that the configuration design of the stacking patterns belongs to the domain knowledge of structural mechanics and fracture mechanics, and is not a part of the 6S Principles. Instead, the focus of this demonstration is for the given grand panel with specific plate stacking patterns such as the one shown in Figure 4.3, what is the best construction sequence that can minimize overall distortion caused by residual stress and strain during construction?

6S Principles for the construction of the grand panel

As a set of six general principles for guiding naval ship construction, the 6S Principles can be seen as a qualitative method at the current stage. Therefore, it will not explicitly provide a detailed procedure on how to construct a structure from the very beginning to the end. Rather, it provides an overarching measure of whether a construction is made based on symmetry and similarity principles. Certainly, many potential construction procedures follow the symmetry and similarity principles, but not necessarily all these procedures are equal and the best. Instead, if there are many options, either maximization or minimization must be the underlying mechanism for performance optimization. Without quantified measures for these principles yet, the 6S Principles can be understood as a procedure to maximize symmetry and similarity in all possible levels and scales. In other words, the optimized solutions obtained from the 6S Principles are unique under a given set of constraints. Based on this understanding, the following 6 executable work steps of the interim products and associated processes can be defined and illustrated in Figure 4.4, which best reflect the spirit of the 6S Principles for the interim product requirements in minimizing distortion and increasing construction quality.

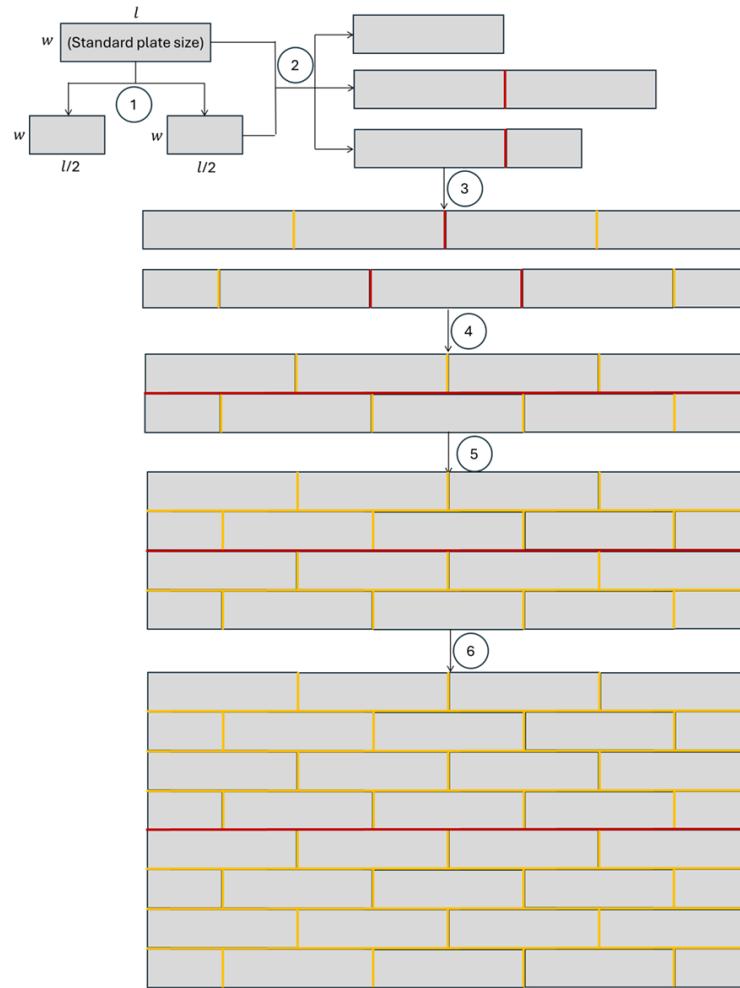


Figure 4.4: The grand panel ($W \times L$) constructed by following the 6S Principles.

Step-1 Cut to size, i.e. ①, cutting plates of standard size into equally-sized two halves

by following the similarity (in both geometry and process) and symmetry (in geometry) principles.

Step-2 Short butt-seam welding I, i.e. ②, welding the full base and/or half plates into partial rows by following the similarity (in process) and symmetry (in process) principles.

Step-3 Short butt-seam welding II, i.e. ③, welding the partial rows into a full row by following the similarity (in both geometry and process) and symmetry (in both geometry and process) principles.

Step-4 Long but-seam welding I, i.e. ④, welding two full rows into 2-row 1/4 panels by following the symmetry (in geometry) principle.

Step-5 Long butt-seam welding II, i.e. ⑤, welding the two 2-row panels into a 1/2 4-row panel by following the symmetry (in geometry) principle.

Step-6 Grand panel assembly-butt seam welding, i.e. ⑥, welding two 4-row 1/2 panels into a complete full grand panel assembly by following the symmetry (in geometry) principle.

It should be noted that only 4 of the 6S Principles are significantly involved in the grand panel construction from Step-1 through Step-6, and they are symmetry and similarity in both geometry or shape and process. The remaining 2 principles, i.e., stability and supportability, are more related to the physical aspects and conditions, which should be maximized theoretically. In other words, the symmetry and similarity are evaluated based on certain stability and supportability conditions.

6S-Generative AI solutions for the construction of the grand panel

To compare the constructed build sequence generated by the guidance of the 6S Principles, extensive AI prompts are conducted with several generative AI tools, e.g., GPT4, UM GPT4.1, UM o4-mini (reasoning), UM-GPT-Image-1 (image generation), and the procedures generated from various given prompts are compared. In this report, only three typical prompts, i.e., Prompt-1, Prompt-2, and Prompt-3, are provided for comparison and for shedding light on the gaps and opportunities. Table 4.1 shows the key features of the generated grand panels and the estimated matching scores, which are subjective, between the grand panel construction under the guidance of 6S Principles and those generated by the generative AI tools. For more details, see the Appendix.

In Prompt-1, the overall dimensions of the grand panel $W \times L$ and the corresponding constituents, i.e., full base plate and half plate, were given. Additionally, the description of the alternating brick stacking patterns: FFFF, HFFFH, was also provided. Finally, the 6S Principles and similar details about the predefined pattern were also added. The GPT tools generated essentially the same results as those obtained from the 6S Principle-based grand panel construction. However, the detailed strategy and algorithms behind the construction sequence were not provided by the GPT tools. The detailed output in terms of text description is listed in the Appendix for reference.

To further identify the potential gaps and opportunities of the 6S principle and the associated AI implementations, two additional scenarios were considered: (1) explicitly add the underlying construction strategy on top of Prompt-1 by using “hierarchical” or other “scaling” mechanisms and algorithms, Prompt-2, (2) eliminate the description of the brick stacking pattern, and with or without including the 6S Principles from Prompt-1, that results in Prompt-3.

6S-Principle Methodology for Constructing a Grand Panel and Comparison with Generative AI Tools

In Prompt-2, it is found that with the explicitly added underlying construction mechanisms, such as “hierarchical” construction procedure or other “scaling” requirements, the generative AI tools not only provide the same results as the 6S Principles suggested, but also provide the construction strategy, such as “divide-and-conquer”, and the associated procedure and logic with scaling concepts, such as assembly, subassembly, sub-panel, super-module etc., see the Appendix.

In Prompt-3, without explicitly setting the brick stacking pattern upfront, the AI tools cannot generate the desired brick patterns, though certain levels of similarity and symmetry were kept for minimizing distortion, see the Appendix.

| | Base full and half plates | With defined stacking pattern | 6S Principles | Scaling | Score in matching |
|----------|---------------------------|-------------------------------|---------------|-----------|-------------------|
| Prompt-1 | Yes | Yes | Yes | No | 90–100 |
| Prompt-2 | Yes | Yes | Yes | Yes | 95–100 |
| Prompt-3 | Yes | No | Yes or No | Yes or No | 40–60 |

Table 4.1: Key features of the generated grand panels generated by the generative AI tools.

The experience with these AI tools has shown that they are powerful assistants, but their effectiveness depends heavily on human oversight, the quality of their training data, and the context of their use. Additionally, the output results varied with the AI tools used, the information in the prompt, and the time. How to frame the training process and design effective queries by embedding some of the 6S Principles in using the AI tools is important to get meaningful results for the comparison study.

4.4 Discussion

First, the 6S Principles provide necessary conditions for shipbuilding, i.e., all well-designed ships must satisfy the 6S Principles, and the designs that fail to meet the 6S Principles would not meet the highest quality standard of shipbuilding. However, the 6S Principles at the current stage do not explicitly provide a specific detailed procedure on how to build ship structures. In fact, there would be many potential permissible designs that satisfy the 6S Principles, but the shipbuilders must choose the optimal or best solutions. Mathematically, one possible way to do it is to maximize the 6S Principles, for example, maximize similarity and symmetry, provided that the 6S Principles can be quantified. Although the current 6S Principles do not explicitly provide a path to a specific construction procedure, maximizing similarity and symmetry may lead to the optimized solution among all candidate pathways. Quantifying similarity and symmetry will be the key to unlocking the full potential of the 6S Principles, but the current symmetry and similarity assessments in the engineering community are essentially based on the binary measure system: yes or no. Normalized continuous similarity and symmetry measures in the range of $[0 - 1]$ are essential for optimization and AI implementations.

Similarity measures are often developed using distance measures such as Euclidean and Manhattan distances, and similarity is often inversely proportional to distance. Additionally, non-dimensional similarity measures provide great opportunities to treat multi-dimensional and mixed-type data in a unified framework. Gower’s distance (Gower, 1971), originally developed in the chemical field, is such a good potential similarity measure for quantification of 6S Principles. Equation 4.1 is the similarity measure based on Gower’s distance and equation 4.2 is the normalized

one.

$$s_{ijk} = 1 - \frac{|x_i - x_j|}{R_k}, \quad (0 \leq s_{i,j,k} \ll 1), \quad (4.1)$$

$$S_{i,j} = \frac{\sum_{k=1}^p w_{i,j,k} s_{i,j,k}}{\sum_{k=1}^p w_{i,j,k}}, \quad i,j = 1, \dots, N, \quad (4.2)$$

where x_i and x_j are two entities to be measured for similarity. R_k is the range of k th variable of feature to be measured, $s_{i,j,k}$ is the normalized similarity between the two entities. $S_{i,j}$ is a single weighted average measure per object pair (i,j) having p descriptors, and $w_{i,j,k}$ is non-negative weights.

Symmetry can cover reflection, rotation, translation, glide reflection, point, spherical, radial, and many other possible aspects. The continuous symmetry measure (CSM) ((Zabrodsky et al., 1992), (Alon et al., 2023)) in the chemical field can be utilized for the quantification of the 6S Principles. The key concept of CSM is: the amount of a given symmetry in a structure is a function of the minimal distance that the vertices of the structure must undergo in order for it to attain the desired symmetry (a normalized root-mean-square deviation from the closest structure with the desired symmetry). Eqn.(3) and Eqn.(4) provide simple mathematical expressions of the CSM.

$$S = \frac{\sum_{k=1}^N |Q_k - P_k|^2}{\sum_{k=1}^N |Q_k - Q_0|^2}, \quad (0 \leq S \leq 1), \quad (3)$$

$$Q_0 = \frac{\sum_{k=1}^N Q_k}{N}, \quad (4)$$

where Q_i is the vector of vertices in distorted structures, P_k is a vector of the vertices of the nearest perfectly G -symmetric object (G being a specific symmetry group), and Q_0 is the reference level (coordinate vector of the center mass of the structure). Both these similarity and symmetry measures can be used to quantify geometry and processes.

Second, by explicitly adding an underlying construction mechanism, such as a hierarchical construction procedure, an overall “divide-and-conquer” strategy, and the associated cross-scale implementation can be revealed, as demonstrated by generative AI tools. The scaling-related terminologies, such as assembly, subassembly, subpanel, super module, etc., generated by the AI tools indicate that the 6S Principles can benefit from explicitly adding regulation mechanisms and scaling parameters across different scales or levels.

Third, for more complex naval structures, such as ship bow, stern, and other irregular or curved hulls, both domain knowledge and quantified 6S Principles need to be considered to determine the coupled problems of optimized grand panel construction patterns and the corresponding build sequence. Advanced AI technology can empower the 6S Principles to fully realize their potential in naval structural construction. Deep learning has revolutionized many machine learning tasks in recent years, and it is especially promising for implementing the 6S Principles. Graph Neural Networks (GNNs) (Wu et al., 2020) are especially attractive for learning over module-interface graphs, embedding each module’s attribute vector into a latent space that respects connectivity. Physics-Informed Neural Networks (PINNs) (Cuomo et al., 2022) bring the known physical basis to the original data-driven AI and provide explainability and interpretability of the results. Kolmogorov–Arnold Networks (KANs) (Liu et al., 2024b,a) are certainly intriguing from a theoretical

standpoint: the Kolmogorov–Arnold representation theorem guarantees that any continuous multi-variable mapping can be written as sums of univariate (one-dimensional) functions composed with simple linear forms. KAN can also provide modular and explicit mathematical relationships between two entities through the edges of the neural networks, which is a much-needed framework for the 6S Principles-enabled AI implementations. The development of a 6S Principles Guided Neural Network (6SPG-NN) by effectively embedding the 6S Principles in the AI tools is an important part of this research. With an established rigorous mathematical framework for generalizing the 6S Principles and AI implementation as outlined here, the effectiveness in achieving smart modularity by selecting several more ship structure components can be further demonstrated. The knowledge, procedures, and tools generated from this project will bring significant value to shipbuilding managers and engineers in improving shipbuilding quality and efficiency.

Finally, the 6S Principles, particularly the similarity and symmetry in both geometry and process, seem to be universal and can benefit other applications, in addition to manufacturing and construction. For example, these principles enabled by AI implementation could provide alternative solutions to complex mechanical and physical problems, such as time-and frequency-domain multi-axial fatigue (Wei and Dong, 2025b), fluid-structure-interaction and wet vibration (Wei and Dong, 2025c), elastic-plastic deformation and low-cycle fatigue of typical engineering structures (Wei and Dong, 2025a), and the characterization of load-path nonproportionality effects on fatigue (Wei et al., 2025), which are still handled diligently with the traditional deterministic methods (Wei and Dong, 2025a,b,c; Wei et al., 2025). Additionally, the analytical solutions of these problems can be further used as a baseline to validate the 6S Principles in broader applications.

4.5 Conclusions

This report presented some general historical information and the domain knowledge that led to the development of the 6S Principles. Through the example of grand panel construction, the 6S Principles demonstrated it's capability in suggesting an accurate and efficient construction procedure even without the involvement of domain knowledge for a simple and commonly used naval ship plate construction pattern. This report also demonstrated the potential significance of the 6S Principles in meeting the US Navy's needs in the smart modularity construction of naval structures. The major observations made in this research can be summarized as follows:

Guided by the 6S Principles, particularly the principles of similarity and symmetry in both geometry and process, a grand panel with a pre-defined plate construction pattern has been completely determined without additional domain knowledge. The generative AI tools provide almost the same construction sequence for the same input in terms of prompts, without providing the underlying mechanisms and algorithms. This clearly demonstrates that the 6S Principles can not only provide general guidance in principle but also provide the capability to realize a detailed procedure by following certain underlying optimization mechanisms, for example, symmetry and similarity maximization.

Quantifying the 6S Principles and introducing algorithms that can maximize similarity and symmetry for more general scenarios seem to be the necessary next steps in the line of development. A simple way to do it is to provide a scaling parameter quantitatively regulating the “divide-and-conquer” strategy across different levels of scale, as demonstrated from the implementation of “hierarchical” in the AI examples. Additionally, combining domain knowledge and the 6S Principles seems to be necessary if the naval structures are complex in nature. The development of a 6S Principles Guided Neural Network seems to enable the AI implementation of the 6S Principles

and has made a significant difference in smart modularity definition and implementation into the construction of marine structures.

Appendix: Generative AI Solutions for Grand Panel Construction

Prompt-1: Full Description of the Grand Panel

A base plate has width w and length l . A half base plate has width w and length $l/2$. The total width and length of a grant panel are $W = 8 * w$ and $L = 4 * l$, respectively. The grand panel is constructed by combining the hull and half plates as follows: one row has 4 full basic plates, and the next row has 3 full plates and 2 half plates. If we use "F" and "H" to represent full and half plate, respectively, then the grand panel is constructed as: [FFFF; HFFFH; FFFF; HFFFH; FFFF; HFFFH; FFFF; HFFFH]. The plates are connected through butt welding. To minimize distortion after welding, the following principles need to be followed: Identify similarities in interim product forms, identify similarities in processes, consider symmetries, shape rigidity, and support surface. Given the information above, suggest a welding sequence that maximizes symmetry. Provide answers step by step, for example, how many FF are combined in the first step, and what sequence a whole row is welded with another row.

GPT4 answers:

To create a welding sequence that maximizes symmetry, we want to distribute heat as evenly as possible and minimize distortion. Given your row pattern [FFFF; HFFFH; FFFF; HFFFH; FFFF; HFFFH; FFFF; HFFFH], let's work through a suggested sequence step-by-step:

Initial Considerations:

1. ****Symmetrical Pattern**:** Since each pair of adjacent rows has a symmetrical counterpart, we need to take advantage of this to manage heat distribution.
2. ****Alternating Rows**:** Note that full rows ("FFFF") and alternating rows ("HFFFH") allow for symmetry.
3. ****Sequential Build**:** We should work from the center outward or alternate sides to achieve balanced tension.

Suggested Welding Sequence:

1. Assemble Partial Rows:

- ****Row 1 (FFFF):****
 - Weld plates 1 and 2 together.
 - Weld plates 3 and 4 together.
 - Weld the pair (1-2) to (3-4).
- ****Row 2 (HFFFH):****
 - Weld half plate 1 to full plate 2.
 - Weld plate 3 to plate 4.
 - Weld plate 5 (half) to plate 4.
 - Weld plate group (H2) to (3-4H).

2. Assemble Entire Rows:

- Now, combine the subcomponents of the rows:
- Weld Row 1 with Row 2 at the common edge.
- Repeat for subsequent sets of rows in the pattern.

3. Sequence the Rows:

- **Center Outward Assembly Approach:**

- Begin with the central rows. If you assume the panel has 8 rows (alternating FFFF and HFFFH), start with welding row segments 3 (FFFF) and 4 (HFFFH) together and do the same for rows 5 (FFFF) and 6 (HFFFH).

- This central approach gives a point of symmetry around which distortion is balanced.

- **Outer Rows:**

- Once central rows are connected, move outward. Combine row 1 (FFFF) with the row 2 (HFFFH) module and the same for row 7 (FFFF) with row 8 (HFFFH).

- Ensure symmetry by alternating welds from the top & bottom pairs (e.g., weld top rows 1 and 2, then bottom rows 7 and 8).

4. Final Assembly:

- Combine the top composite (Rows 1-2, 3-4) with the bottom composite (Rows 5-6, 7-8).

- Ensure the final weld is done in a back-and-forth manner to distribute heat evenly over the seam.

Additional Tips:

- Monitor the cooling and adjust the sequence as required.

- Clamp and fixturing to maintain alignment and rigidity.

- Consider pre-heating based on material type to reduce thermal gradient.

This method should provide you with minimal distortion by taking advantage of symmetry and balanced welding practices. By working outward from the center and balancing welding between fully symmetrical sections, you can minimize the common issues associated with welding distortion.

Prompt-2: Adding Additional Welding Sequence Mechanisms

A full base plate has width w and length l . A half base plate has width w and length $l/2$. A grand panel to be built from the combinations of the full base plates and half base plate has width $W = 8 * w$ and length $L = 4 * l$. If "F" and "H" respectively represent full base plate and half base plate, and the grand panel will be built with the following patterns: first row: HFFFH; second row: FFFF; third row: HFFFH; fourth row: FFFF; fifth row: HFFFH; sixth row: FFFF; seventh row: HFFFH; eighth row: FFFF. The plates are welded together to form the grand panel. To minimize distortion after welding, the following principles need to be followed: identify similarities in interim product forms, identify similarities in processes, consider symmetries in geometry and welding process, shape rigidity, and support surface, group technology, clustering technique, and hierarchical welding process. Given the information above, what is the welding sequence that can maximize symmetry? Provide step-by-step, detailed answers. Provide a sketch or graphic illustration of the welding procedure.

Solutions from UM-GPT 4.1

To develop an optimal welding sequence for the grand panel, maximizing symmetry and following the manufacturing guidelines you listed, let's proceed step-by-step. We'll reference the given dimensions and constraints:

Given:

Full base plate (F): width = w , length = l

Half base plate (H): width = w , length = $l/2$

Grand panel: width $W = 8w$, length $L = 4l$

Plate arrangement:

6S-Principle Methodology for Constructing a Grand Panel and Comparison with Generative AI Tools

Rows 1, 3, 5, 7: H F F F H

Rows 2, 4, 6, 8: F F F F

There are 8 rows in total, alternating between these two patterns.

| Row | Plates |
|-----|-----------|
| 1 | H F F F H |
| 2 | F F F F |
| 3 | H F F F H |
| 4 | F F F F |
| 5 | H F F F H |
| 6 | F F F F |
| 7 | H F F F H |
| 8 | F F F F |

Table 4.2: Visual layout (each letter indicates a plate, each row is under the previous).

Key Principles to Follow

Symmetry: Balance weld heat-distribution to minimize distortion.

Grouping/Clustering: Build interim sections with similar geometry.

Hierarchical: Weld subassemblies first, then assemble the grand panel.

Shape Rigidity: Build subassemblies with higher rigidity first when possible.

Step 1: Subassemblies (Interim Products)

Overlapping similarities among the rows allows us to define subassemblies:

Type A Subassembly: One "H F F F H" row (occurs on rows 1, 3, 5, 7).

Type B Subassembly: One "F F F F" row (occurs on rows 2, 4, 6, 8).

Construct four A's and four B's as interim subpanels:

A₁ = Row 1; A₂ = Row 3; A₃ = Row 5; A₄ = Row 7

B₁ = Row 2; B₂ = Row 4; B₃ = Row 6; B₄ = Row 8

Step 2: Subassembly Construction Sequence

a. Assemble all "A" rows:

For each A:

Symmetrically weld the two "H" plates at both ends (start from the center, move outwards).

Weld the 3 "F" plates in the center together (in order: F2—F3, then F1—F2).

Connect the 'F' subassembly' to the 'H' plates on both ends.

Weld sequence within a row: Start at the middle "F" plates, then outward to the "H" ends.

[Symmetry preserved in each subassembly.]

b. Assemble all "B" rows:

For each B:

Weld F2—F3 in the middle.

Weld F1—F2, then F3—F4.

Again, start at the center and move outward for symmetry.

Step 3: Grand Panel Assembly (Vertical Stacking/Clustering)

Now, alternate and stack the A and B subassemblies:

Stack as: A₁ (Row 1) + B₁ (Row 2) + A₂ (Row 3) + B₂ (Row 4) + ... + A₄ (Row 7) + B₄ (Row 8)

Symmetric welding sequence across the height: Weld from both ends and meet in the center

6S-Principle Methodology for Constructing a Grand Panel and Comparison with Generative AI Tools

(i.e., weld A_1 to B_1 at top and bottom, then proceed inward: A_2 to B_2, then meet at the center).

Step 4: Hierarchical Symmetric Welding Sequence

Detailed Final Welding Steps:

Layout all 8 subassemblies ("A" and "B") in order.

Start by tack welding at 4 corners to hold assembly.

Weld at both ends first:

Weld A_1 to B_1 (topmost pair), and A_4 to B_4 (bottommost pair).

Weld next pair inward:

A_2 to B_2, and A_3 to B_3 (moving in from each end toward center).

Finish by welding the central seams simultaneously, if possible (A_2—B_2 to A_3—B_3), maintaining symmetry.

For each seam, alternate the starting point left/right, and top/bottom, to allow even heat dissipation and minimal distortion.

Step 5: Symmetric Welding Within the Full Grand Panel

For maximum symmetry within the complete panel:

Divide the panel into left and right halves (vertical axis).

Weld seams and joints on both sides simultaneously (left/right mirroring).

Alternate welding order: start from the center and move outward, then from outward back to the center.

| Stage | Components | Method (Symmetry) |
|-------|----------------|--|
| 1 | A rows (HFFFH) | Center-out, mirror assembly |
| 2 | B rows (FFFF) | Center-out, mirror assembly |
| 3 | Stack A+B rows | Top/bottom in, left/right mirror welding |
| 4 | Join center | Meet at center from both ends/halves |
| 5 | Final seams | Alternate, mirrored, for thermal balance |

Table 4.3: Summary Table of Hierarchical Sequence.

An alternative prompt resulting in similar answers (Group Technology used in GPT-4.1)

Recommended Step-by-step Welding Sequence:

Step 1: Assemble and Weld identical rows as Submodules

Build all HFFFH rows (1, 3, 5, 7):

Lay out H, F, F, F, H in a single line, tightly abutted.

Weld joints between each plate, working from center outwards for symmetry:

Start with center two F's, weld their joint.

Weld F-H joints on both ends.

Weld remaining F's (i.e., F to H) on both sides simultaneously.

This maintains symmetry and reduces possible distortion as weld contraction happens evenly outwards.

Build all FFFF rows (2, 4, 6, 8):

Lay out four F plates.

Weld two center F's together first, then symmetrically add and weld the outer F's.

Step 2: Pair Rows into Double-row Supermodules

Take HFFFH (Row 1) + FFFF (Row 2) and weld them together along their matching seams.

Repeat for (Row 3 + 4), (Row 5 + 6), (Row 7 + 8).

6S-Principle Methodology for Constructing a Grand Panel and Comparison with Generative AI Tools

Welds should be made from the center of the seam outward, on both sides, to keep heat input balanced and maintain flatness.

Step 3: Assemble and Weld Supermodules

Now you have four supermodules, each covering the full panel width ($8w \times 2l$).

Place the first pair of supermodules (1+2 and 3+4) adjacent, align, clamp, then weld their interface (middle of the panel). Weld from the center line outward symmetrically.

Do the same for the bottom half (5+6 and 7+8).

Step 4: Final Assembly – Combine Halves

Now, weld the two halves (upper and lower; each $8w \times 2l$) together along the final interface (a horizontal seam).

Again, weld from the central line outward, ensuring clamping and support for flatness.

Prompt-3: Without Predefined Plate Patterns

A full base plate has width w and length l . A half base plate has width w and length $l/2$. A grand panel to be built from the combinations of the full base plates and half base plates has a width $W = 8 * w$ and a length $L = 4 * l$. The plates are welded together to form the grand panel. Given the information provided above, what is the welding sequence that can minimize the distortion of the grand panel after welding? Provide a sketch or graphic illustration of the welding procedure.

Solutions from UM-GPT O4-mini (reasoning)

Here's one practical way to lay out and sequence your welds so that you always "balance" the heat input around the geometric center of the 8×4 panel and thus minimize net distortion. In what follows we think of your full plates ($w \times l$) and half-length plates ($w \times l/2$) as laid out on a grid of 4 rows (in the L-direction) by 8 columns (in the W-direction), and we always weld in small intermittent "blocks," pairing seams symmetrically about both the center-line in length and the center-line in width.

1. PREP & TACKING

- Lay all plates in a flat fixture, clamp them so that the grid is held to flatness tolerances.
- Put light tack welds at every intersection, paying special attention to the two central seams (row 2/3 at column 4/5).

2. WELD THE CENTRAL BAND (row 2 \longleftrightarrow row 3)

- Identify the four seams in the center band:
 - Seam A: between row 2 & 3 at columns 4–5 (this straddles the mid-width line $W/2$).
 - Seam B: between row 2 & 3 at columns 3–4
 - Seam C: between row 2 & 3 at columns 5–6
 - Seam D: between row 2 & 3 at columns 2–3 and 6–7, etc.
- Tack-re-check A, then run two or three short weld "ladders," alternating left and right of the center-line, e.g.:
 - Weld a 100 mm segment on the left side of seam A
 - Jump to the corresponding 100 mm on the right side of seam A
 - Back-step and repeat until seam A is complete.
- Move out to B & C as a pair: weld a 100 mm block on the left side of seam B, then its mirror on seam C, then back-step, etc.
- Finally, do D & its mirror on the far side of the center in the same block-and-mirror fashion.

3. ADD THE ADJACENT ROWS

You have now welded all of the center-band seams. Next, weld the seams between row 1 \longleftrightarrow 2 and between row 3 \longleftrightarrow 4—again in pairs, always matching one weld on the "top" half of the panel with its mirror in the "bottom" half:

- Pair 1: seam (row 1–2 at col 4–5) with seam (row 3–4 at col 4–5)
- Pair 2: seam (row 1–2 at col 3–4) with seam (row 3–4 at col 5–6)
- ... and so on, moving out toward the edges.

Within each pair, you still use short runs (block welds) and alternate back and forth.

4. FINISH THE OUTER PERIMETER

Once all inboard seams are done, stitch-weld the outer border joints in one continuous loop or in two halves (left side then right side), again using short segments and alternating top/bottom or left/right sides to avoid heat build-up.

5. KEY POINTS TO MINIMIZE DISTORTION

- Always start in the geometric center (that first center band of half-length plates).
- Work outward in matched pairs so that each weld's shrinkage is counteracted by its mirror weld.
- Keep individual weld runs short (100–200 mm), back-step frequently, and allow cooling between blocks.
- Reserve the outermost perimeter till last—by then the panel is stiffened by the interior welds, so final shrinkage has less lever arm.

By following a “center-out, mirror-block” welding sequence, you ensure that heat is introduced symmetrically and any shrinkage in one zone is immediately balanced by its opposite, thus driving net distortion toward zero.

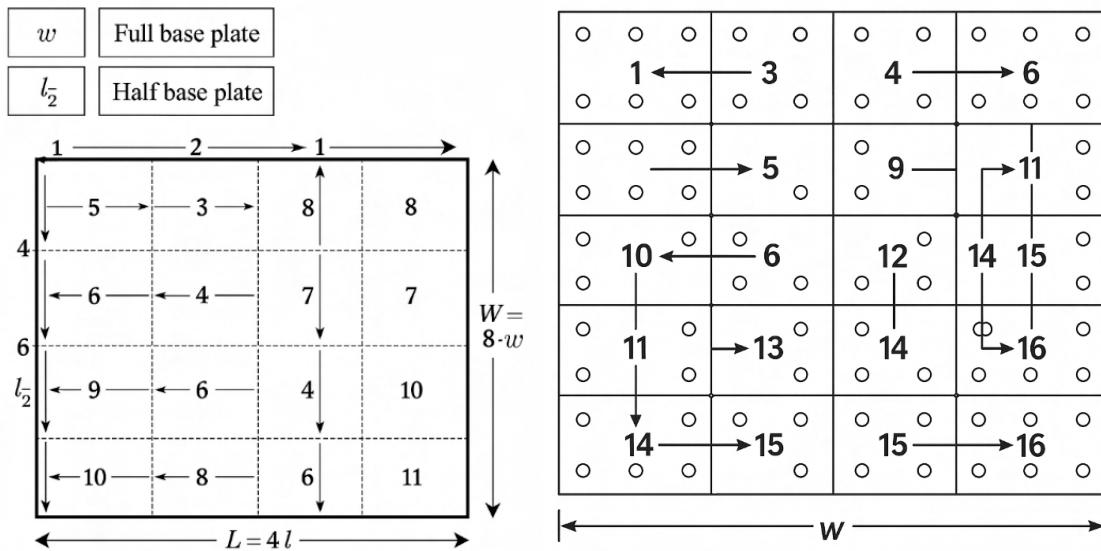


Figure 4.5: Two sketches of welding sequence and patterns provided by an AI tool for Prompt-1.

Figure 4.5 shows the sketches generated from UM-GPT-Image-1(image generation) for two similar prompts. It is clear from the sketches and the corresponding texts showing the steps to generate the sketches that the brick stacking patterns are not automatically generated, which often require domain knowledge such as fracture mechanics and structural mechanics. Without domain knowledge, the half plate and full plate will not make any difference as long as the full dimension is guaranteed and certain overall symmetry and similarity are maintained. It should be noted that the procedure in the text and the two sketches were generated from different sub-platforms of GPT, therefore, the correspondence between the test and sketch is not guaranteed. Additionally, the output results are random in nature; for the same input at different times, a little bit of difference in the prompt

often results in different procedures.

References

G. Alon, Y. Ben-Haim, and I. Tuvi-Arad. Continuous symmetry and chirality measures: approximate algorithms for large molecular structures. *Journal of Cheminformatics*, 15(1):106, 2023.

S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli. Scientific machine learning through physics-informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.

J. C. Gower. A general coefficient of similarity and some of its properties. *Biometrics*, pages 857–871, 1971.

Z. Liu, P. Ma, Y. Wang, W. Matusik, and M. Tegmark. Kan 2.0: Kolmogorov-arnold networks meet science. *arXiv preprint arXiv:2408.10205*, 2024a.

Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark. Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*, 2024b.

Z. Wei and P. Dong. An analytical structural strain method for low-cycle fatigue evaluation of thin-walled pipes under bending. *Journal of Thin-Walled Structures*, 2025a. (In press).

Z. Wei and P. Dong. Loading-path dependent time- and frequency-domain multiaxial variable amplitude fatigue analysis for marine structural reliability. In *16th International Symposium on Practical Design of Ships and Other Floating Structures PRADS*, Ann Arbor, MI, USA, October 2025b.

Z. Wei and P. Dong. Wet vibration analysis and uncertainty quantification of ship plates. In *16th International Symposium on Practical Design of Ships and Other Floating Structures PRADS*, Ann Arbor, MI, USA, October 2025c.

Z. Wei, P. Dong, and S. Wu. A load-path nonproportionality parameter for time-domain and frequency-domain multiaxial fatigue evaluation. *International Journal of Fatigue*, 2025. subimtted.

Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1):4–24, 2020.

H. Zabrodsky, S. Peleg, and D. Avnir. Continuous symmetry measures. *Journal of the American Chemical Society*, 114(20):7843–7851, 1992.

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

ANDRÁS SZABÓ & KRISHNAN MAHESH

Global linear stability analysis is utilized to understand the development of the asymmetric flow around a prolate spheroid. The base flow, the direct modes, the base flow plus the perturbations, and the skin friction lines are analyzed to characterize the flow. The study of aspect ratio effects at zero angle of attack establishes that the structure of the asymmetry is the same whether the axisymmetric body is bluff or more streamlined. Stability of the flow field around the 6:1 spheroid is then investigated for angles of attack $\alpha \in [0^\circ, 90^\circ]$ as a function of Reynolds number. Comparison of the low angle of attack results to the experimental results of Ashok *et al.* (*J. Fluid Mech.* 774, 2015, pp. 416-442) shows that the asymmetry observed in the experiments is similar to the global mode predicted by the stability calculations. It is conjectured that the experimental asymmetry is triggered by the weak cross-stream circulating flow. The leading eigenmode is a stationary asymmetric mode in the angle of attack $\alpha \in [0^\circ, 65^\circ]$ range, while above $\alpha = 65^\circ$, the leading mode is an oscillatory shedding mode. The change from a stationary to an oscillatory instability between $\alpha = 65^\circ$ and 70° is linked to the ability/inability of the vortex sheets to roll up and reattach to the body in the former/latter cases, respectively. The difference in the separation patterns and the similarity between the eigenmodes indicate that asymmetry of the flow field is governed by the same mechanism across a wide angle of attack range, regardless of whether the flow is like a bluff body wake, a streamlined body wake, or a vortex wake. Previous studies have argued that the asymmetry of vortex pairs emerges either because of a vortex instability or a separation/reattachment related mechanism; since the development of the asymmetry cannot be linked to specific features in the separation pattern in the investigated configurations, our results support the former argument.

Keywords: prolate spheroid, asymmetry, global stability analysis

5.1 Introduction

The prediction of hydrodynamic forces acting on a body in relative motion to the surrounding fluid has tremendous practical relevance in a wide range of engineering applications. Oddly, when such a configuration is symmetric, the flow field may be asymmetric (on long time scales) despite the symmetry of the arrangement. An asymmetric flow field in symmetric configurations has been observed in the case of diffusers, bumps, jets in crossflow, and wakes of both slender and bluff bodies, to mention a few examples. Despite its occurrence in a wide range of practically relevant problems, this phenomenon is not well understood (Williams and Smits, 2025).

In the case of geometries such as the prolate spheroid and the DARPA SUBOFF that mimic the main characteristics of underwater vessels, although most research did not report any asymmetry, several recent studies did. At high Re_D values ($\mathcal{O}(10^5 - 10^6)$), large eddy simulations (LES) did

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

not find asymmetry in the flow field (Plasseraud et al., 2023; Plasseraud and Mahesh, 2025), while low Reynolds number direct numerical simulations (DNS) and LES displayed a strong asymmetry without any perturbation to trigger it. For $Re \in [1000, 8000]$ at an angle of attack of 45° , a series of studies using DNS (Jiang et al., 2014, 2015, 2016; Strandenes et al., 2019a,b; Andersson et al., 2019) reported an asymmetry in the mean flow without the presence of any external excitation. Nidhan et al. (2025) reported asymmetry in the wake of the spheroid at an angle of attack of 10° degrees and $Re_L = 30000$ ($Re_D = 5000$), when the Froude number was high (the stratification was weak) in their LES.

Although most experimental research did not report asymmetry in the flow field around the spheroid, one experimental study did report a very robust flow asymmetry. Ashok et al. (2015) investigated the DARPA suboff at angles of attack $\alpha \in [0, 4, 8, 12]^\circ$ and $1.1 \cdot 10^6 < Re_L < 67 \cdot 10^6$. They documented an asymmetry in the flow field which was robust to variations in the pitch, yaw, trip wire geometry, level of free-stream turbulence, and surface roughness; furthermore, the asymmetry was observed in two different models in a water and a wind tunnel. The reason for the flow asymmetry was unknown.

There are flow configurations that share similarities with the spheroid flow; studying these, one may gain insight into the dynamics of spheroid flow. One class of problems is the flow field around slender bodies at an angle of attack, which includes cylinders with a hemisphere/conical/ogive cap, cones, and delta wings. These flows are characterized by a pair/pairs of counter-rotating vortices that separate from the body (which has also been reported in the case of the flow around the spheroid), which, depending on the parameters such as the Reynolds number, angle of attack, apex angle, leading edge/nose roundedness, tend to be asymmetric. Although these configurations have a great importance in aerospace engineering, and thus they have been receiving continued attention in the scientific community, the mechanism underlying the asymmetry of the flow field is still unknown (Bridges, 2006; Deng et al., 2008; Williams and Smits, 2025).

Asymmetry has also been observed in bluff body wakes — either bodies with axisymmetry (Rigas et al., 2014) or reflectional symmetry such as Ahmed bodies (Grandemange et al., 2013). In the case of these flows, even though the mechanism behind the asymmetry is not fully understood, there are some important observations. It has been found that the stationary asymmetric flow structures that can be observed in the low Re flow also appear in the high Re case in both the reflectionally symmetric (Grandemange et al., 2012, 2013) and axisymmetric cases (Rigas et al., 2014, 2016; Zhu et al., 2025). In addition, these flow structures can be used to derive models that describe the main parameters of the flow field (Rigas et al., 2015). Recently, Chiarini et al. (2025) used global stability analysis to identify the active regions of the flow and confirm the stability criterion of the symmetry loss of axisymmetric wakes conjectured by Magnaudet and Mougin (2007). The criterion is based on the maxima of the radial derivative of the azimuthal vorticity: the instability originates from the azimuthal vorticity wanting to follow both the rigid body, but also being transported by the free stream, which creates large gradients.

The usefulness of low Re linear stability analysis in the case of bluff body wakes suggests that it may provide insight into the mechanism of flow asymmetries in the other configurations too. Tezuka and Suzuki (2006) conducted low Reynolds number stability analysis of a 4:1 prolate spheroid at zero and $\alpha = 10^\circ$ angle of attack. They found that in both cases, when the Reynolds number is increased, the flow becomes unstable to a stationary asymmetric eigenmode. Then, when the Reynolds number is further increased, the flow starts to produce an additional shedding mode. They noted that the structure of the asymmetric mode was the same in the two cases; however, they did not provide an explanation for why the asymmetry arises in the flow field.

Thus, this work investigates the 6:1 prolate spheroid flow using global linear modal stability analysis for incompressible flow. The calculations are carried out using the open-source finite

element library FreeFEM (Hecht, 2012). First, the flow around a sphere is investigated to verify the methodology. Then, starting from the sphere, the aspect ratio of the spheroid is progressively increased to six in steps of one, and the stability of the flow is characterized. Finally, the angle of attack of the 6:1 spheroid is varied between 0° and 90° .

The rest of this report is organized as follows. Section 5.2 discusses the problem definition, modeling assumptions, and the numerical solution techniques. The results of the calculations are presented in Sec. 5.3: three separate subsections are dedicated to the verification of the implementation of the modeling equations, the zero angle of attack spheroid while the aspect ratio is varied, and the angle of attack effects in the case of the 6:1 spheroid. Concluding remarks are drawn in Sec. 5.4.

5.2 Methodology

Flow configuration and notations

The flow configuration is illustrated in Fig. 5.1, where a rigid spheroid is placed in a uniform incompressible flow. The origin of the Cartesian coordinate system in which the results are presented is located at the center of the spheroid. The x axis is aligned with the major axis of the spheroid; a change in the angle of attack α corresponds to rotations around the z axis; the "normal" coordinate y is perpendicular to the free stream at 0° angle of attack and aligned with it at 90° angle of attack. Dimensional quantities are denoted with a hat ($\hat{\square}$), while nondimensional quantities are denoted with regular letters. In the nondimensionalization of the equations, the velocity scale is the undisturbed free stream velocity \hat{U}_∞ , and the length scale is the largest diameter of the cross-section of the spheroid D , which yields the Reynolds number:

$$\text{Re} = \frac{\hat{D}\hat{U}_\infty}{\hat{\nu}}, \quad (5.1)$$

where $\hat{\nu}$ is the kinematic viscosity of the fluid.

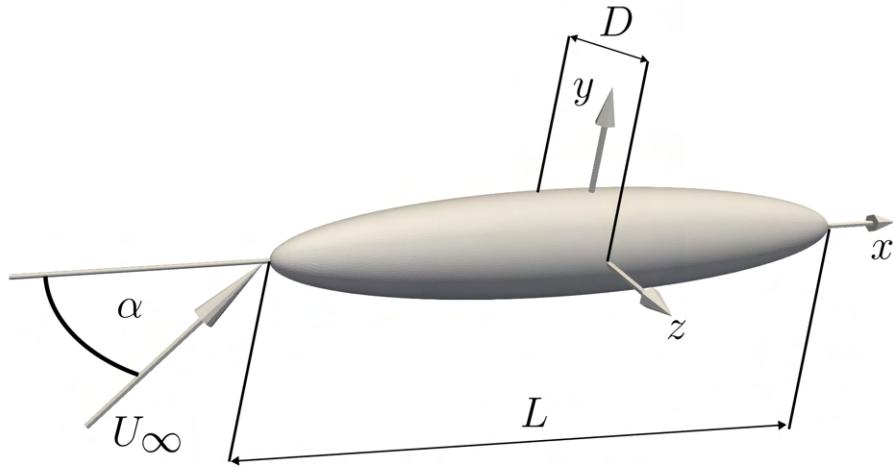


Figure 5.1: Sketch of the configuration.

Modeling techniques

Global modal linear stability analysis is utilized in the incompressible limit. In this framework, the total velocity field is decomposed into a stationary equilibrium solution, called base flow, and a small amplitude perturbation around the base flow. Let $\mathbf{U} = [U, V, W]^T$ and P denote the base flow velocity and pressure, and $\mathbf{u}' = [u', v', w']^T$ and p' denote the velocity and pressure of the small-amplitude perturbations. Then, the nondimensional governing equations of the base flow are

$$(\mathbf{U} \cdot \nabla) \mathbf{U} = -\nabla P + \frac{1}{\text{Re}} \nabla^2 \mathbf{U}, \quad (5.2)$$

$$\nabla \cdot \mathbf{U} = 0. \quad (5.3)$$

Once the stationary equilibrium base flow is obtained, its stability can be assessed using the linearized Navier-Stokes equations:

$$\frac{\partial \mathbf{u}'}{\partial t} + (\mathbf{U} \cdot \nabla) \mathbf{u}' + (\mathbf{u}' \cdot \nabla) \mathbf{U} = -\nabla p' + \frac{1}{\text{Re}} \nabla^2 \mathbf{u}', \quad (5.4)$$

$$\nabla \cdot \mathbf{u}' = 0. \quad (5.5)$$

Since the base flow is stationary, Eq. (5.4-5.5) can be further simplified by seeking the disturbances in the form

$$\mathbf{q}'(x, y, z, t) = \mathbf{q}(x, y, z) e^{\sigma t} + \text{c.c.}, \quad (5.6)$$

where $\mathbf{q} = [u, v, w, p]$, the complex number σ characterizes the temporal variation of the disturbances and c.c. stands for complex conjugate. The real part of σ , σ_r is the growth rate of the disturbances, and can be used to determine the stability of the flow: if its positive, the flow is unstable, while negative values correspond to a stable flow. The imaginary part σ_i is the angular frequency that characterizes the oscillatory nature of the disturbances. Searching for the disturbances in this form, in essence, corresponds to utilizing Laplace transform since the coefficients in the linearized Navier-Stokes equations are homogeneous in time. The ansatz of Eq. (5.6) simplifies the solution of Eqs. (5.4-5.5): the temporal variation is eliminated, but instead an eigenvalue problem needs to be solved, so the relevant flow structures can be computed directly. Typically, the few most unstable eigenvalues are sought at different Reynolds numbers, and the goal is to find the critical Reynolds number Re_c , where the most unstable eigenmode is marginally stable ($\sigma_r = 0$), since this is where the disturbances reach only small amplitudes, and linear stability analysis is appropriate.

The disturbances in the form Eq. (5.6) can be obtained by solving the modal version of the linearized Navier-Stokes equations:

$$\sigma \mathbf{u}' + (\mathbf{U} \cdot \nabla) \mathbf{u}' + (\mathbf{u}' \cdot \nabla) \mathbf{U} = -\nabla p' + \frac{1}{\text{Re}} \nabla^2 \mathbf{u}', \quad (5.7)$$

$$\nabla \cdot \mathbf{u}' = 0., \quad (5.8)$$

which are also called TriGlobal stability equations (Theofilis, 2003). It is also convenient to write the linearized equations as follows:

$$\sigma \mathbf{B} \mathbf{q}' = \mathbf{A} \mathbf{q}' \quad (5.9)$$

which clearly shows that finding the disturbances in the form of Eq. (5.6) corresponds to an eigenvalue problem. The elements of the matrices in the equation above of the matrices are straightforward to derive.

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

It is important that the eigenmodes are normalized appropriately. Using the standard inner product over the computational domain Ω

$$\langle \mathbf{q}_1, \mathbf{q}_2 \rangle = \int_{\Omega} \mathbf{u}_1^* \cdot \mathbf{u}_2 dV, \quad (5.10)$$

the eigenmodes are normalized so that $\langle \mathbf{q}, \mathbf{q} \rangle = 1$.

Numerical methods

The equations in the previous section are solved using computer programs developed in the open-source finite element package FreeFEM (Hecht, 2012). The equations are discretized using Arnold-Brezzi-Fortin MINI-elements (P1b for the velocity components and P1 for the pressure). For each configuration (aspect ratio and angle of attack), a separate unstructured tetrahedra mesh is created using the open-source software GMSH (Geuzaine and Remacle, 2009). Creating a separate mesh for each configuration is necessary to capture the structure of the wake region while keeping to size of the problem modest. The typical mesh size is between $3 - 4 \cdot 10^6$ elements.

The equations are solved in a coordinate system (ξ, η, ζ) in which the ξ axis is aligned with the free-stream, since this simplifies the specification of the boundary conditions; however, the results are presented in the coordinate system aligned with the spheroid. The rectangular computational domain, in which the spheroid is placed, is illustrated in Fig. 5.2. The center of the spheroid is also the center of the computational domain. The spheroid is placed symmetrically in the η and ζ directions; the sizes of the bounding box are detailed in Table 5.1.

| $L_{\xi,0}$ | $L_{\xi,1}$ | L_{η} | L_{ζ} |
|-------------|-------------|------------|-------------|
| -15 | 120 | 50 | 50 |

Table 5.1: The size of the computational domain.

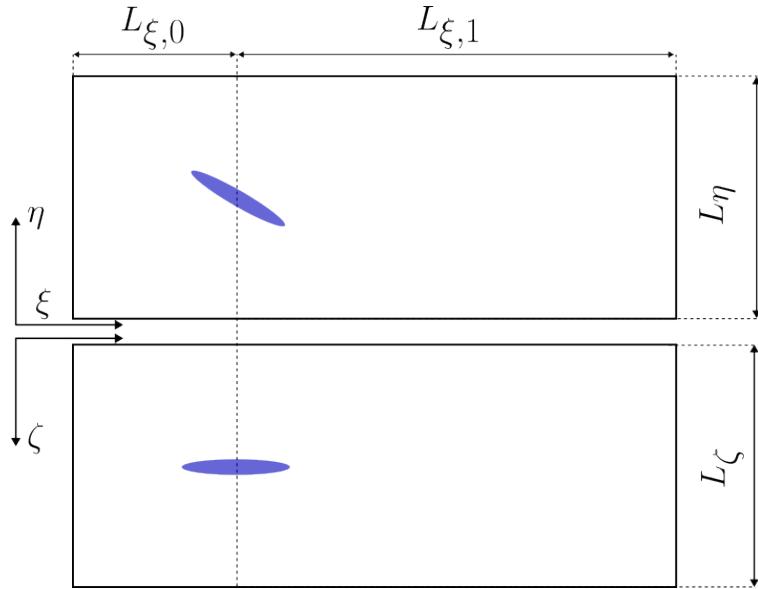


Figure 5.2: The computational domain.

The boundary conditions are the following. At the inlet, the free stream velocity is specified as the base flow ($\mathbf{U} = [1, 0, 0]$), and the perturbations are zero ($\mathbf{u} = 0$). At the wall of the spheroid,

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

no-slip conditions are enforced for both the base flow and the perturbations ($\mathbf{U} = \mathbf{u}' = [0, 0, 0]^\top$). On the side wall of the computational box, a free-slip boundary condition is prescribed (the normal velocity is zero, and the normal derivative of the tangential velocity components is zero). Finally, at the downstream end of the computational domain, a free-stress condition, which is automatically satisfied in the finite element formulation, constitutes the boundary condition.

There are two main classes of techniques that can be used to obtain the base flow and the eigenmodes. The first is the so-called matrix-forming method, and the second is the timestepper method. In the case of matrix-forming methods, the base flow is typically found using the Newton-Raphson method, in which the Jacobian is explicitly formed and its inverse is applied using LU factorization. Then, the matrices of the eigenvalue problem are calculated, followed by calculating the eigenvalues using the Krylov-Schur method combined with the shift-invert technique. In both the base flow calculation and the eigenvalue problem, the computational bottleneck is the large memory requirement of the LU factorization, which limits the problem size to roughly 20 million unknowns in the case of very sparse matrices. As far as the authors know, although there have been efforts to develop efficient iterative methods for the inversion of the linearized Navier-Stokes operator (e.g. Moulin et al. (2019)), they only offer an alternative to direct solver in the low Re limit (up to $Re \approx 200$), which is not sufficient for the present study.

In the timestepper technique, a modified CFD solver is used both to find the base flow and the eigenmodes. In the case of the base flow, the instabilities can be filtered out by various techniques (Åkervik et al., 2006; Citro et al., 2017); alternatively, some symmetries can be prescribed that prohibit the development of the disturbances. In the case of the eigenvalue calculation, advancing the linearized equations by one timestep of dt corresponds to the application of $\exp(\mathbf{A}dt)$ restricted to the velocities, where \mathbf{A} in Eq. (5.9). The advantage of the timestepper techniques over the matrix-forming methods is that they do not have a strong memory requirement that limits the problem size. However, the timestepper approach requires long simulations, even up to $\mathcal{O}(100)$ convective units, thus they have a strong computational requirement.

Based on the advantages and disadvantages of the two families of methods, similarly to the studies utilizing the resolvent (input-output) framework (Ribeiro et al., 2023; Rolandi et al., 2024) for stability analysis, the matrix-forming method is chosen to investigate the stability of the flow. The Newton-Raphson method and the shift-invert technique, both combined with sparse LU factorization, are used to obtain the equilibrium base flow and the relevant part of the eigenspectrum, respectively. The matrix forming method requires a sophisticated mesh that is sufficiently refined at the dynamically relevant regions of the flow, so for each angle of attack and aspect ratio, a separate mesh was carefully tailored to achieve this goal. The implementation of the Newton and Krylov-Schur methods in the PETSc (Balay et al., 2021) and SLEPc (Roman et al., 2022) libraries is utilized in the calculations.

The calculations are performed on the Lighthouse cluster of the University of Michigan Advanced Research Computing. The simulations are run on 20 nodes, using 48 of the 192 processors per node. The FreeFEM and PETSc libraries were compiled using Intel compilers and MPI.

5.3 Results

Verification of the implementation

To verify our methodology, critical Reynolds number of the sphere was calculated for an axisymmetric base flow to find both the first stationary asymmetric mode and also the second shedding mode that breaks the temporal symmetry (an asymmetric base flow would result in a different critical Reynolds number for the shedding mode). The critical Reynolds number, which was ob-

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

tained by interpolation, is compared to the reference values from the literature in Table 5.2. Our results match closely with the references, especially with Citro et al. (2016) and Meliga et al. (2009) (for the first and second bifurcations, respectively) who utilized a finer mesh than Natarajan and Acrivos (1993).

In addition, the leading eigenmodes of three canonical flow problems are calculated, and the results are compared with the literature: Poiseuille flow (Juniper et al., 2014), the 3D lid driven cavity (Giannetti et al., 2009; Gómez et al., 2014; Regan and Mahesh, 2017) and the 4:1 spheroid at zero angle of attack (Tezuka and Suzuki, 2006). To keep the discussion concise, the comparison of the results is not presented here, but in all cases, the agreement between our calculations and the literature is very good.

| | Natarajan and Acrivos (1993) | Citro et al. (2016) | Present calculation |
|--------------------|------------------------------|----------------------|---------------------|
| First bifurcation | 210 | 212.4 | 212.1 |
| | Natarajan and Acrivos (1993) | Meliga et al. (2009) | Present calculation |
| Second bifurcation | 277.5 | 280.7 | 280.1 |

Table 5.2: Critical Reynolds number of the first two bifurcations in the case of the sphere.

Effect of varying the aspect ratio

As the first investigation of the asymmetry, the aspect ratio is varied between 1 (sphere) and 6:1 at zero angle of attack. The base flows are calculated over a wide range of Reynolds numbers, the relevant part of the eigenspectrum of the flow is calculated, and the most unstable mode is identified. At each aspect ratio, there is a single leading mode, which is a stationary asymmetric mode, and all other eigenmodes are stable and have low growth rates. This agrees with previous investigations of axisymmetric bluff bodies (Rigas et al., 2016).

The critical Reynolds number of the first unstable mode as a function of the angle of attack is displayed in Fig. 5.3. As the aspect ratio is increased and the body becomes more streamlined, the flow becomes unstable at increasingly higher Reynolds numbers. The recirculation ratio (not shown for the sake of conciseness) monotonically decreases with an increase in the aspect ratio for a given Re value. Since recirculation is crucial in the self-sustained global mode as it is responsible for the feedback of disturbances by convecting them upstream (Chomaz, 2005), the destabilization of the flow at higher Reynolds numbers when the aspect ratio is increased may be attributed to a weaker recirculation.

Angle of attack effects for the 6:1 spheroid

In the case of the 6:1 spheroid, the main objective of the present work is to find the first bifurcation of the flow at every angle of attack to construct a stability chart as a function of α and Re , and then characterize the modes and the base flow. The stability chart is presented in Fig. 5.4 in the $Re - \alpha$ plane. The black dots denote where the base flow is stable; the blue and red dots correspond to stationary asymmetric and oscillatory shedding modes, respectively; the \times symbols mark the marginally stable states, which are approximated by interpolating the growth rate and finding $\sigma_r = 0$. For the sake of completeness, abscissas with both diameter based Re_D and the length based Re_L are presented, and the values of the critical Reynolds number values are tabulated in Tables 5.3 and 5.4.

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

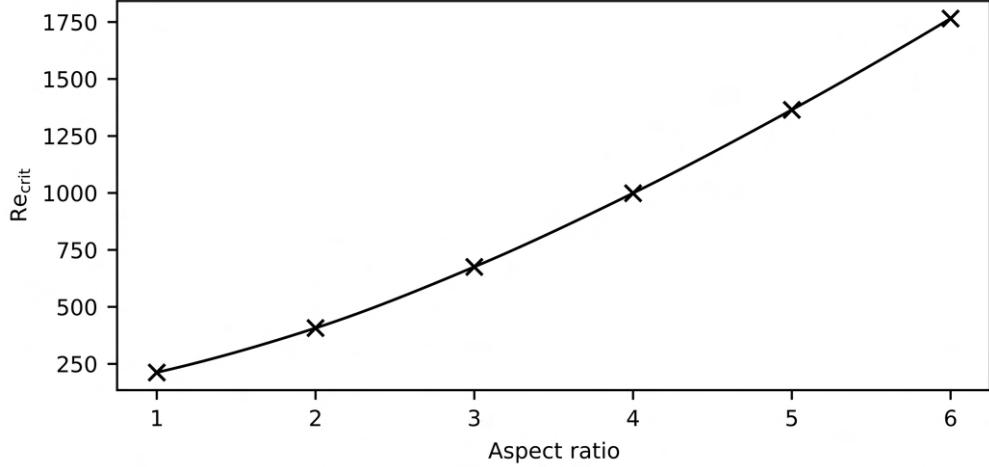


Figure 5.3: Critical Reynolds number as a function of the aspect ratio at zero angle of attack.

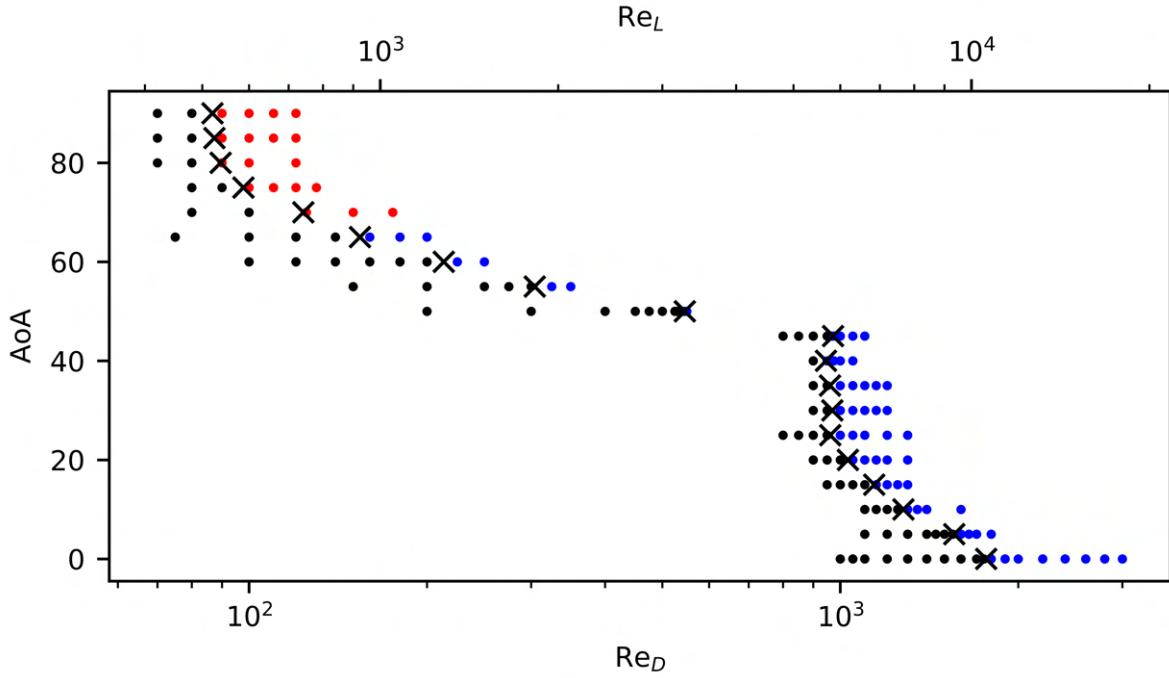


Figure 5.4: Stability chart of the 6:1 spheroid as functions of the Reynolds number and the angle of attack. The dots denote parameter combinations at which the stability calculation was carried out. The black dots correspond to a stable flow; blue and red dots mark stationary asymmetric and oscillatory shedding modes, respectively. The \times symbols denote the marginally stable states.

A number of interesting trends stand out from the figure. As the angle of attack is increased from zero, the critical Reynolds number of the stationary asymmetric leading mode continuously decreases until it reaches 25° ; then, it stays around $Re_D \approx 950$ in the interval $\alpha \in [25, 45]$. When the angle of attack is increased beyond 45° , the critical Reynolds number starts to decrease drastically, while the leading eigenmode remains stationary and asymmetric. However, at $\alpha = 70^\circ$, the most unstable eigenmode becomes an oscillatory shedding mode, and no stationary asymmetric mode is observed in the eigenspectrum. Finally, increasing the angle of attack past 80° does not influence the critical Reynolds number in a meaningful manner: its value stays around $Re_D \approx 90$.

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

| Angle of attack | 0 | 5 | 10 | 15 | 20 | 25 | 30 |
|-----------------|----------|---------|---------|---------|---------|---------|---------|
| Re_D | 1765.28 | 1558.95 | 1278.74 | 1141.12 | 1030.27 | 961.71 | 969.30 |
| Re_L | 10591.71 | 9353.72 | 7672.46 | 6846.70 | 6181.61 | 5770.29 | 5815.82 |
| Angle of attack | 35 | 40 | 45 | 50 | 55 | 60 | 65 |
| Re_D | 960.71 | 945.76 | 972.38 | 545.93 | 304.30 | 213.48 | 154.01 |
| Re_L | 5764.27 | 5674.54 | 5834.29 | 3275.55 | 1825.78 | 1280.90 | 924.08 |

Table 5.3: Diameter- (Re_D) and length-based (Re_L) critical Reynolds numbers of the first bifurcation of the spheroid flow at various angles of attack. Stationary asymmetric mode.

| Angle of attack | 70 | 75 | 80 | 85 | 90 |
|-----------------|--------|--------|--------|--------|--------|
| Re_D | 123.55 | 97.83 | 89.52 | 87.40 | 86.73 |
| Re_L | 741.31 | 586.97 | 537.14 | 524.39 | 520.38 |

Table 5.4: Diameter- (Re_D) and length-based (Re_L) critical Reynolds numbers of the first of the first bifurcation of the spheroid flow at various angles of attack. Oscillatory shedding mode.

The calculated critical Re values match the observations of previous studies. In the case of a 6:1 spheroid at an angle of attack of 45° , Andersson et al. (2019) reported that in the DNS, the flow around the spheroid was perfectly symmetric at $Re_D = 800$, while at $Re_D = 1000$ a modest asymmetry was observed, meaning that the asymmetry in the flow field could be discerned only starting from $x/D = 8$ downstream of the spheroid (the origin of their coordinate system is at the center of the spheroid, and the x direction aligns with the undisturbed free stream). The fact that their observed asymmetry was weak implies that their simulation is close to the neutrally stable Reynolds number, which matches well with the present prediction of $Re_D = 972.46$.

In the case of the 90° spheroid, Khouri et al. (2012) performed DNS at low Re values. They reported a symmetric flow field at $Re_D = 75$ while at $Re_D = 100$ the flow field was oscillatory, which agrees well with our prediction of the bifurcation at $Re_D = 86.73$. Nidhan et al. (2025) found using LES that at 10° angle of attack and $Re_D = 5000$ ($Re_L = 30000$ in their paper) the mean flow field was asymmetric. The Reynolds number of Nidhan et al. (2025) is substantially higher than the one in this work at $\alpha = 10^\circ$, so the agreement between the two studies cannot be considered as a strong verification; nevertheless the lack of disagreement still serves as a reassurance of the correctness of the presently utilized modeling framework.

In the following, the skin friction lines on the surface of the spheroid, the base flow, and the base flow plus the perturbation, as illustrations of the total flow field, are used to characterize the changes in the flow field around the spheroid. It is important to note that although the other velocity components are not shown for conciseness, both the base flow and the perturbation are strongly three-dimensional. For each angle of attack, a case near the critical Reynolds number is analyzed to describe the flow field near the neutrally stable conditions.

The zero angle of attack case is not shown, as it shares the characteristics of the sphere flow which is well documented: an axisymmetric, circular separation pattern can be observed in the skin friction lines, which reattaches at $y = z = 0$; the base flow is composed of a shear layer that detaches from the body and reattaches at the aft of the spheroid, forming a recirculation bubble, and the perturbations displace the flow to one side.

At $\alpha = 5^\circ$, the topology of the flow changes compared to the zero angle of attack case, which

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

is apparent by examining the skin friction lines that are depicted in Fig. 5.5. In the skin friction line figures, the x axis is the axial direction, while on the y axis the azimuthal angle φ is presented; since the base flow is symmetric, only angles in the $\varphi \in [0, \pi]$ interval are shown. A separation line (Se_1) that is different from the axisymmetric zero angle of attack case is clearly visible at the rear part of the spheroid. The separation is fed mainly by the upstream flow, but it is also connected to the recirculation of the wake (Re_w). The two separation lines at the two sides of the spheroid are connected at $\varphi = \pi$ by the saddle point Sa_1 being their unstable directions, and the stable directions of Sa_1 are the upstream inflow and the wake reverse flow.

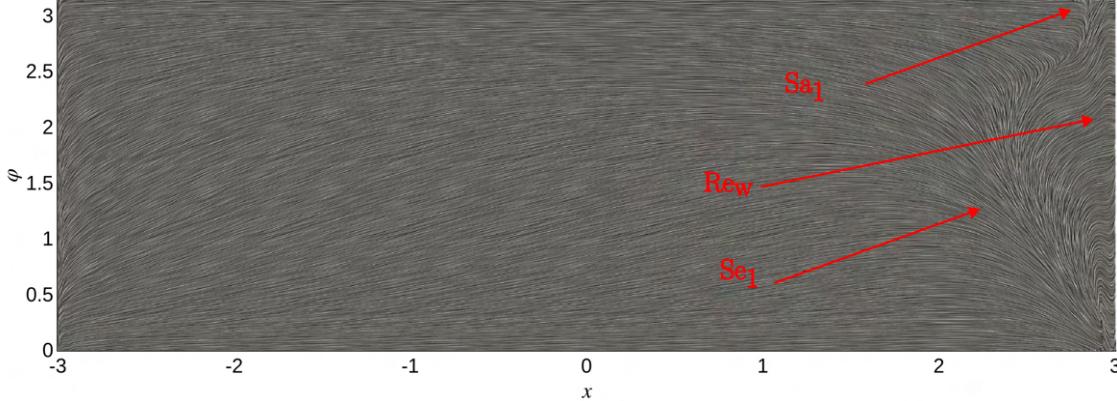


Figure 5.5: Skin friction lines of the base flow on the spheroid 6:1 spheroid at an angle of attack of 5° near the first bifurcation at $Re = 1550$.

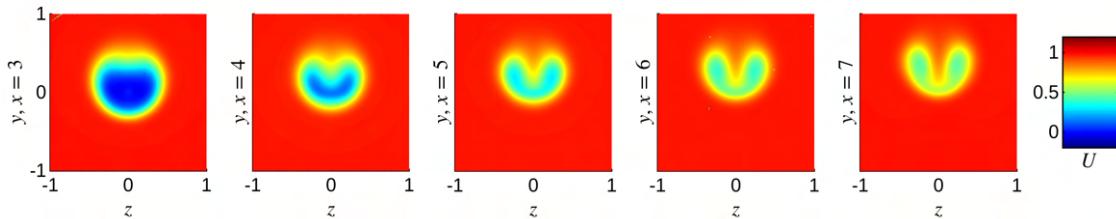


Figure 5.6: The axial component of the base flow at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 5° near the first bifurcation at $Re = 1550$.

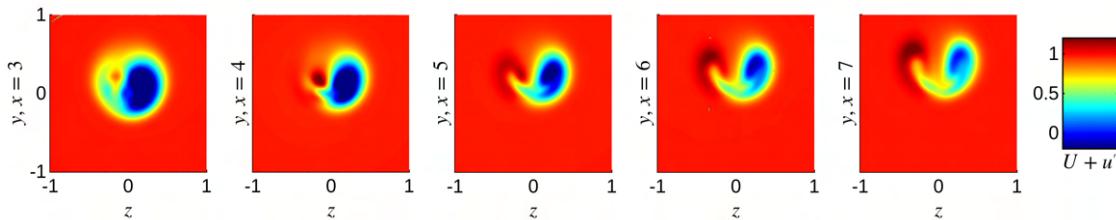


Figure 5.7: The axial component of the base flow plus perturbation at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 5° near the first bifurcation at $Re = 1550$.

Although the separation pattern is located at the aft of the body, it already results in vortices rolling up downstream of the spheroid, which is clearly highlighted by the axial component of the base flow presented at successive axial stations in Fig. 5.6. Initially, the velocity profile resembles

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

that of a wake; then, moving downstream, the vortex rollup becomes distinguishable. A similar trend can be observed in the total velocity field shown in Fig. 5.7, which is obtained by adding the direct eigenmode with a prescribed amplitude to illustrate the change in the structure of the flow. Initially, the flow field looks like a wake mode, while further downstream it resembles the well-documented vortex axisymmetry of slender body flows, in which one vortex becomes stronger and “overcomes” the other. The modes in Fig. 5.7 are compared to the mean streamwise velocity measurement of Ashok (2014) and Ashok et al. (2015), who investigated the flow field around the DARPA SUBOFF. The experimental results are presented in Fig. 5.11. Despite the three orders of magnitude difference in the Reynolds number and the dissimilarities in the configurations, the qualitative agreement between the stability analysis and the measurement is striking. The agreement strongly suggests that the asymmetric flow field in the experiments might also be caused by some kind of imperfection triggering this stationary global mode.

Next, the case with an angle of attack $\alpha = 10^\circ$ is discussed. The skin friction lines are presented in Fig. 5.8. Compared to $\alpha = 5^\circ$, significant changes can be observed in the topology of the skin friction lines. Although the saddle point Sa_1 , separation region Se_1 , and recirculation region Re_w can all be identified as at $\alpha = 5^\circ$, two new flow features emerge. Below the saddle point Sa_1 , a stable focus Fo can be identified, which corresponds to a specific flow separation that may only arise in three-dimensional flows. This type of separation has previously been observed in the experimental investigation of 2:1, 3:1, and 4:1 aspect ratio spheroids by Wang et al. (1990). Another new feature is the saddle point Sa_2 , which connects the focus Fo and the separation line Se_1 .

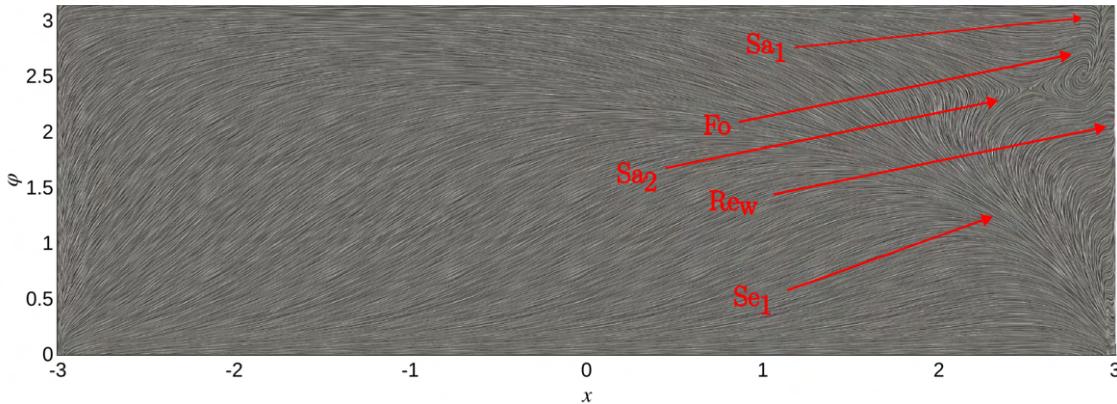


Figure 5.8: Skin friction lines of the base flow on the spheroid 6:1 spheroid at an angle of attack of 10° near the first bifurcation at $Re = 1300$.

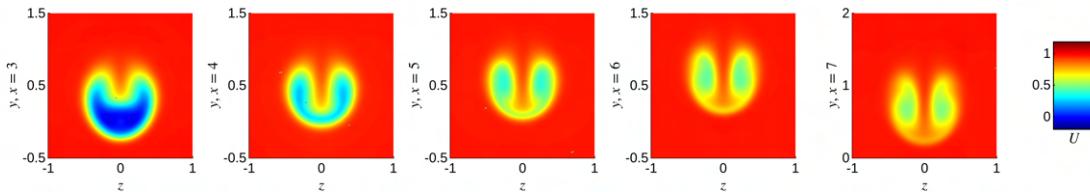


Figure 5.9: The axial component of the base flow at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 10° near the first bifurcation at $Re = 1300$.

The base flow presented in Fig. 5.9 shows that a pronounced vortex rollup can already be observed in the aft of the spheroid at $x = 3$, and the vortices become even more distinguishable as they evolve downstream. Regarding the perturbed flow field depicted in Fig. 5.10, similar observations can be drawn as in the $\alpha = 5^\circ$ case. The structure of the flow field is similar to

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

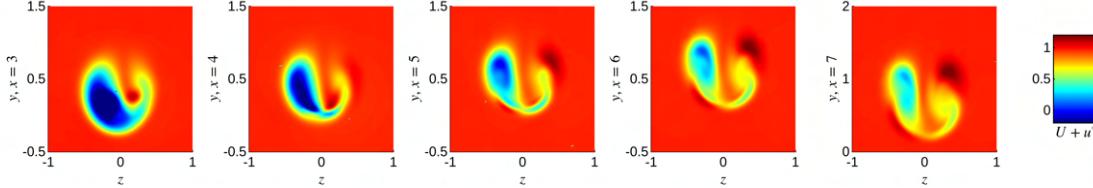


Figure 5.10: The axial component of the base flow plus perturbation at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 10° near the first bifurcation at $\text{Re} = 1300$.

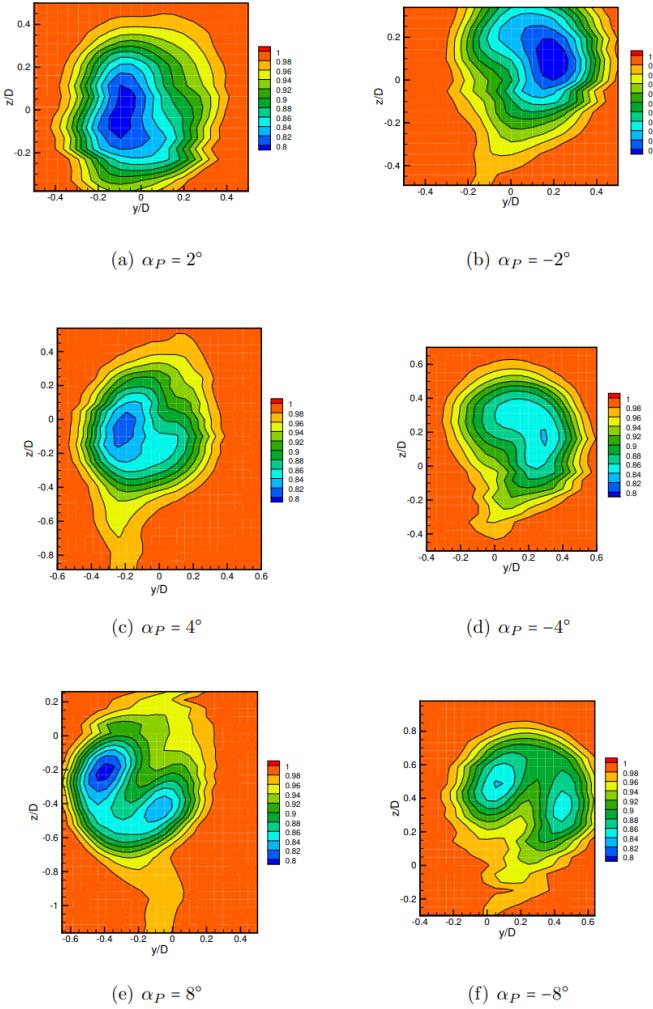


Figure 5.11: Mean streamwise velocity in the wake of high Reynolds number DARPA SUBOFF experiments (Ashok, 2014, Fig. 4.3. Reproduced with permission).

the vortex asymmetry described in the case of slender bodies, and it is similar to the asymmetry reported by Ashok (2014) and Ashok et al. (2015); Fig. 4.3 of Ashok (2014) is presented in Fig. 5.11. The fact that the total flow field is similar to the ones measured by Ashok (2014) and Ashok et al. (2015) suggests that, despite the three orders of magnitude difference in the Reynolds number and the dissimilarities configurations (DARPA SUBOFF versus spheroid, differences in the angles of attack), the same mode was triggered by an imperfection in the experiment that was found solving the eigenvalue problem. We hypothesize that the weak in-plane velocity in the y direction (Ashok

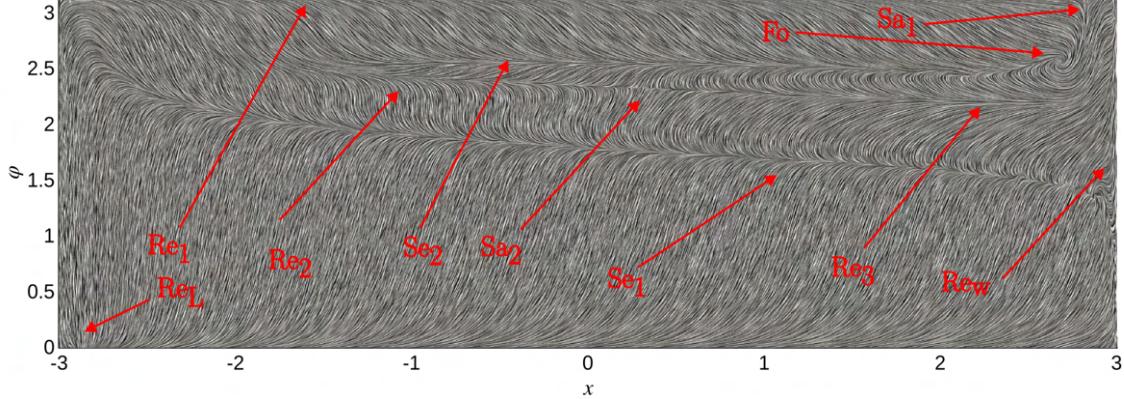


Figure 5.12: Skin friction lines of the base flow on the spheroid 6:1 spheroid at an angle of attack of 45° near the first bifurcation at $Re = 975$.

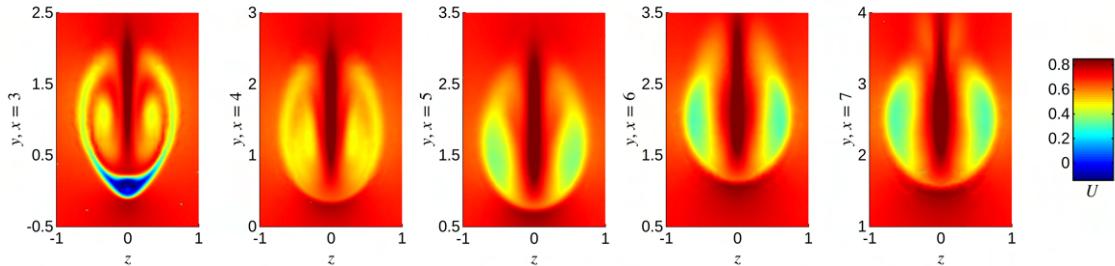


Figure 5.13: The axial component of the base flow at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 45° near the first bifurcation at $Re = 975$.

et al., 2015, Fig. 8) is responsible for exciting the eigenmode in the experiments, since it is an effect with a clear directionality that can potentially steer the flow into one of the two possible asymmetric states. Another piece of information that supports this hypothesis is that when the pitch of the configuration was varied from positive to negative angles (Fig. 5.11), the orientation of the asymmetry changed: altering the sign of α corresponds to a change in the direction of the side flow relative to the body. In addition, it is important to note that despite the topological changes in the skin friction lines and separation patterns, the perturbed flow field is similar in the $\alpha = 5^\circ$ and $\alpha = 10^\circ$ cases, implying a continuous change in the flow field with the variation of α and meaning that the same mode is observed in the two cases.

The change in the flow structure is continuous between angles of attack $\alpha = 10^\circ$ and $\alpha = 45^\circ$; thus, for the sake of brevity, only $\alpha = 45^\circ$ is presented. The skin friction lines are shown in Fig 5.12. Although the structure of the flow is more complicated than at $\alpha = 10^\circ$, most of the key flow features are the same in the two cases: the saddle points Sa_1 and Sa_2 , the focus Fo , the separation line Se_1 , and the recirculation region Re_w are all clearly identifiable. However, there are additional flow regions that become distinguishable at a higher α . At $\varphi = \pi$, the flow reattaches and becomes the main reattachment line Re_1 . In addition to the main separation line Se_1 , a secondary separation line Se_2 can be observed. Furthermore, the unstable directions of the saddle point Sa_2 become two, clearly distinguishable reattachment lines: upstream, the secondary reattachment line Re_2 , and downstream the reattachment line Re_3 which is a region of reverse flow, stemming from the recirculating wake in the aft of the spheroid. Finally, a region of reverse flow Re_L is apparent near the tip of the spheroid, which is a consequence of the fact that the stagnation point moved from near the tip downstream.

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

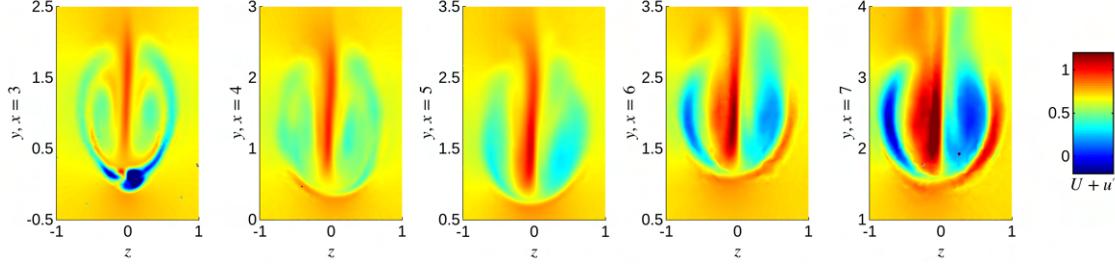


Figure 5.14: The axial component of the base flow plus perturbation at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 45° near the first bifurcation at $\text{Re} = 975$.

In the axial component of the base flow, which is presented in Fig 5.13, the two counter-rotating vortices can be clearly seen. Initially, as the flow passes the spheroid, the vortices on each side exhibit a complicated structure having two patches, but as the flow evolves downstream, the two patches merge into a single coherent vortex. The accelerated flow at the attachment line at $z = 0$ can also be clearly observed. In the total flow field, displayed in Fig 5.14, the asymmetry is clearly apparent, and the flow structure is similar to the one reported by Jiang et al. (2014).

The case with $\alpha = 50^\circ$ is omitted for the sake of brevity and $\alpha = 55^\circ$ is discussed to illustrate the changes in the flow structure when the angle of attack is further increased. At $\alpha = 55^\circ$, the critical Reynolds number is significantly lower than at $\alpha = 45^\circ$: it is $\text{Re}_{\text{crit}} = 304.30$, while at $\alpha = 45^\circ$ it is $\text{Re} = 972.38$. Note that the critical Reynolds number starts to decrease at $\alpha = 50^\circ$. The decrease in the Reynolds number with the increase in the angle of attack indicates the strong susceptibility of the flow to asymmetry. The destabilization of the flow with the angle of attack in the case of the spheroid aligns with the observations made in the case of slender bodies: when the angle of attack is increased, beyond a certain angle of attack, the asymmetry becomes much more pronounced (Bridges, 2006).

The skin friction lines are presented in Fig. 5.15. The continued increase in the angle of attack simplifies the topology of the flow. The secondary separation and reattachment lines (Re_2 , Re_3 and Se_2), which are relatively close at $\alpha = 45^\circ$, are no longer present: as the angle of attack increases, they come closer together, then merge, eliminating themselves and also destroying the saddle point Sa_2 and focus Fo . The gradual coalescence of the secondary separation and reattachment lines is clearly visible at $\alpha = 50^\circ$, but not shown for the sake of conciseness. As a result of the merging process, only the main separation line Se_1 and the reattachment line at Re_1 at $\varphi = \pi$ remain, along with the saddle point Sa_1 and the reverse flow region Re_L near the tip of the spheroid. The region of reverse flow Re_w can also be observed: it is located near where previously the secondary separation and reattachment lines were present at a lower α value, and it connects the separation and reattachment lines (Se_1 and Re_1).

Regarding the rest of the flow features, the axial base flow velocity and the axial total velocity presented in Figs. 5.16 and 5.17, respectively, most of the flow features are the same as in the $\alpha = 45^\circ$ case. However, there are also differences in the flow field. The vortex pair occupies a much larger region in the y direction, and the structure of the vortex is also simpler (c.f. $x = 3$ in Figs. 5.13 and 5.16), which is likely connected to the simplification of the separation pattern that is apparent in the skin friction lines. Examination of the total flow field suggests that the same mode can be observed at $\alpha = 45^\circ$ and $\alpha = 55^\circ$. Although, because of the difference in the color scales, the similarity may not be apparent at first glance, comparison with $\alpha = 50^\circ$ (not shown) supports that the modes are the same. So, despite the remarkable changes in the separation pattern, our analysis shows that the same asymmetric mode is present in both configurations.

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

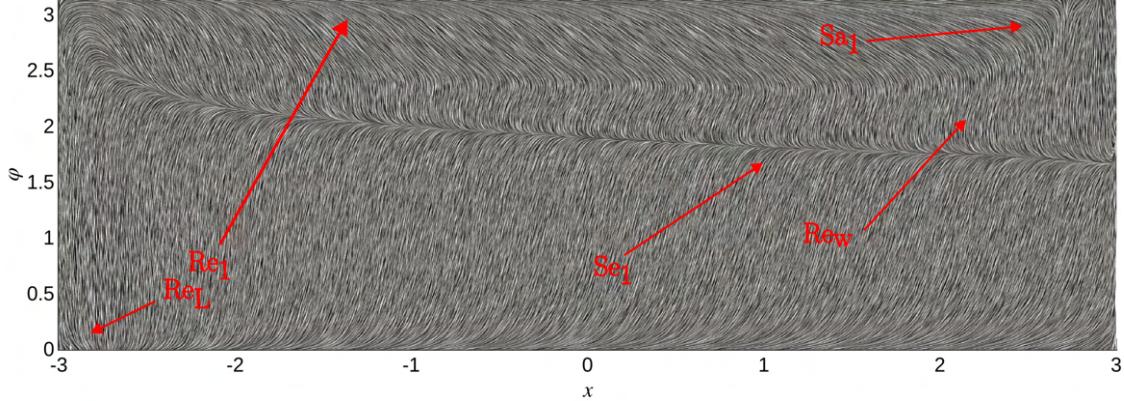


Figure 5.15: Skin friction lines of the base flow on the spheroid 6:1 spheroid at an angle of attack of 55° near the first bifurcation at $Re = 300$.

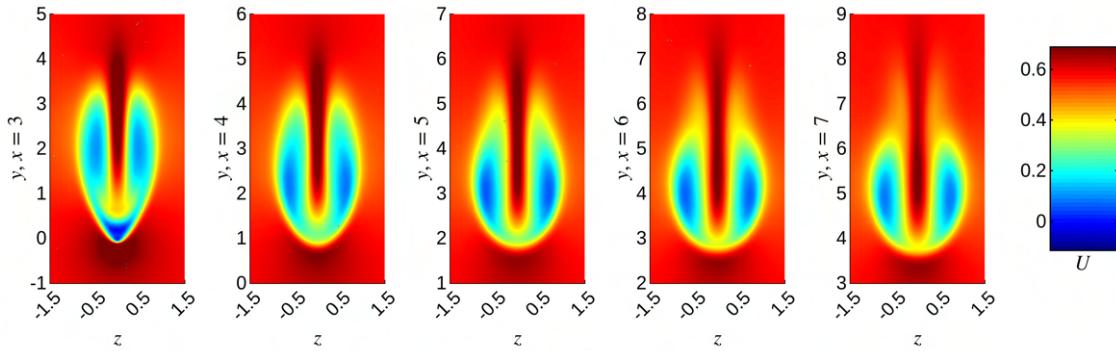


Figure 5.16: The axial component of the base flow at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 55° near the first bifurcation at $Re = 300$.

Next we consider $\alpha = 65^\circ$ and $\alpha = 70^\circ$, focusing on the discrepancies between the two cases to unveil why the most unstable mode is a stationary asymmetric and an oscillatory shedding mode in the two cases, respectively. First, the case $\alpha = 65^\circ$ is analyzed, where the critical Reynolds number drops further to $Re = 154.01$. The skin friction lines are presented in Fig. 5.18. The primary separation line Se_1 , the reattachment line Re_1 , the reverse flow region Re_w on the leeward side, the reverse flow Re_L near the tip, and the saddle point Sa_1 can all be identified as before. However, a new feature, a source So can be found near $x \approx 2.4$, which marks a closed recirculation region, and is a characteristic of wakes of bluff bodies.

There are further signs that the flow around the spheroid starts to become more wake-like. The saddle point Sa_1 moves upwards along the axis of the spheroid, thus the recirculation occupies a larger portion of the surface. In addition, both in the case of the attachment line at $\varphi = 0$ and the reattachment of the flow (Re_1), the flow starts to get more aligned with the azimuthal direction instead of the axial one. This suggests that the evolution of the flow cannot be solely interpreted by examining successive axial locations: the azimuthal direction tends to dictate the dynamics more than the axial one. Furthermore, the extent of the reattachment in the axial direction (Re_1) is shorter, which implies that although the flow above the spheroid dives behind the spheroid, which plays a role in the rollup of the vortices, this effect is less pronounced than in the case of lower angles of attack.

The fact that the flow starts to exhibit wake-like behavior is also apparent in the plots of the

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

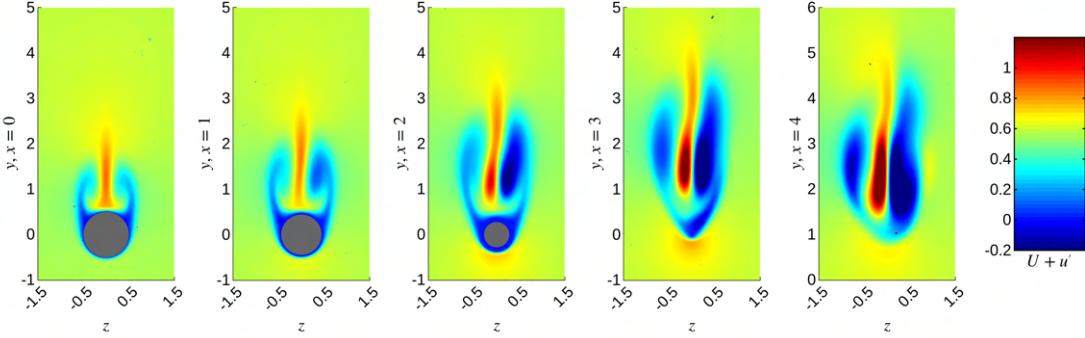


Figure 5.17: The axial component of the base flow plus perturbation at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 55° near the first bifurcation at $\text{Re} = 300$.

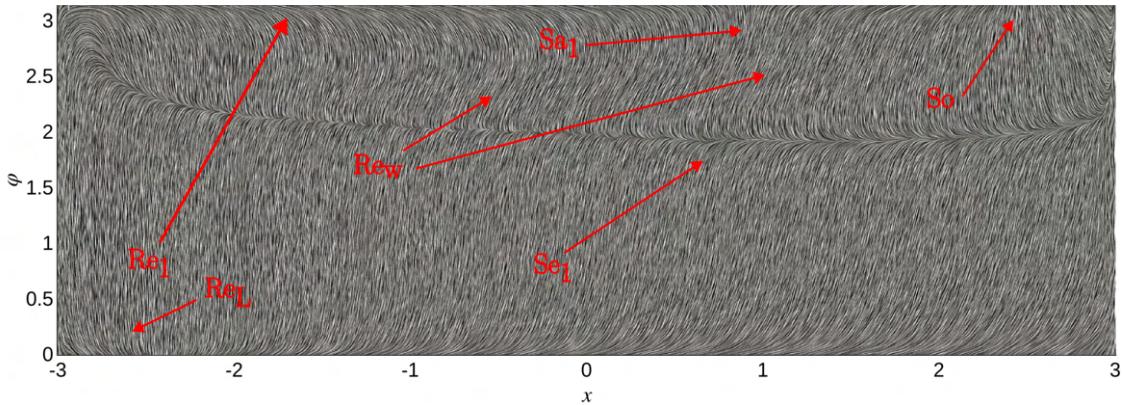


Figure 5.18: Skin friction lines of the base flow on the spheroid 6:1 spheroid at an angle of attack of 65° near the first bifurcation at $\text{Re} = 160$.

base flow presented in Fig. 5.19. The vortex rollup in shear layers is less pronounced than at $\alpha = 55^\circ$ (Fig. 5.16), and flow starts to be reminiscent of the wake behind a cylinder at low Re . The difference between the vortex- and wake-like regions can be observed between the axial stations $x < 2$ and $x \geq 2$: in these two flow regions, which are separated by the saddle point Sa_1 the skin friction lines in Fig. 5.18, display reattachment with axial flow in the x direction and wake-like reattachment with reverse flow, respectively.

Despite the fact that the flow field displays wake-like characteristics which imply a shedding-type instability, the eigenmode that first becomes unstable is a stationary asymmetric one, which is apparent by examining Fig. 5.20. The main features of the total flow field remain similar to $\alpha = 55^\circ$: the flow is displaced to one side, and along the axis the amplitude of the disturbance becomes more pronounced.

Next, the skin friction lines for the $\alpha = 70^\circ$ case are presented in Fig. 5.21. Almost all the main flow features are the same as in the case of $\alpha = 65^\circ$; yet, there is one major difference which explains why in this case the leading eigenmode is a shedding mode. The saddle point Sa_1 moves backwards along the x axis, and it is very close to the tip of the spheroid. Thus, there is almost no vortex rollup and reattachment that was still present at $\alpha = 65^\circ$ (Re_1 in Fig 5.18). Therefore, the flow behaves like a bluff body wake, and a sufficiently high Re starts to display vortex shedding. The total flow field of the axial flow presented in Figs. 5.23 clearly illustrates the shedding behavior by showing the well-known characteristics of an oscillatory wake mode.

The conclusion drawn from the comparison of cases with $\alpha = 65^\circ$ and $\alpha = 70^\circ$ is the following

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

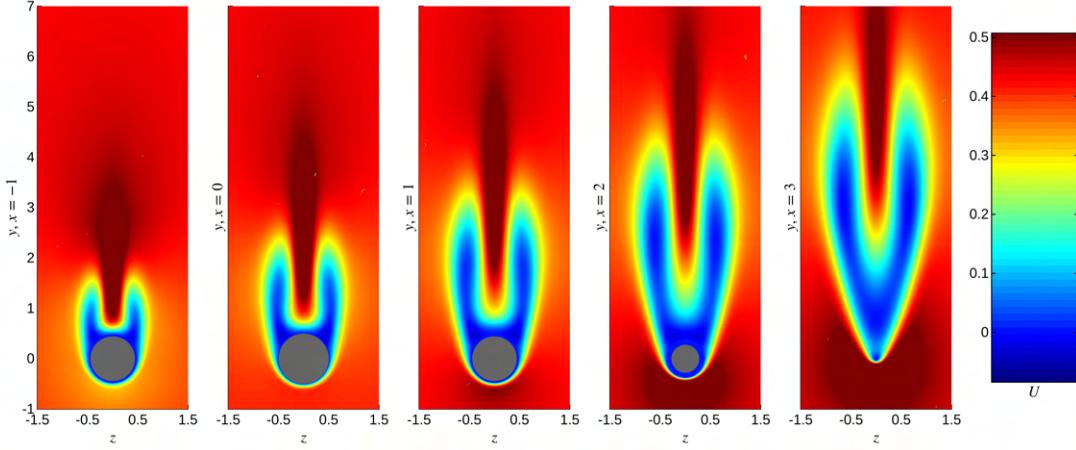


Figure 5.19: The axial component of the base flow at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 65° near the first bifurcation at $\text{Re} = 160$.

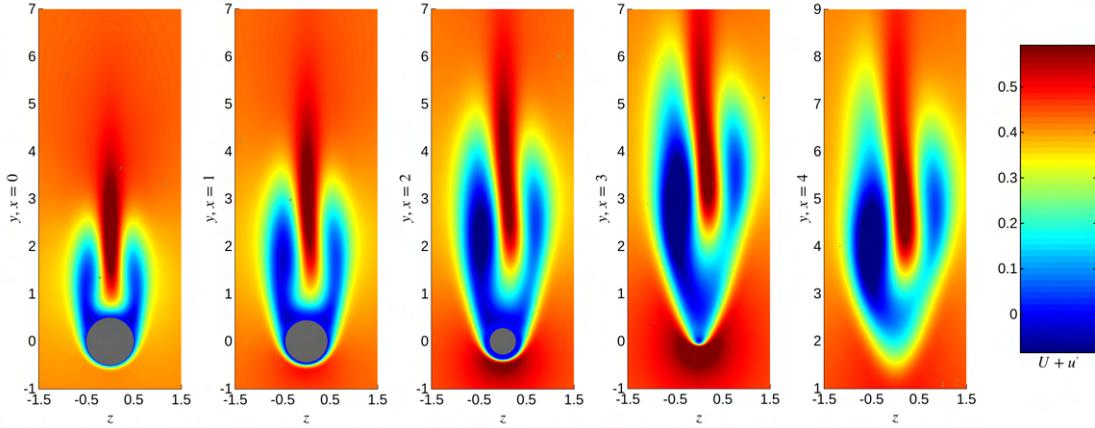


Figure 5.20: The axial component of the base flow plus perturbation at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 65° near the first bifurcation at $\text{Re} = 160$.

conjecture: the reason for the stationary asymmetry is that coherent vortices are sufficiently close to each other, since the lack of vortex rollup is the major difference between cases with $\alpha = 65^\circ$ and $\alpha = 70^\circ$. This, in essence, supports the argument of Keener and Chapman (1977), who suggested that the vortices need to be “crowded” together for the asymmetry to be able to develop. Reflecting on the low α cases, an axisymmetric wake can be thought of as a vortex ring at the aft of the rigid body, which is generated by the flow that surrounds the body. Thus, this conjecture does not contradict the changes in the flow topology at low α values; on the other hand, the low α flow topology supports this conjecture.

Increasing the angle of attack further past $\alpha = 70^\circ$ does not induce significant changes in the characteristics of the flow; therefore, for the sake of conciseness, no figures are shown, and only the main trends are reported without displaying the flow field. Continuing to increase the angle of attack slightly lowers the critical Reynolds number (see Table 5.4), but the changes are minor. The structure of the base flow and the modes, consequently, is also subject to only small changes: whereas at lower angles of attack, the modes are localized higher axial coordinates, they gradually move backward along the x axis with the increase of α until the core of the dynamics is located around $x = 0$.

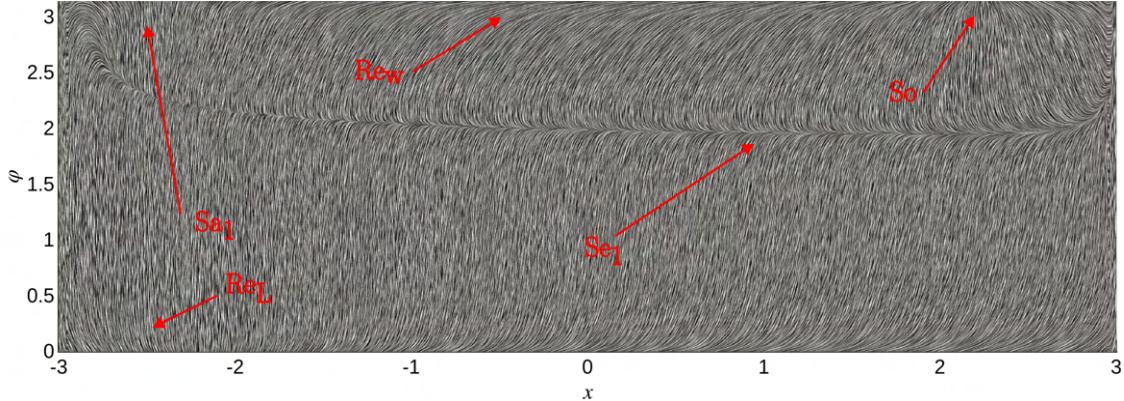


Figure 5.21: Skin friction lines of the base flow on the spheroid 6:1 spheroid at an angle of attack of 70° .

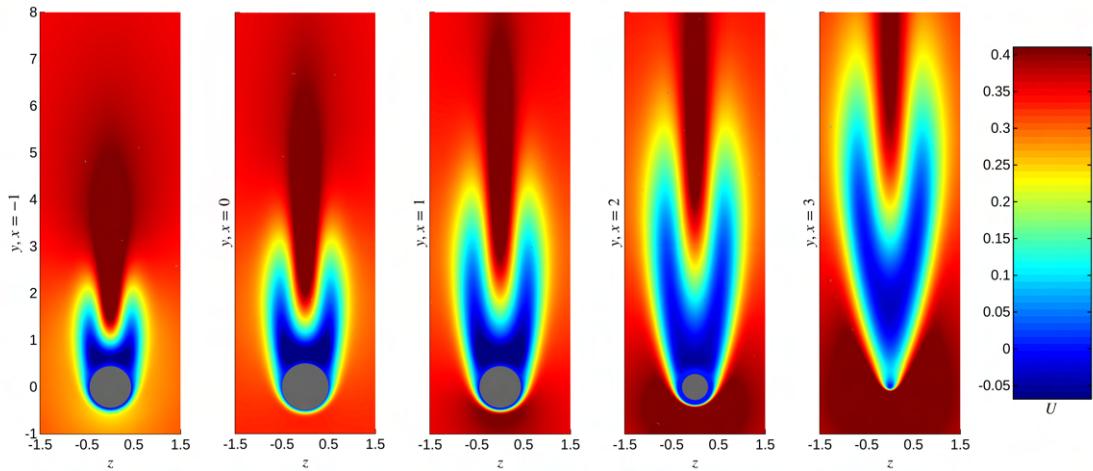


Figure 5.22: The axial component of the base flow at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 70° near the first bifurcation at $\text{Re} = 125$.

5.4 Conclusion

This study is concerned with characterizing the first bifurcation of the flow around the 6:1 spheroid at low Reynolds numbers while the angle of attack is varied to gain a deeper understanding of how flow asymmetries arise in symmetric flow configurations. An efficient modeling framework was established using the matrix forming technique, the Newton-Raphson method to find the equilibrium base flow, and the shift-invert technique combined with the Krylov-Schur method to calculate the relevant part of the eigenspectrum. The application of the inverse operators in both cases is carried out using a sparse LU factorization library. A new set of codes was developed using the FreeFEM finite element library, and the implementation was verified in the case of the lid-driven cavity, channel flow, and the flow around the sphere and 4:1 spheroid at zero angle of attack.

First, the aspect ratio of the spheroid was varied from 1:1 to 6:1 at zero angle of attack, and it was found that the leading eigenmode is a stationary asymmetric mode, which shares the features of axisymmetric wakes. The critical Reynolds number of the leading mode increases monotonically with the aspect ratio, which is linked to the more streamlined shape of the body and weaker re-circulation. Then, the effect of changing the angle of attack is investigated in the case of the 6:1 spheroid, and the stability map of the flow is constructed. Below $\alpha = 70$, the first unstable eigen-

Investigation of the Asymmetric Flow Field around the Prolate Spheroid using Global Stability Analysis

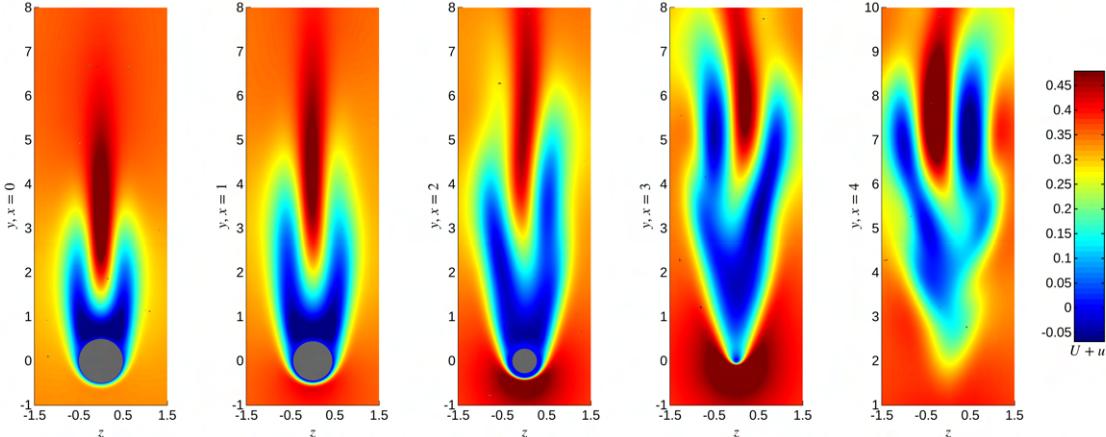


Figure 5.23: The axial component of the base flow plus perturbation at $x = \text{const.}$ cross-sections. 6:1 spheroid at an angle of attack of 70° near the first bifurcation at $\text{Re} = 125$.

mode is a stationary asymmetric mode; for $\alpha \geq 70^\circ$, the leading instability is an oscillatory shedding mode. The observations that beyond a certain angle of attack, vortex shedding is preferred instead of a stationary asymmetry, and at a certain angle of attack, the flow becomes more susceptible to asymmetry, which is reflected by a drop in the critical Reynolds number, shows that the flow around the spheroid shares the characteristics of the flow around other slender bodies (Bridges, 2006).

At various angles of attack, the flow fields are analyzed near the critical Reynolds number by examining the skin friction lines to characterize the separation pattern, the axial component of the base flow, and the axial component of the total velocity field (base flow plus the leading mode with a prescribed amplitude). It is found that despite the changes in the flow field as the angle of attack is increased — from an axisymmetric wake to vortices rolling up with a complicated separation pattern, whose structure becomes simpler as the angle of attack is increased — a large number of flow fields can support asymmetry. In addition, as the angle of attack is changed, the structure of the asymmetry changes continuously, implying that the same mode, resulting from the same mechanism, is observed. The common feature of the flow fields is that when long time-scale asymmetry is observed, stationary vortices with opposing signs can interact. When vortex shedding is observed, the separating flow does not reattach to the body, so no concentrated vortex patches are formed, and the flow looks like a cylinder wake. At low α , the axisymmetric wake can be thought of as a vortex ring attached to the aft of the body, which makes the $\alpha = 0$ limit consistent with the observation that asymmetry is related to the interaction of vortices with opposing signs. The observation regarding the relation of the vortex rollup to asymmetry supports the hypothesis of Keener and Chapman (1977) that the vortices need to be closely “crowded” for the asymmetry to develop.

The total flow field obtained by adding the perturbation to the base flow looks remarkably similar to the asymmetry in the experiments of Ashok et al. (2015) (also Ashok (2014)), despite the difference in the geometry and the three orders of magnitude difference in the Reynolds number. We conjecture that the weak in-plane flow in one direction is responsible for triggering the asymmetric mode, since varying the sign of the pitch changed the orientation of the asymmetry in the experiments.

References

E. Åkervik, L. Brandt, D. S. Henningson, J. Høpffner, O. Marxen, and P. Schlatter. Steady solutions of the Navier-Stokes equations by selective frequency damping. *Physics of Fluids*, 18:068102, 2006.

H. I. Andersson, F. Jiang, and V. L. Okulov. Instabilities in the wake of an inclined prolate spheroid. In A. Gelfgat, editor, *Computational Modelling of Bifurcations and Instabilities in Fluid Dynamics*, pages 311–352. Springer International Publishing, 2019.

A. Ashok. The turbulent wake of submarine model in pitch and yaw, 2014. Doctoral Dissertation, Princeton University, NJ.

A. Ashok, T. V. Buren, and A. Smits. Asymmetries in the wake of a submarine model in pitch. *Journal of Fluid Mechanics*, 774:416–442, 2015.

S. Balay, S. Abhyankar, M. F. Adams, S. Benson, J. Brown, P. Brune, K. Buschelman, E. Constantinescu, L. Dalcin, A. Dener, V. Eijkhout, W. D. Gropp, V. Hapla, T. Isaac, P. Jolivet, D. Karpeev, D. Kaushik, M. G. Knepley, F. Kong, S. Kruger, D. A. May, L. C. McInnes, R. T. Mills, L. Mitchell, T. Munson, J. E. Roman, K. Rupp, P. Sanan, J. Sarich, B. F. Smith, S. Zampini, H. Zhang, H. Zhang, and J. Zhang. PETSc/TAO users manual. Technical Report ANL-21/39 - Revision 3.16, Argonne National Laboratory, 2021.

D. Bridges. The asymmetric vortex wake problem - asking the right question. In *36th AIAA Fluid Dynamics Conference and Exhibit*, San Francisco, CA, 2006. AIAA 2006-3553.

A. Chiarini, R. Gauthier, and E. Boujo. On the symmetry-breaking instability of the flow past axisymmetric bluff bodies. *Journal of Fluid Mechanics*, 1016:A52, 2025.

J.-M. Chomaz. Global instabilities in spatially developing flows: Non-normality and nonlinearity. *Annual Review of Fluid Mechanics*, 37:357–392, 2005.

V. Citro, J. Tchoufag, D. Fabre, F. Giannetti, and P. Luchini. Linear stability and weakly nonlinear analysis of the flow past rotating spheres. *Journal of Fluid Mechanics*, 807:62–86, 2016.

V. Citro, P. Luchini, F. Giannetti, and F. Auteri. Efficient stabilization and acceleration of numerical simulation of fluid flows by residual recombination. *Journal of Computational Physics*, 344:234–246, 2017.

X. Y. Deng, W. Tian, B. F. Ma, and Y. K. Wang. Recent progress on the study of asymmetric vortex flow over slender bodies. *Acta Mechanica Sinica*, 24:475–487, 2008.

C. Geuzaine and J.-F. Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.

F. Giannetti, P. Luchini, and L. Marino. Linear stability analysis of three-dimensional lid-driven cavity flow. In *Atti del XIX Congresso AIMETA di Meccanica Teorica e Applicata*, pages 7381–73810. Aras Edizioni, 2009.

F. Gómez, R. Gvmez, and V. Theofilis. On three-dimensional global linear instability analysis of flows with standard aerodynamics codes. *Aerospace Science and Technology*, 32:223–234, 2014.

M. Grandemange, O. Cadot, and M. Gohlke. Reflectional symmetry breaking of the separated flow over three-dimensional bluff bodies. *Physical Review E*, 86:35302, 2012.

M. Grandemange, M. Gohlke, and O. Cadot. Turbulent wake past a three-dimensional blunt body. Part 1. Global modes and bi-stability. *Journal of Fluid Mechanics*, 722:51–84, 2013.

F. Hecht. New development in freefem++. *Journal of Numerical Mathematics*, 20(3-4):251–266, 2012.

F. Jiang, J. P. Gallardo, and H. I. Andersson. The laminar wake behind a 6:1 prolate spheroid at 45° incidence angle. *Physics of Fluids*, 26:113602, 2014.

F. Jiang, J. P. Gallardo, H. I. Andersson, and Z. Zhang. The transitional wake behind an inclined prolate spheroid. *Physics of Fluids*, 27(9):093602, 2015.

References

F. Jiang, H. I. Andersson, J. P. Gallardo, and V. L. Okulov. On the peculiar structure of a helical wake vortex behind an inclined prolate spheroid. *Journal of Fluid Mechanics*, 801:1–12, 2016.

M. P. Juniper, A. Hanifi, and V. Theofilis. Modal stability theory: Lecture notes from the FLOW-NORDITA Summer School on Advanced Instability Methods for Complex Flows, Stockholm, Sweden, 2013. *Applied Mechanics Reviews*, 66, 2014.

E. R. Keener and G. T. Chapman. Similarity in vortex asymmetries over slender bodies and wings. *AIAA Journal*, 15:1370–1372, 1977.

G. K. E. Khoury, H. I. Andersson, and B. Pettersen. Wakes behind a prolate spheroid in crossflow. *Journal of Fluid Mechanics*, 701:98–136, 2012.

J. Magnaudet and G. Mougin. Wake instability of a fixed spheroidal bubble. *Journal of Fluid Mechanics*, 572:311–337, 2007.

P. Meliga, J. M. Chomaz, and D. Sipp. Unsteadiness in the wake of disks and spheres: Instability, receptivity and control using direct and adjoint global stability analyses. *Journal of Fluids and Structures*, 25:601–616, May 2009.

J. Moulin, P. Jolivet, and O. Marquet. Augmented Lagrangian preconditioner for large-scale hydrodynamic stability analysis. *Computer Methods in Applied Mechanics and Engineering*, 351:718–743, 2019.

R. Natarajan and A. Acrivos. The instability of the steady flow past spheres and disks. *Journal of Fluid Mechanics*, 254:323–344, 1993.

S. Nidhan, S. Jain, J. L. Ortiz-Tarin, and S. Sarkar. Stratified wake of a 6:1 prolate spheroid at a moderate pitch angle. *Journal of Fluid Mechanics*, 1009:A58, 2025.

M. Plasseraud and K. Mahesh. Vortex topology in the lee of a 6:1 prolate spheroid. *arXiv preprint arXiv:2507.03187*, 2025.

M. Plasseraud, P. Kumar, and K. Mahesh. Large-eddy simulation of tripping effects on the flow over a 6:1 prolate spheroid at angle of attack. *Journal of Fluid Mechanics*, 960:A3, 2023.

M. A. Regan and K. Mahesh. Global linear stability analysis of jets in cross-flow. *Journal of Fluid Mechanics*, 828:812–836, 2017.

J. H. M. Ribeiro, C.-A. Yeh, and K. Taira. Triglobal resolvent analysis of swept-wing wakes. *Journal of Fluid Mechanics*, 954:A42, 2023.

G. Rigas, A. Oxlade, A. Morgans, and J. Morrison. Low-dimensional dynamics of a turbulent axisymmetric wake. *Journal of Fluid Mechanics*, 755:R5, 2014.

G. Rigas, A. S. Morgans, R. D. Brackston, and J. F. Morrison. Diffusive dynamics and stochastic models of turbulent axisymmetric wakes. *Journal of Fluid Mechanics*, 778:R2, 2015.

G. Rigas, L. Esclapez, and L. Magri. Symmetry breaking in 3D wakes. In *Proceedings of the Summer Program*. Center for Turbulence Research, Stanford University, 2016.

L. V. Rolandi, J. H. M. Ribeiro, C. A. Yeh, et al. An invitation to resolvent analysis. *Theoretical and Computational Fluid Dynamics*, 38:603–639, 2024.

J. E. Roman, C. Campos, L. Dalcin, E. Romero, and A. Tomas. SLEPc users manual. Technical Report DSIC-II/24/02 - Revision 3.18, D. Sistemes Informàtics i Computació, Universitat Politècnica de València, 2022.

H. Strandenes, F. Jiang, B. Pettersen, and H. I. Andersson. Near-wake of an inclined 6:1 spheroid at Reynolds number 4000. *AIAA Journal*, 57:1364–1372, 2019a.

H. Strandenes, F. Jiang, B. Pettersen, and H. I. Andersson. Low-frequency oscillations in flow past an inclined prolate spheroid. *International Journal of Heat and Fluid Flow*, 78:108421, 2019b.

A. Tezuka and K. Suzuki. Three-dimensional global linear stability analysis of flow around a spheroid. *AIAA Journal*, 44:1697–1708, 2006.

References

V. Theofilis. Advances in global linear instability analysis of nonparallel and three-dimensional flows. *Progress in Aerospace Sciences*, 39:249–315, 2003.

K. C. Wang, H. C. Zhou, C. H. Hu, S. Harrington, and M. J. Lighthill. Three-dimensional separated flow structure over prolate spheroids. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 429: 73–90, 1990.

O. J. Williams and A. J. Smits. Asymmetries in nominally symmetric flows. *Annual Review of Fluid Mechanics*, 57: 35–60, 2025.

T. Zhu, G. Rigas, and J. F. Morrison. Near wake coherent structures of a turbulent axisymmetric bluff body wake. *International Journal of Heat and Fluid Flow*, 112:109668, 2025.

Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow

LOKESH SILWAL, RODRIGO VILUMBRALES-GARCIA, & ANCHAL SAREEN

This study investigates the impact of surface indentations shaped as dimples on the flow dynamics of a pitching foil under zero-freestream conditions. A series of systematic experiments were conducted employing flow field measurements using Particle Image Velocimetry. The dimple depth ratio (d/D , where d is the dimple depth and D is the dimple diameter) was varied from 0.022 to 0.088 across Reynolds numbers ($Re = V_{TE_{max}} c / \nu$, where $V_{TE_{max}}$ is the maximum trailing edge velocity, c is the foil chord, and ν is the fluid kinematic viscosity) of 3700, 10000 and 20000. The impact of dimples on the wake characteristics was evaluated by analyzing the time-averaged jet behavior and vortex dynamics. The results reveal that the deepest dimpled case modified the far wake of the pitching foil, particularly at higher Reynolds numbers. Under these conditions, the vortices shed from the trailing edge persisted longer, and the jet exhibited greater coherence. The dimples appear to influence dipole interactions in the wake, reducing the jet deflection. These findings suggest that surface roughness can be strategically employed to modulate wake dynamics and improve the stability of the jet, potentially enhancing the propulsion efficiency of bio-inspired flapping foil systems.

Keywords: flow-structure interaction, vortex dynamics, swimming/flying

6.1 Introduction

Flapping foils have been shown to be efficient propulsors and when their kinematics are properly tuned, large amounts of thrust can be achieved with high efficiencies, comparable to rotary turbines (Anderson et al., 1998). The main parameter that governs the propulsive efficiency and wake characteristics of a flapping foil is the Strouhal number ($St = Af/U_\infty$), where A is the characteristic width of the jet flow created, f is the oscillation frequency, and U_∞ is the free-stream flow velocity. Previous work has shown that the most efficient propulsion is achieved for values of St between 0.25 and 0.4 for heave amplitude to chord ratios of order one and the maximum pitching angles between 15° to 25° (Anderson et al., 1998; Lagopoulos et al., 2019; Mackowski and Williamson, 2015).

In this range of St , leading edge vortices (LEVs) with moderate strength are formed that interact with trailing edge vortices (TEVs) to form a reverse von Kármán street. The phase angle between the heave and pitch motions is the critical parameter affecting the timing of formation and shedding of LEVs. These flow features add momentum in the wake, in contrast to the drag producing von Kármán street, which creates a momentum deficit. As St increases beyond 0.4, the propulsion efficiency drops, and the wake structures no longer exhibit the reverse von Kármán street. For $St > 0.5$, the foil produces four vortices per cycle, with two dipoles of counter-rotating sign (Anderson et al., 1998; Godoy-Diana et al., 2008). The formation of the dipole has been shown

Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow

to induce wake deflection, which leads to a loss of efficiency caused by a lower thrust generation (Godoy-Diana et al., 2008, 2009; Gursul and Cleaver, 2019; Zhong and Quinn, 2024). At $St = \infty$, corresponding to a zero-free-stream condition, wake instabilities, such as deflection, can be expected to intensify. Shinde and Arakeri (2013) studied a pitching foil in quiescent flow and observed a wake characterized by random meandering and a lack of coherence. These wake characteristics resulted from destructive interactions between closely-spaced trailing edge vortices due to the absence of freestream convection. Consequently, the strength of the jet was reduced, leading to reduced propulsive performance of the foil (Shinde and Arakeri, 2018).

Several strategies have been developed to limit wake deflection and meandering at $St = \infty$, all incorporating foil flexibility. Previous studies (Shinde and Arakeri, 2013, 2014, 2018; Heathcote and Gursul, 2007; Heathcote et al., 2008) have shown that incorporating a flexible trailing-edge flap can significantly alter the wake dynamics. The higher motion amplitude, resulting from the flap's flexibility, prevents vortex clustering, reducing destructive interactions and suppressing jet meandering. This promotes the formation of a reverse von Kármán street, improving thrust production. By analyzing various flap stiffnesses, the authors (Shinde and Arakeri, 2014) found that optimizing flap stiffness could enhance mean thrust by up to 50%.

The effect of surface texture modifications on flapping foils has also been explored as a strategy to enhance thrust production and propulsion efficiency. Oeffner and Lauder (2012) tested short-fin mako shark skin samples on both rigid and flexible flapping plates. They found that while shark denticles reduced the propulsive effectiveness of the rigid plate, they increased the swimming speed of the flexible plate by 12%. However, it remains unclear whether the kinematics were consistent across test cases. Later, Wen et al. (2014) examined an undulating plate covered with 3D-printed denticles scaled 100 times larger than natural denticles and reported a 6.6% increase in efficiency. Again, the consistency of kinematics between smooth and rough surfaces was not clearly stated. Another study by Guo et al. (2021) investigated the effect of biomimetic shark skin in dynamic flow separation conditions (maneuvers) and found that local roughness may mitigate boundary layer separation through instantaneous modification of the local pressure-gradient distribution. For the accelerating shark skin, a substantial decrease in the size of the shear layer and remarkable differences in the near-wake velocity distribution were observed due to the combination of acceleration and the presence of shark skin.

To investigate the interaction between kinematics and surface roughness shaped as an egg carton, Massey et al. (2023) conducted high-resolution simulations of a self-propelled swimming plate at a transitional Reynolds number of $Re = 12000$. By matching the kinematics of rough and smooth plates, they found that roughness can reduce the required swimming power and decrease the unsteady amplitude of forces. Their results highlighted a nonlinear interaction between roughness and kinematics, demonstrating that findings from static roughness studies do not directly translate to unsteady swimmers. This idea was further reinforced by Vilumbrales-Garcia et al. (2024), who examined the effect of generic semi-hemispherical bumps on purely pitching foils. They found that unsteady flapping systems exhibit greater robustness to surface roughness variations, reporting no performance improvements for rough pitching foils compared to smooth foils. Additionally, they observed that protruding roughness increased form drag, resulting in a drag penalty rather than a propulsive benefit. These studies on the effect of surface protrusions on flapping foils and plates underscore the importance of optimizing the shape and size of roughness elements to achieve meaningful performance gains. While surface roughness has shown potential for enhancing propulsion efficiency, its effectiveness depends on careful tuning of the roughness elements to foil kinematics and flow conditions. Furthermore, whether these benefits extend to quiescent flow remains an open question.

Surface indentations shaped as dimples have also shown effectiveness in improving aerodynamic

and hydrodynamic performance for flat plates and static airfoils. For example, van Nesselrooij et al. (2016) observed a 4% drag reduction for a flat plate with dimples arranged in a staggered pattern. In that study, the dimple depth-to-diameter ratio was 2.5%, with a coverage area ratio of 33%, and the most significant drag reduction occurred at a Reynolds number of 40,000. Similarly, Ali et al. (2024) investigated a static airfoil (NACA 65(3) – 218) with three rows of staggered dimples placed at 15%, 50% & 85% of the chord length. The Reynolds number based on the chord was 250000. The study found that dimples reduced drag by up to 6.6% without negatively affecting lift. This improvement was attributed to the dimples energizing the boundary layer, which delayed flow separation and reduced pressure drag. In addition to improving aerodynamic performance, dimples have been shown to influence small-scale vortical structures and reduce acoustic signatures in static foils (Ananthan et al., 2022; Kumar et al., 2024). These results suggest that dimples can modify boundary layer behavior and vortex shedding characteristics in unsteady flows, such as those encountered in flapping foil configurations. Despite this, the effects of dimples on the wake dynamics and propulsive performance of flapping foils have yet to be explored. Furthermore, whether these benefits extend to quiescent flow remains an open question. To the best of the authors' knowledge, there are no reported studies on the effect of surface roughness on flapping foils under quiescent flow conditions. This study aims to address this gap by systematically investigating the influence of surface roughness—in the form of dimples—on the wake dynamics of a purely pitching foil in a quiescent environment. For this purpose, we selected surface indentations rather than protrusions, and chose dimples over more intricate geometries like shark scales, owing to their simplicity in design and manufacturing. Dimples also allow for precise control of roughness through variations in depth. The key unanswered questions are:

- How do dimples influence the wake characteristics of a pitching foil in quiescent flow?
- What is the effect of the Reynolds number (based on maximum trailing edge velocity)?
- How does the dimple depth to diameter ratio impact the flow?

In this study, we conducted systematic experiments to address the aforementioned questions. Section 6.2 details the experimental methods and operating conditions, while Section 6.3 presents and discusses the results. Finally, Section 6.4 summarizes our findings and draws conclusions.

6.2 Experimental methods and operating conditions

The experiments were carried out in a water tank with dimensions of $(1.22 \times 0.53 \times 0.32) \text{ m}^3$. A brief schematic of the experimental setup is shown in Figure 6.1. The pitching motion of the foil was driven by a StepSERVO motor (Applied motion TSM17P-3AG NEMA 17). The motion conformity with the design parameters (trailing edge amplitude and motion frequency) was verified using a rotary optical incremental encoder (E5 US Digital) mounted on the transmission shaft.

A NACA0015 foil with a chord of $c = 0.09 \text{ m}$ and a span of $s = 0.135 \text{ m}$ was used as the test model (see Figure 6.1). Nominal two-dimensional flow characteristics were enforced by enclosing the wing between the bottom of the tank and a surface plate, ensuring a gap of less than $2\%c$ similar to previous studies (Fuglsang et al., 1998; Shinde and Arakeri, 2014; Vilumbrales-Garcia et al., 2024). These boundary conditions are widely used to suppress spanwise flow and tip vortices, effectively minimizing three-dimensional effects in flapping foil experiments (Quinn et al., 2014; Wong and Rival, 2015). The planar PIV measurements were performed at the mid-span of the foil to capture

Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow

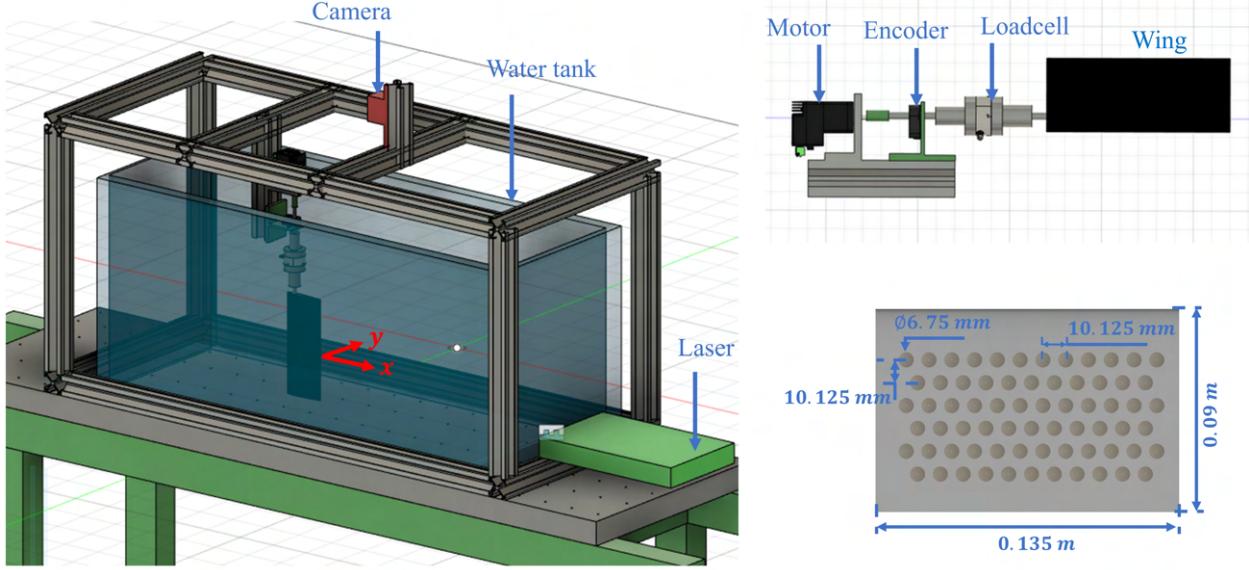


Figure 6.1: A brief schematic of the experimental setup used in the current study. The figure on the top right shows the foil actuation assembly and the one on the bottom right shows foil dimensions and details of the dimples employed.

the core dynamics of the wake under these controlled conditions. For the smooth foil, we observed no significant signs of large-scale three-dimensional flow features, supporting the assumption of a nominally two-dimensional wake. It is important to point out that the introduction of dimples could induce localized three-dimensionality, such as streamwise vortices or spanwise pressure gradients near the surface. However, it is emphasized that based on the symmetry of the foil and motion, and the lack of irregularities in the planar wake fields observed, any three-dimensional effects induced by the dimples is expected to be small in scale and localized, without compromising the interpretation of the mid-span dynamics.

With current dimensions, the resulting side wall clearance was 8 times the maximum trailing edge excursion ($\Delta_{TE_{max}}$), the leading edge of the wing was 8 times $\Delta_{TE_{max}}$ from the front wall and 26 times $\Delta_{TE_{max}}$ from the back wall of the hydrodynamic tank. These dimensions are on par with previous studies conducting pitching foil studies in a quiescent tank facility (Shinde and Arakeri, 2013, 2014).

All four foils were 3D printed with a stereolithography (SLA) resin 3D printer (Form 3, Form-Labs) with a resolution of $25\mu\text{m}$. The first foil, serving as the control case, had a smooth surface without indentations. For the other foils, the dimple depth was varied as $d = 0.15 \text{ mm}$, $d = 0.3 \text{ mm}$, and $d = 0.6 \text{ mm}$, while the dimple diameter remained constant at $D = 6.75 \text{ mm}$. The dimples were arranged in a staggered pattern for all cases, since it has been shown that the staggered pattern offers higher drag reduction in flat plate studies (Veldhuis and Vervoort, 2009; van Nesselrooij et al., 2016). The dimple depth-to-diameter ratios of $d/D = 0.022$, $d/D = 0.044$, and $d/D = 0.088$, were chosen with a total surface coverage of 20.27%. This range of dimple depths covers the recommended values from both the flat plate studies (van Nesselrooij et al., 2016) and the static airfoil studies (Ali et al., 2024). The coverage ratio is inspired from the flat plate studies (van Nesselrooij et al., 2016).

The foil undergoes sinusoidal pitching motion about a pivot point located at the quarter-chord from the leading edge. The motion is described as $\theta = \theta_{max} \sin(2\pi ft)$ with a fixed maximum pitch amplitude of $\theta_{max} = 15^\circ$. The frequency f is varied at 0.37, 1 and 2 Hz to change the Reynolds

| Dimple depth ratio (d/D) | Re | Coverage ratio | θ_{\max} |
|------------------------------|--------------------|----------------|-----------------|
| 0.022 | 3700, 10000, 20000 | 20.27% | 15° |
| 0.044 | 3700, 10000, 20000 | 20.27% | 15° |
| 0.088 | 3700, 10000, 20000 | 20.27% | 15° |

Table 6.1: Summary of operating conditions. Here, d is the dimple depth, D is the dimple diameter, Re is the Reynolds number based on the trailing edge velocity, and θ_{\max} is the maximum pitch amplitude.

number (Re). Since this study is conducted under quiescent flow conditions, the Reynolds number is defined based on the maximum trailing edge velocity following previous studies (Shinde and Arakeri, 2013):

$$Re = V_{TE_{max}} c / \nu$$

where ν is the fluid kinematic viscosity and c is the chord length. The maximum trailing edge velocity is calculated as $V_{TE_{max}} = \theta_{max} 2\pi f c$. The resulting Re are 3700, 10000 & 20000, which are selected based on the operational limitations of the current experimental facility. Flow measurements were conducted using 2D-2C Particle Image Velocimetry (PIV) for all test conditions. Each test was repeated four times to ensure repeatability. The confidence intervals in the figures represent the standard deviation (σ) of the mean across the four trials. The summary of these operating conditions is provided in Table 6.1.

Flow measurements

The flow field was captured using planar Particle Image Velocimetry (PIV) in $x - y$ plane, passing through the foil's midspan. The seeding was achieved using 55 μm polyamide particles with the specific gravity of 1.01 g/cm³. The resulting Stokes number (St) for the current flow conditions is estimated to be 0.02 which is below the recommended value of $St \leq 0.05$ (Samimy and Lele, 1991). An Evergreen 200 dual-pulsed laser with a pulse energy of 200 mJ was used to illuminate the particles. A Photron NOVA R3-4K camera with a resolution of 4096 \times 2304 pixels² was used to capture the PIV images. The camera was equipped with a 60 mm Nikon lens to provide a field of view of 2.4 c \times 1.2 c . A total of 1080 image pairs were captured over 72 seconds at a sampling rate of 15 Hz for all cases. The time delay between pulses (δt) was optimized for each case to ensure accurate velocity measurements. The images were captured and processed in DaVis 11 (LaVision) software. The vector field was obtained by processing the images using the multi-grid, multi-pass cross-correlation technique with initial and final interrogation window of size 64 \times 64 pixels and 24 \times 24 pixels, respectively, with an overlap of 50%. The resulting vector field had a spatial resolution of 0.416 mm (0.0046 c). The uncertainties in the velocity measurements were quantified using the correlation statistics method (Wieneke, 2015), where the uncertainty is proportional to the residual positional disparity between the matched correlation peaks (Sciacchitano, 2019). The resulting uncertainty in the velocity across the flowfield for all cases was $\leq 1\%$ of the average velocity in the flow field.

The vector fields were post processed in MATLAB for time-averaging, phase-averaging, and vortex tracking. Phase-averaging was performed using encoder data synchronized with the PIV measurements. The vortex tracks and their circulation were quantified using the Γ_1 and Γ_2 criteria, as given by Graftieaux et al. (2001). The Γ_1 criterion was used to identify vortex centers and track their trajectories. This involves computing the non-dimensional scalar parameter (Γ_1) at any point P in the discrete vector field. The parameter is defined by the following equation:

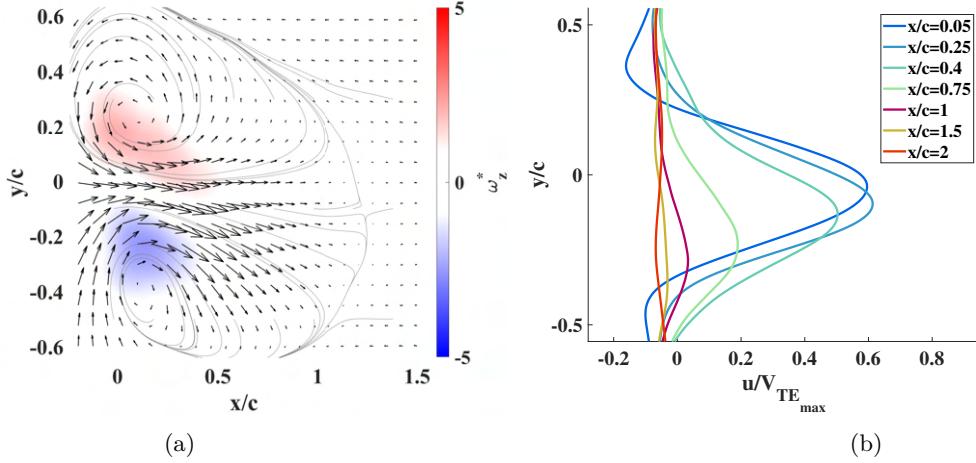


Figure 6.2: (a) Time-averaged normalized z-vorticity ($\omega_z^* = \omega c / V_{TE_{max}}$) contours overlaid with time-averaged streamlines (based on velocities in x and y direction) for the smooth foil at $Re = 3700$. (b) Normalized x-component velocity at different downstream locations for the smooth foil at $Re = 3700$.

$$\Gamma_1(P) = \frac{1}{N} \sum_s \frac{(r_{PM} \times U_M) e_z}{\|r_{PM}\| \|U_M\|} \quad (6.1)$$

where S is two-dimensional area surrounding P , M is any other point inside S , U_M is the velocity at point M , and e_z is the unit vector normal to the measurement plane.

The Γ_2 criterion was then calculated, which identifies regions where the flow is locally dominated by rotation. For Γ_2 values greater than $2/\pi$, the flow is considered to be dominated by rotation, thus defining the vortex core. The vorticity inside the vortex core is then integrated to quantify the vortex circulation. The Γ_2 criterion is given by the following equation:

$$\Gamma_2(P) = \frac{1}{N} \sum_s \frac{(r_{PM} \times (U_M - \widehat{U}_P)) e_z}{\|r_{PM}\| \|U_M - \widehat{U}_P\|} \quad (6.2)$$

where \widehat{U}_P is the local convective velocity at point P . The variations in the vortex center and the normalized circulation determined using this technique over the four repetitions were within 4% of the chord and 5% of the peak normalized circulation, respectively.

6.3 Results and discussions

First, the results of the time-averaged flow characteristics for the smooth foil control case are presented and compared with the existing literature on foil flapping under quiescent conditions. Next, the influence of the Reynolds number on the flow characteristics of the control case is analyzed. Finally, the effects of dimples on flow characteristics are examined.

Wake dynamics of a smooth pitching foil

The results for a smooth pitching foil at $Re = 3700$ are presented first, as this Reynolds number closely aligns with the conditions studied by Shinde and Arakeri (2013). The time-averaged flow field averaged over 50 cycles and further averaged across four independent runs for the smooth foil, is shown in Figure 6.4a (see 6.4 for the convergence study). Although the motion profile is symmetric, the wake exhibits asymmetry, with the jet inclined downward and an imbalance in the

vorticity of the resulting dipole, as depicted in Figure 6.2. Jet deflection is a well-documented phenomenon in flapping foils operating at $St > 1$, as reported in previous literature (Jones et al., 1998; Emblemsvag et al., 2002; Heathcote and Gursul, 2007; Shinde and Arakeri, 2013).

Jones et al. (1998) conducted numerical simulations and experiments on a purely plunging foil at high St . For $St > 1$, they observed a dual-mode deflected wake in both cases. For numerical simulations, the deflection direction was determined by the initial conditions, whereas in experiments, wake-switching occurred randomly. The authors hypothesized that even small disturbances could trigger wake-deflection switching in experimental measurements. Similar observations were made by Emblemsvag et al. (2002) in their numerical simulation of a pitching flat plate about the leading edge at $St \approx 1$, where the jet deflection direction depended on the initial pitching motion.

These jet characteristics are maximized at the limiting condition of $St = \infty$, which corresponds to quiescent flow conditions. Heathcote and Gursul (2007) conducted experimental measurements on a purely plunging foil in quiescent flow conditions and observed periodic variations in the jet deflection angle, influenced by plunging frequency, amplitude and airfoil stiffness. However, the flow characteristics of purely plunging and purely pitching motions differ. In the plunging case, the two trailing edge vortices (TEVs) shed per cycle convect as a vortex dipole and persist over longer downstream distances (Heathcote and Gursul, 2007). In contrast, in purely pitching motion, the two TEVs interact destructively in the near wake (Shinde and Arakeri, 2013). Likely due to these differences, Shinde and Arakeri (2013) observed random wake deflection switching in a purely pitching foil under quiescent conditions at $Re = 3234$.

These conditions closely resemble those of the current measurements. However, unlike Shinde and Arakeri (2013), no switching in the wake deflection angle is observed; instead, the wake consistently inclines downward. Multiple repetitions of the measurements yielded consistent results. To assess the influence of the initial conditions, the initial motion was reversed, yet the wake remained deflected downward (see 6.4). For all the repetitions, initially, the wake remained centered for the first few cycles before settling into a downward-deflected state for the remainder of the cycles, as shown in the supplementary material S1. Consequently, the time-averaged flow exhibits asymmetric characteristics. The time-averaged jet characteristics were further analyzed by extracting the streamwise velocity component (x -component) at various downstream locations, as summarized in Figure 6.2b. The velocity is normalized by the maximum trailing edge velocity ($V_{TE_{max}}$). As the jet develops, the peak velocity occurs around $x/c \approx 0.2$, then gradually diminishes at further downstream locations, reaching zero beyond $x/c > 1$. These observations are consistent with those reported by Shinde and Arakeri (2013).

Since the jet remains symmetric for the first few cycles, it suggests that the pitching motion alone does not introduce an immediate bias. However, the vortex dipole formed in the wake is inherently unstable, making it highly sensitive to even the slightest perturbations. As the dipole evolves, any minor asymmetry—whether due to experimental imperfections, background disturbances, or nonlinear interactions within the wake—can grow over time, leading to a self-reinforcing deflection. Once the dipole begins to tilt, the induced velocity from the vortices further destabilizes the symmetric configuration, favoring a consistent downward jet inclination. This process is similar to symmetry breaking in other unsteady fluid systems, where an initially balanced state is disrupted by small perturbations that grow due to the system's inherent instability. This hypothesis aligns with the observation that reversing the initial motion does not alter the deflection direction—suggesting that the instability, rather than initial conditions, dictates the final wake asymmetry.

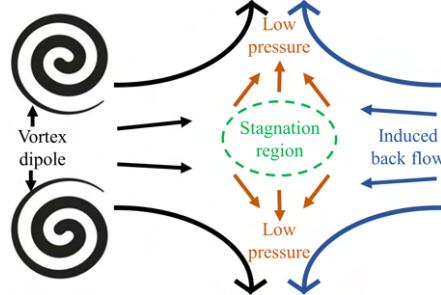


Figure 6.3: Schematic representation of the flow characteristics observed in the time-averaged wake of a smooth pitching foil under quiescent conditions.

Reynolds Number effects

The effects of Reynolds number on the flow characteristics of a smooth foil were analyzed, and the time-averaged flow fields for $Re = 3700, 10000$ & 20000 are shown in Figure 6.4. At $Re = 3700$, the wake exhibited noticeable jet asymmetry, a feature that persisted at $Re = 10000$. Interestingly, at $Re = 20000$, the wake became nearly symmetric, indicating that higher Reynolds numbers reduce the jet's sensitivity to disturbances that cause deflection at lower Re . At low Reynolds numbers of $Re = 3700$, the flow is more viscously dominated, resulting in thicker boundary layer and less coherent separation. These conditions can promote asymmetric shear layer development and unsteady vortex shedding, which can lead to imbalanced dipoles and deflected jets. Additionally, residual vortices from previous cycles persist longer in the wake due to slower convection, increasing the likelihood of dipole-wake interactions that can further exacerbate jet deflection. In contrast, at higher Reynolds numbers of $Re = 20000$, the flow becomes more momentum-dominated, boundary layers are thinner, and separation occurs more sharply and coherently. As a result, vortex shedding is more symmetric and repeatable, producing stronger but more balanced dipoles that generate a narrowly focused and stable jet. The faster convection of shed vortices at high Re also reduces cross-cycle interactions, contributing to reduced jet deflection and improved wake alignment.

Another key observation was the presence of a stagnation region downstream, located around $x/c \approx 1$, as indicated by the streamlines. This region forms due to the interaction between the jet produced by the vortex dipole and the backflow from the far wake region. The backflow results from low pressure created by fluid entrainment in the near wake by the vortex dipole (see Figure 6.3). As Re increases, the stagnation region becomes more pronounced, consistent with the expectation that stronger dipoles at higher Re generate greater backflow.

The presence of a stagnation region suggests that the jet diminishes around the same location. This was further investigated by plotting two jet properties: the peak and mean streamwise velocities in the wake. The streamwise velocities were integrated as $\int_y u dy$, where u is the streamwise jet velocity, along y -axis between $-0.6c$ and $0.6c$ at different downstream locations to extract the mean streamwise velocity. The results shown in Fig 6.5 reveal two phases of the jet: the development phase, where the velocity increases to a maximum, and the decay phase, where the velocity gradually decays to zero. The stagnation region, where velocity approaches zero, is located around $x/c \approx 1$ and remains consistent across all Reynolds numbers. However, Reynolds number significantly influences velocity magnitudes. At higher Re , the stronger dipole generates higher jet velocities, accompanied by an increase in induced backflow (see Figure 6.5b). These findings illustrate the impact of Reynolds number on the wake structures of the baseline case.

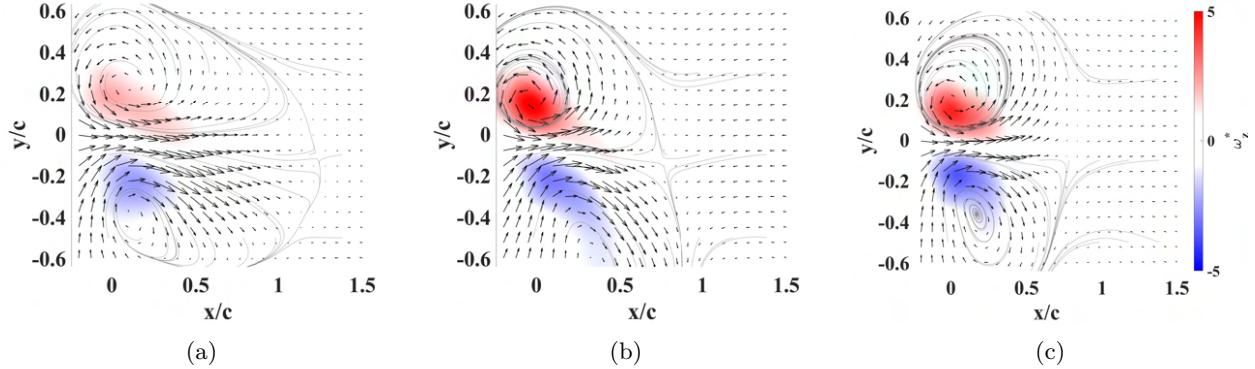


Figure 6.4: (a) Time-averaged normalized z-vorticity contours overlaid with time-averaged streamlines (based on velocities in x and y direction) for the smooth foil at (a) $Re = 3700$ (b) $Re = 10000$ and (c) $Re = 20000$.

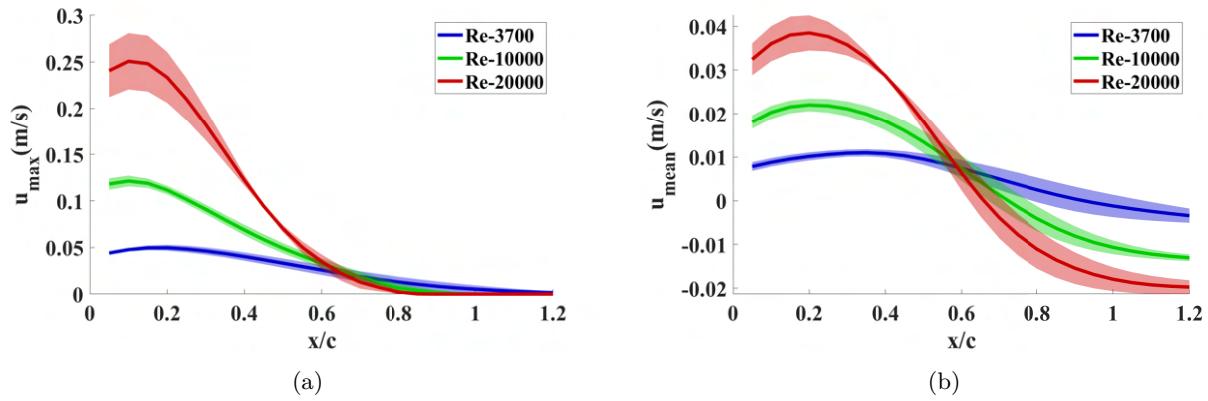


Figure 6.5: Development of the (a) peak velocity and (b) mean velocity in the smooth pitching foil wake for Reynolds number cases of $Re = 3700$, 10000 and 20000 . The shaded region represents the standard deviation of the mean over the four repetitions.

Effects of dimples

In this section, time-averaged flow field characteristics are analyzed for different dimple depth ratios and Reynolds numbers. Next, the vortex characteristics and their influence on the time-averaged flow features are discussed.

Time-averaged flow characteristics

The time-averaged flow characteristics for the three dimpled cases at all Reynolds numbers are shown in Figure 6.6. When comparing the smooth foil case (Figure 6.4a) with the dimpled cases (Figures 6.6a, 6.6d, and 6.6g) at $Re = 3700$, notable differences in the vortex dipole behavior were observed. For the dimpled cases, the dipoles were more diffused, and the jet inclination was less pronounced compared to that of the smooth foil. Additionally, the stagnation region observed in the near-wake of the smooth foil case was absent in the dimpled cases. As the Reynolds number increased to $Re = 10000$ & 20000 , these effects were only present for the case with the deepest dimple ($d/D = 0.088$), while the flow fields for the shallower dimples resembled those of the smooth foil. Specifically, for $d/D = 0.022$ and $d/D = 0.044$, the jet was more inclined, and the stagnation region appeared around $x/c \approx 1$. These observations suggest that dimples influence vortex dipole and jet characteristics, particularly for deeper dimples, as the effects persisted across all Reynolds number

Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow

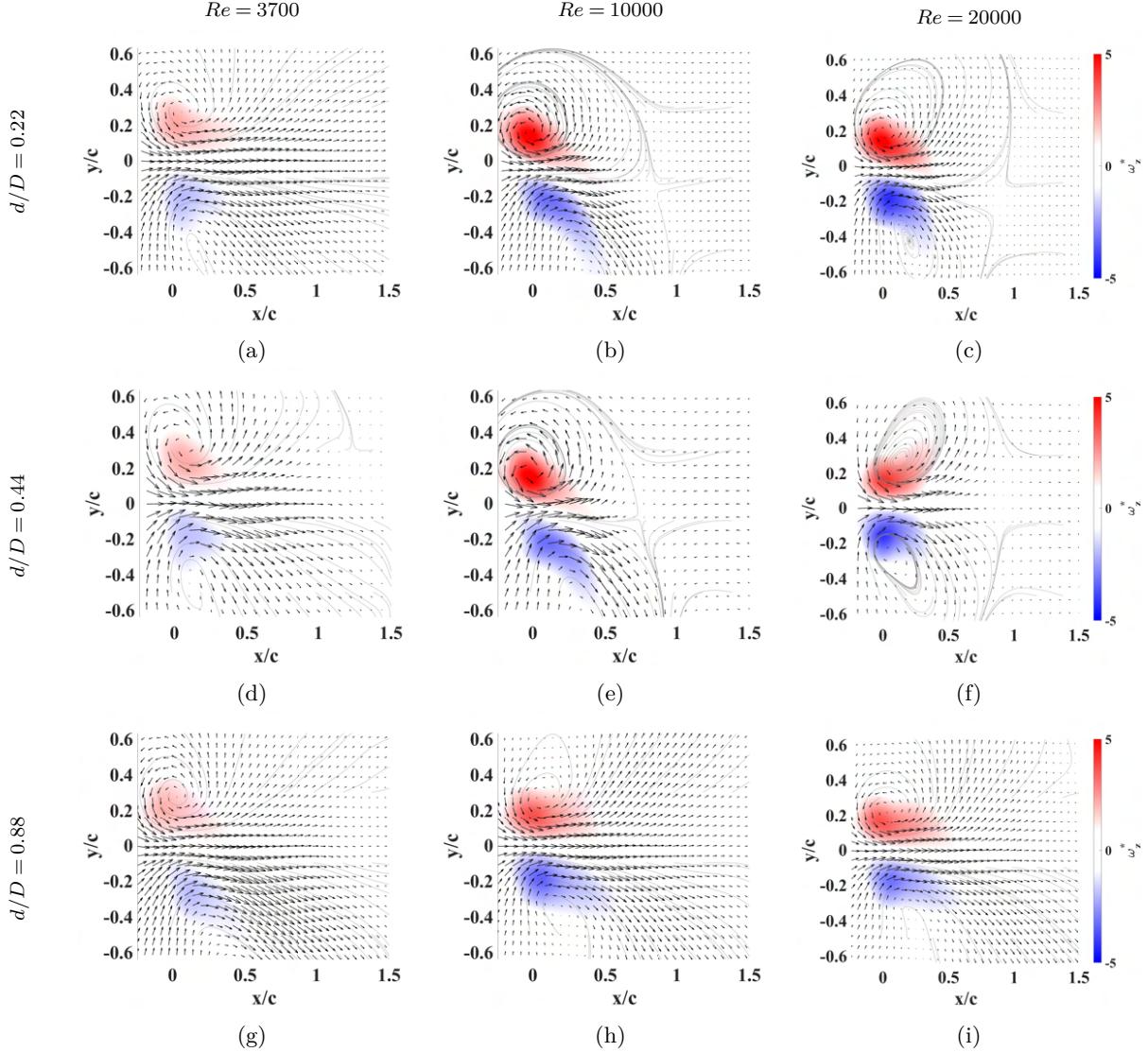


Figure 6.6: Time-averaged normalized z-vorticity contours overlaid with time-averaged streamlines (based on velocities in x and y direction) for; dimple depth of $d/D = 0.022$ at (a) $Re = 3700$, (b) $Re = 10000$, (c) $Re = 20000$; dimple depth of $d/D = 0.044$ at (d) $Re = 3700$, (e) $Re = 10000$, (f) $Re = 20000$; dimple depth of $d/D = 0.088$ at (g) $Re = 3700$, (h) $Re = 10000$, and (i) $Re = 20000$.

conditions. These effects are further quantified by extracting the peak and mean jet velocities at different downstream locations, with the results shown in Figure 6.7.

The upper row of Figure 6.7 represents the maximum streamwise jet velocity normalized by the maximum trailing edge velocity, while the lower row shows the mean streamwise jet velocity for different Reynolds numbers. At $Re = 3700$, while the qualitative flowfield characteristics between the smooth and dimpled foils were different, the differences in the quantitative properties fall within the margin of uncertainty and can not be discerned. At $Re = 10000$, the jet characteristics for the smooth foil and the dimpled cases with $d/D = 0.022$ & 0.044 were similar. However, the deepest dimple case ($d/D = 0.088$) exhibited significantly higher peak velocities during both the jet development and decay phases (as high as 50% larger than the control foil). Additionally, the jet was sustained longer for the dimpled case ($d/D = 0.088$), persisting up to $x/c = 2$, compared to

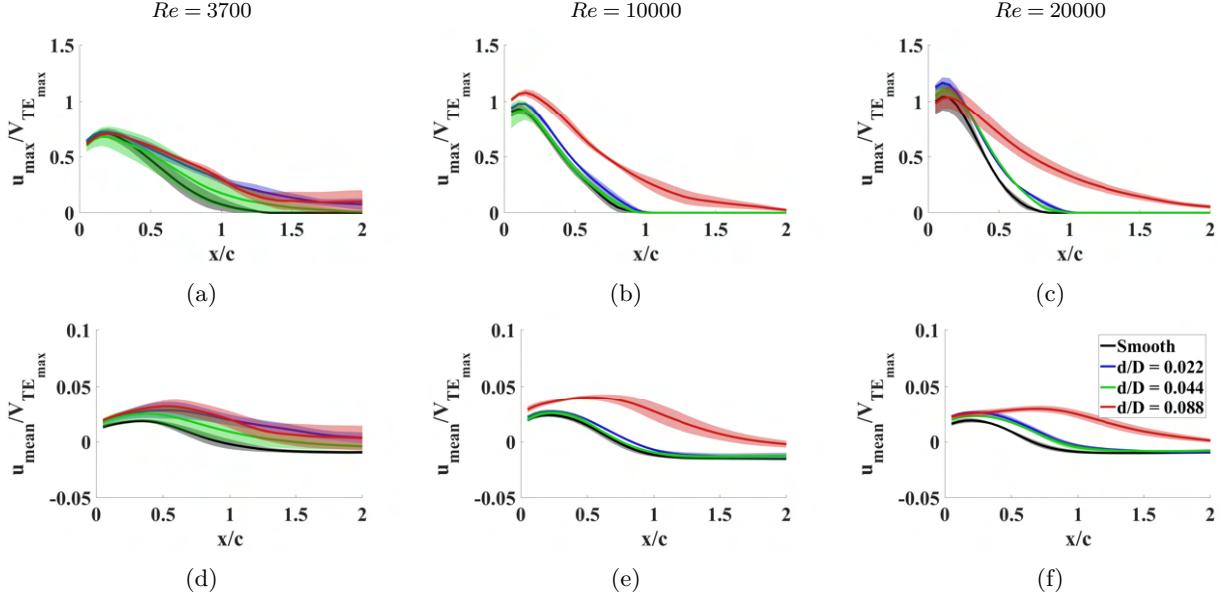


Figure 6.7: Comparisons of peak velocity evolution between smooth foil and dimpled foils with dimple depths of $d/D = 0.022, 0.044$ & 0.088 at Reynolds numbers of (a) $Re = 3700$, (b) $Re = 10000$, and (c) $Re = 20000$. Comparisons of mean velocity evolution between smooth foil and dimpled foils with dimple depths of $d/D = 0.022, 0.044$ & 0.088 at Reynolds numbers of (d) $Re = 3700$, (e) $Re = 10000$, and (f) $Re = 20000$. The shaded region represents the standard deviation of the mean over the four repetitions.

only $x/c = 1$ for the smooth foil. The differences between the smooth and dimpled ($d/D = 0.088$) foils are substantially larger than the standard deviation of repeated measurements. For example, the increase in the maximum jet velocity observed for the dimpled foils exceeded the experimental variability by a factor of 10, indicating that these trends are robust and not attributable to random fluctuations or measurement noise.

At the highest Reynolds number ($Re = 20000$), the dimpled cases with $d/D = 0.022$ & 0.044 showed higher velocities than the smooth foil during the development and decay phases, but the jet decayed more quickly, reaching zero velocities at a similar location to the smooth case. In contrast, for the deepest dimple case ($d/D = 0.088$), the jet velocity during the development phase was comparable to that of the smooth case, but the decay phase exhibited significantly higher velocities (again, as high as 50% larger than the control foil), with the jet persisting for a longer duration. These trends align with the qualitative flow characteristics shown in Figure 6.6. For example, at $Re = 10000$, the time-averaged flow fields for the smooth foil (Figure 6.4c) and the dimpled cases with $d/D = 0.022$ & 0.044 (Figs. 6.6b, 6.6e, and 6.6h) exhibited similar features, including jet inclination and the presence of a stagnation region, resulting in comparable jet characteristics. However, for the deepest dimple case ($d/D = 0.088$), the vortex dipoles appeared more diffused, jet inclination was reduced, and the near-wake stagnation region was absent, leading to higher jet velocities compared to the other cases.

Similarly, at $Re = 20000$, the stagnation region for $d/D = 0.022$ & 0.044 cases remained at approximately the same downstream location as the smooth foil, causing the jet velocity to decay to zero at a similar location. In contrast, the deepest dimpled case ($d/D = 0.088$) (Figure 6.6i) exhibited reduced jet asymmetry and an absence of the stagnation region in the near wake, allowing the jet to sustain for a longer duration. Due to the lack of stagnation and backflow region in the $d/D = 0.088$ case, the mean velocities remained positive even up to $x/c = 2$, unlike the smooth foil and dimpled cases with $d/D = 0.022$ & 0.044 .

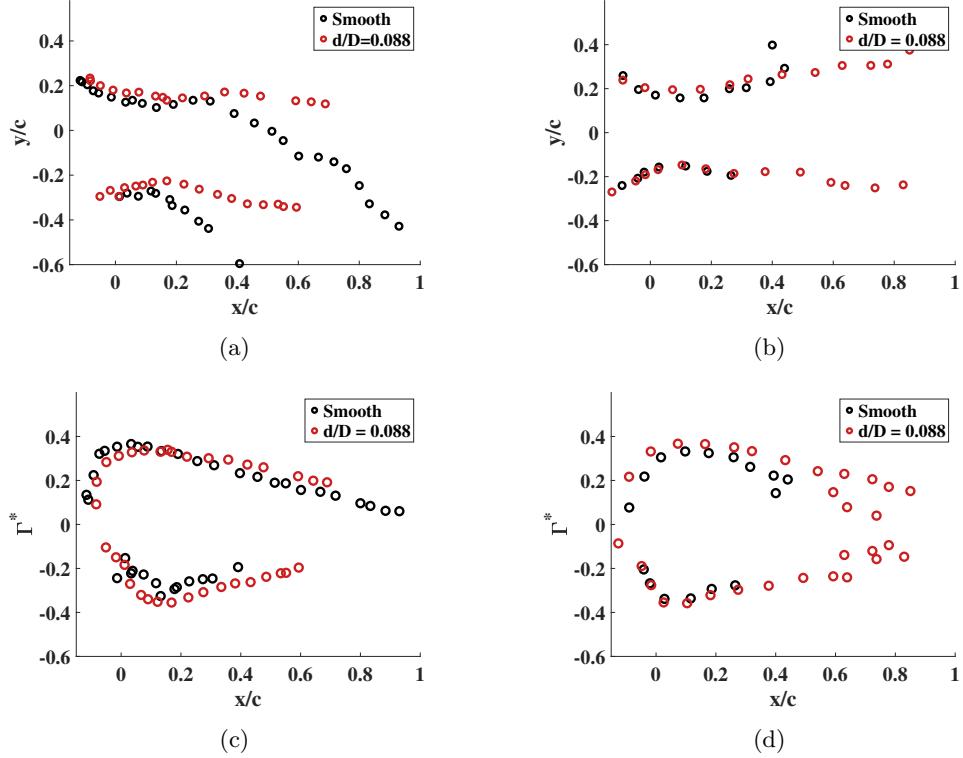


Figure 6.8: Comparison of the vortex center tracks at (a) $Re = 10000$, and (b) $Re = 20000$ and normalized circulation ($\Gamma^* = \Gamma/V_{TE_{max}}c$) at (c) $Re = 10000$, and (d) $Re = 20000$ between smooth foil and dimpled case ($d/D = 0.088$).

To summarize the discussion so far, this study demonstrates, for the first time, that dimples can alter the wake dynamics of a pitching foil in quiescent flow. The findings show that higher dimple depth ratios ($d/D = 0.088$) and Reynolds numbers ($Re = 20,000$) result in reduced jet deflection and prolonged sustainment of the time-averaged jet. In the following section, the vortex characteristics are quantified and compared across the different cases to further understand their impact on the observed flow characteristics.

Vortex dynamics

The flow field was phase-averaged over ten pitching cycles to analyze the vortex dynamics (ten cycles were deemed sufficient based on the convergence test presented in 6.4). The vortex trajectory and circulation were quantified and compared between the smooth foil and the dimpled foil ($d/D = 0.088$), as this case showed significant differences in time-averaged flow characteristics in the Reynolds numbers tested. The vortex trajectory was tracked by following the vortex center identified using the Gamma 1 (Γ_1) criterion and the circulation was quantified using the Gamma 2 (Γ_2) criterion, which are discussed in Sec. 6.2. The results for $Re = 10000$ and $Re = 20000$ are summarized in Figure 6.8.

At $Re = 10000$, the vortex trajectories differed significantly between the smooth and dimpled ($d/D = 0.088$) foils. For the smooth foil, the vortices exhibited greater downward inclination compared to the dimpled case (see Figure 6.8a). The normalized circulation ($\Gamma^* = \Gamma/V_{TE_{max}}c$) plot revealed a pronounced asymmetry between the shed vortices, which presumably contributed to the downward deflection of the vortex tracks and the inclination observed in the time-averaged flowfield (see Figure 6.8c). In this case, destructive vortex interactions (Shinde and Arakeri, 2013)

Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow

led to mutual neutralization, causing jet velocities to decay to near zero in this region. In contrast, the dimpled foil ($d/D = 0.088$) exhibited minimal asymmetry, reducing vortex track inclination. Consequently, both positive and negative circulation vortices propagated further downstream, sustaining jet velocity over longer distances. This enhanced downstream convection also eliminated the stagnation region observed in the wake of the smooth foil.

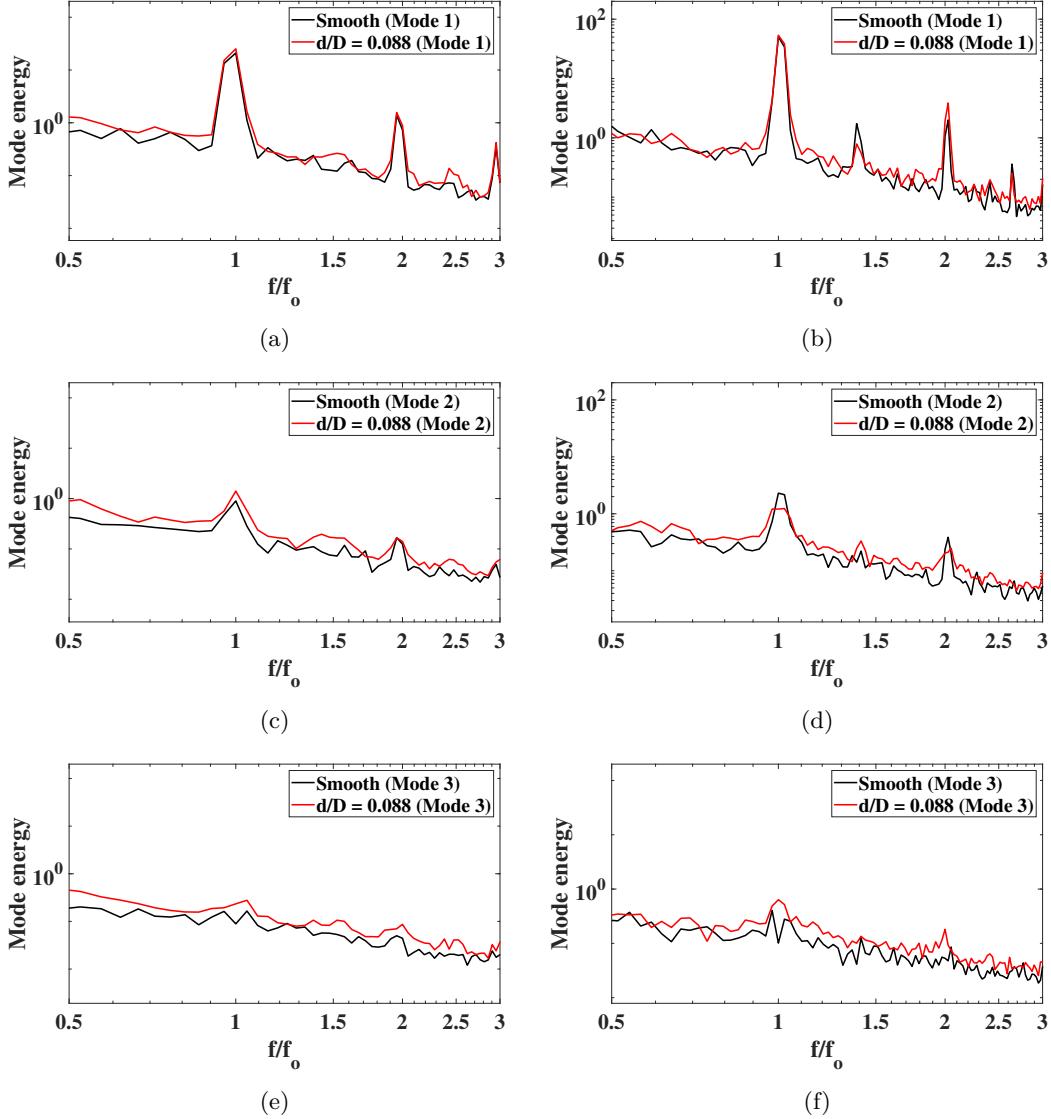


Figure 6.9: Characteristics of the modes from the Spectral Proper Orthogonal Decomposition (SPOD) analysis: (a) Mode 1 at $Re = 10000$, (b) Mode 1 $Re = 20000$, (c) Mode 2 at $Re = 10000$, (d) Mode 2 $Re = 20000$, (a) Mode 3 at $Re = 10000$, (b) Mode 3 $Re = 20000$. In the x-axis, f/f_0 is the modal frequency normalized by the pitching frequency.

At $Re = 20,000$, the vortex trajectories for the smooth foil exhibited no significant asymmetry or inclination, resulting in a largely symmetric time-averaged flow field (see Figure 6.8b). Similarly, the dimpled foil showed no inclination in its vortex tracks. However, its vortices convected further downstream compared to those of the smooth foil. Circulation analysis revealed that during both the development and decay phases, the shed vortices exhibited nearly identical circulation in both Reynolds number cases. However, vortices shed from the dimpled foil maintained their coherence longer, convecting farther downstream and sustaining jet velocities over greater distances. These

trends were also evident in the time-averaged flowfield data. This indicates that the dimples primarily influence dipole interactions in the near wake.

To quantify the vortex shedding frequency and the energy content, Spectral Proper Orthogonal Decomposition (SPOD) (Schmidt and Colonius, 2020) was performed for $Re = 10000$ and 20000 . As most of the energy was contained within the first three modes, the analysis focuses on these modes. The first mode represents the dominant vortex dipole structure seen in the time-averaged flow field, while the second and third modes capture the downstream-shed vortices. Comparison of the energy content among the modes reveals negligible differences in the first mode between the smooth and dimpled cases (see Figs. 6.9a and 6.9b). However, the dimpled foil exhibits slightly higher energy in the second and third modes (see Figs. 6.9c, 6.9d, 6.9e, and 6.9f). This supports earlier observations suggesting that surface dimples influence the vortex interactions in the near wake. The SPOD frequency spectra further show that the dominant frequencies correspond to the pitching frequency and its harmonics, with no significant differences between the smooth and dimpled foils. This aligns with previous findings that the actuation frequency predominantly governs wake shedding dynamics in pitching foils (Vilumbrales-Garcia et al., 2024; Zurman-Nasution et al., 2021). These results contrast with bluff body flows, where surface dimples can alter the shedding frequency (Li et al., 2015).

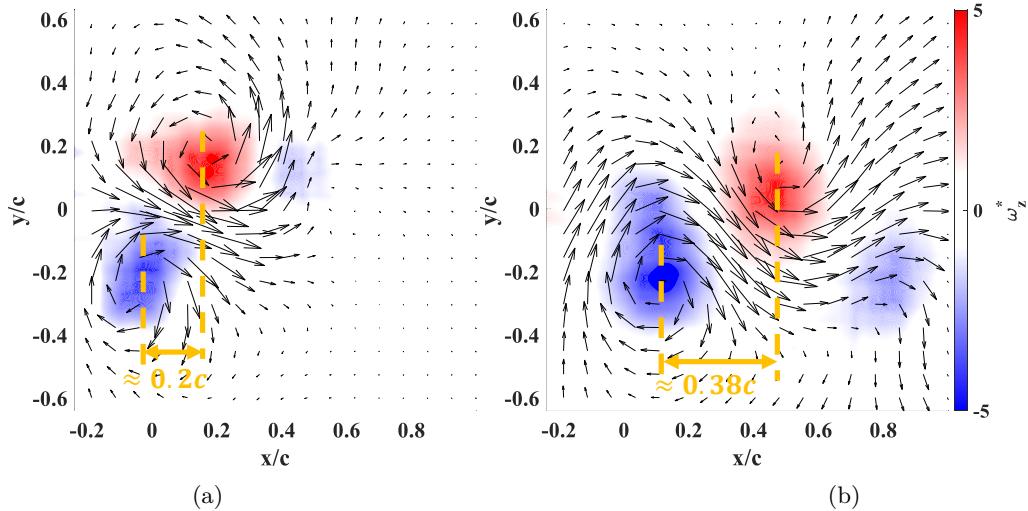


Figure 6.10: Comparison of the instantaneous phase-averaged vorticity field between (a) smooth foil and (b) dimpled foil ($d/D = 0.088$) at $Re = 10000$.

A comparison of phase-averaged vorticity snapshots between the smooth and dimpled foils at the same instant (see Figure 6.10) for $Re = 10000$ reveals that the spacing between the shed vortices increases noticeably for the dimpled case. This observation was consistent for the $Re = 20000$ case as well. An increase in vortex spacing can alter vortex-vortex interactions, potentially reducing destructive interference between counter-rotating vortices, as suggested by prior studies (Shinde and Arakeri, 2014; Heathcote et al., 2008). The observed change in vortex spacing for the dimpled foil is attributed to boundary layer modifications induced by the surface roughness. However, the current measurements are limited to wake dynamics, and a more detailed investigation of the near-wall flow and boundary layer evolution is necessary to fully understand the underlying mechanisms driving these changes. Based on the analysis presented in this section, it is conjectured that the presence of dimples alters the timing of vortex formation or vortex detachment, which in turn influences the dipole interactions in the wake and contributes to the observed flow modifications.

In summary, the vortex dynamics analysis revealed that although the circulation of vortices shed

from dimpled foils was comparable to that of the smooth foil, the presence of dimples notably altered vortex interactions and the evolution of the jet. At $Re = 10,000$, dimples reduced jet deflection and enhanced streamwise momentum by up to 50%. At $Re = 20,000$, while the jet was already more aligned, dimples further improved its coherence and extended its longevity. Instantaneous vorticity fields also showed increased vortex spacing in the dimpled case relative to the smooth foil. These observations suggest that dimples perhaps modify the timing of vortex formation and detachment by disrupting the boundary layer through localized surface indentations. This presumably leads to a more coherent and sustained wake with reduced deflection. These findings are consistent with previous bioinspired studies showing that surface indentations can passively manipulate vortex shedding to enhance propulsive performance and stabilize unsteady wakes. For example, Oeffner and Lauder (2012) and Domel et al. (2018) demonstrated that surface textures inspired by biological swimmers improve wake coherence and thrust. Similarly, Wen et al. (2014) and Mallah et al. (2021) found that surface roughness elements can modify boundary layer dynamics to delay separation and regularize vortex shedding. More recently, Guo et al. (2021) and Vilumbrales-Garcia et al. (2024) showed that passive surface features such as dimples or grooves significantly impact the timing and strength of trailing-edge vortex formation, improving the spatiotemporal organization of wakes in unsteady propulsion.

6.4 Summary and conclusions

The influence of surface roughness in the form of dimples on wake characteristics was investigated. The dimples were arranged in a staggered configuration with an area coverage ratio of 20.27%. While the dimple diameter remained constant, the dimple depth was varied as $d/D = 0.022, 0.044 \& 0.088$. A smooth foil without surface modifications served as the control case. The key findings of this study are summarized below:

1. The time-averaged wake of the smooth foil exhibited a vortex dipole near the trailing edge, with a stagnation region and backflow downstream. The backflow was induced by the low-pressure region created by the vortex dipole, while the stagnation region resulted from the interaction between the jet and backflow. The time-averaged jet exhibited two phases: an initial development phase with peak velocity, followed by a decay phase where the velocity gradually decreased to near zero.
2. The wake characteristics of the smooth foil varied with Reynolds number. At lower Reynolds number conditions, the vortex dipole exhibited noticeable asymmetry, and the jet displayed pronounced inclination. However, as the Reynolds number increased, both the asymmetry of the vortex dipole and the jet inclination were reduced.
3. The effects of dimples became prominent at higher Reynolds numbers ($Re = 10000$ and 20000), especially for the deepest dimpled configuration ($d/D = 0.088$). Under these conditions, the deepest dimpled case exhibited higher streamwise jet velocities during the decay phase and sustained the jet for a longer duration compared to both the smooth foil and the shallower dimpled cases ($d/D = 0.022 \& 0.044$).
4. The vortex dynamics analysis revealed that the $d/D = 0.088$ case exhibited weaker destructive interactions between shed vortices and maintained vortex coherence for a longer duration than both the smooth foil and the other dimpled configurations. This resulted in a more aligned jet with reduced deflection and enhanced longevity.

In conclusion, this study presents the first evidence that surface roughness in the form of dimples can influence key jet characteristics of the wake in pitching foils under quiescent flow conditions. These findings are valid at least within the geometric configurations and parameter space investigated in the present study. The effects were found to depend on both the dimple depth ratio and Reynolds number. The dimples enhanced the jet momentum in the wake, and the vortex dipole maintained coherence for longer durations. Although direct force or thrust measurements were not performed in this study due to experimental limitations, prior literature has established a strong correlation between wake jet coherence and thrust generation in flapping foils (Jones et al., 1998; Jallas et al., 2017). Based on these established links, we conjecture that the observed improvements in jet coherence and stability due to surface dimples could translate to improved propulsion efficiency in bio-inspired systems operating in quiescent or hovering conditions. This hypothesis remains to be tested through future studies involving direct force measurements. Further investigations are required to assess the broader applicability of these findings to other geometries, actuation regimes, and flow environments. This study contributes to a growing body of work exploring passive mechanisms for modifying near-wake dynamics in unsteady flows. By demonstrating that surface dimples can enhance jet stability and reduce asymmetry in a canonical pitching foil configuration, the findings offer insights that may inform the design of surface textures in low-speed maneuvering or hovering propulsion systems, such as those used in small underwater vehicles, aerial drones, or robotic swimmers.

Acknowledgments

A preprint has been submitted to *International Journal of Heat and Fluid Flow*.

Appendix

The influence of the initial condition on the direction of jet inclination for the smooth foil was examined by altering the direction of the first pitching cycle. The results for the three Reynolds numbers tested ($Re = 3700, 10000 \& 20000$) are presented in Figure 6.11. Despite the modification in the initial condition, the jet inclination remained downward for $Re = 3700 \& 10000$. For the $Re = 20000$ case, the degree of jet inclination was reduced across all cases. These findings indicate that, within the scope of the current study, the jet inclination is independent of the initial conditions.

Convergence tests for time and phase averaging were conducted by examining streamwise velocity profiles at downstream locations of $0.25c$, $0.5c$, and $0.75c$ from the trailing edge. As shown in Figure 6.12, the time-averaged velocity profiles for selected cycles of the smooth foil across all tested Reynolds numbers stabilized after the first ten cycles, with variations remaining within 5%. This validates the use of 50 cycles for time-averaging and 10 cycles for phase-averaging. This is particularly important for phase-averaging, as time-series data was used in the analysis. Since the acquisition frequency of the flow field measurements and the pitching frequency were not synchronized, phase-averaging over a larger number of cycles would have led to high temporal error propagation due to asynchronous measurements. Limiting the phase-averaging to ten cycles helped minimize these errors, and the convergence analysis confirms that sufficient convergence was achieved.

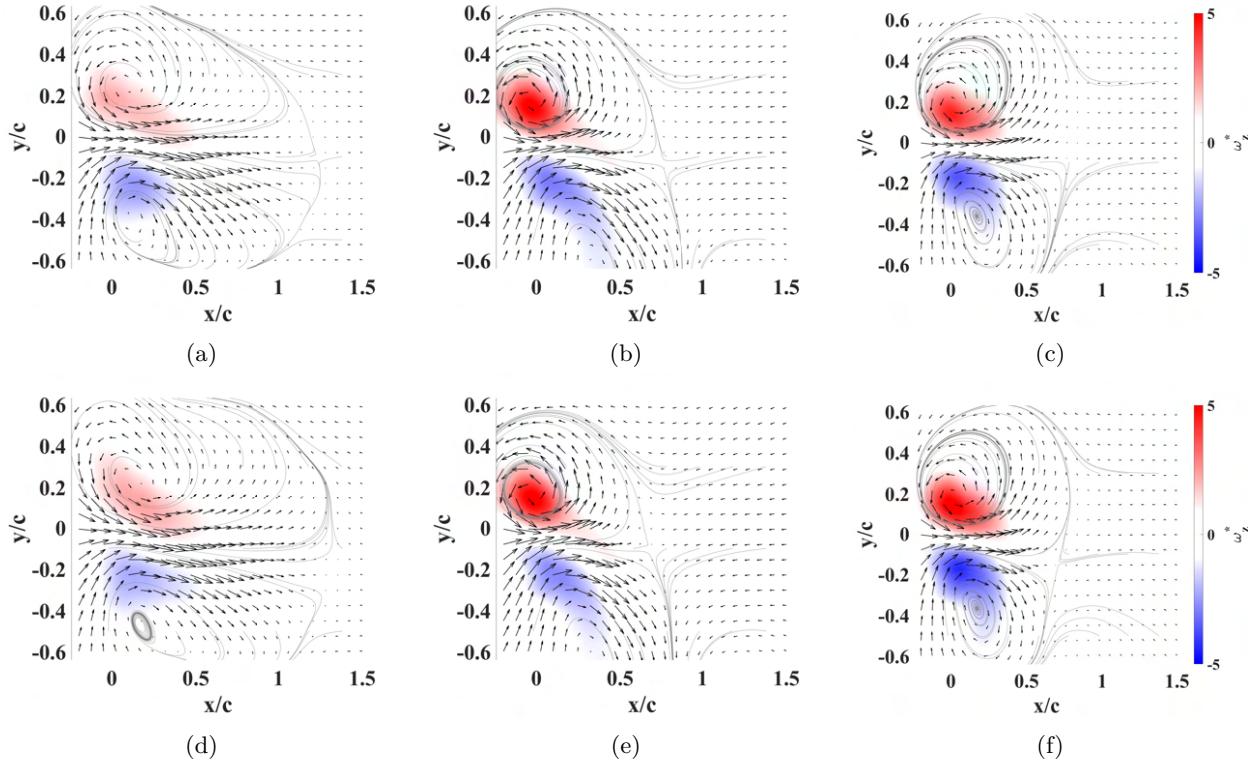


Figure 6.11: Time-averaged normalized vorticity ω_z^* for a smooth foil with non-inverted initial condition at (a) $Re = 3700$, (b) $Re = 10000$, (c) $Re = 20000$. Time-averaged flow field for a smooth foil with the inverted initial condition at (d) $Re = 3700$, (e) $Re = 10000$, (f) $Re = 20000$.

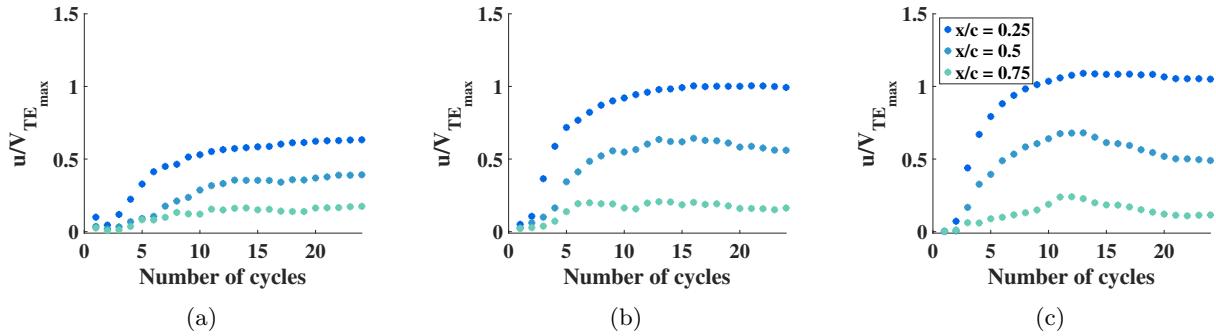


Figure 6.12: Moving average of the streamwise velocity profile at $x/c = 0.25$ between pitching cycles of one to thirty-five at (a) $Re = 3700$, (b) $Re = 10000$, and (c) $Re = 20000$.

References

H. Ali, M. R. Rasani, Z. Harun, and M. A. Shahid. Passive flow-field control using dimples for improved aerodynamic flow over a wing. *Scientific Reports*, 14(1):12918, 2024.

V. B. Ananthan, R. A. Akkermans, T. Hu, P. Q. Liu, and N. Rathje. Trailing-edge noise reduction potential of a locally applied shallow dimpled surface. *Journal of Sound and Vibration*, 525:116745, 2022.

J. M. Anderson, K. Streitlien, D. Barrett, and M. S. Triantafyllou. Oscillating foils of high propulsive efficiency. *Journal of Fluid Mechanics*, 360:41–72, 1998.

References

A. G. Domel, G. Domel, J. C. Weaver, M. Saadat, K. Bertoldi, and G. V. Lauder. Hydrodynamic properties of biomimetic shark skin: effect of denticle size and swimming speed. *Bioinspiration & Biomimetics*, 13(5):056014, 2018.

J.-E. Emblemsvag, R. Suzuki, and G. Candler. Numerical simulation of flapping micro air vehicles. In *32nd AIAA Fluid Dynamics Conference and Exhibit*, page 3197, 2002.

P. Fuglsang, I. Antoniou, N. N. Sørensen, and H. A. Madsen. Validation of a wind tunnel testing facility for blade surface pressure measurements. *Risø National Laboratory*, (Risø-R-1041(EN)), 1998.

R. Godoy-Diana, J.-L. Aider, and J. E. Wesfreid. Transitions in the wake of a flapping foil. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 77(1):016308, 2008.

R. Godoy-Diana, C. Marais, J.-L. Aider, and J. E. Wesfreid. A model for the symmetry breaking of the reverse Bénard–Von Kármán vortex street produced by a flapping foil. *Journal of Fluid Mechanics*, 622:23–32, 2009.

L. Graftieaux, M. Michard, and N. Grosjean. Combining PIV, POD and vortex identification algorithms for the study of unsteady turbulent swirling flows. *Measurement Science and Technology*, 12(9):1422, 2001.

P. Guo, K. Zhang, Y. Yasuda, W. Yang, J. Galipon, and D. E. Rival. On the influence of biomimetic shark skin in dynamic flow separation. *Bioinspiration & Biomimetics*, 16(3):034001, 2021.

I. Gursul and D. Cleaver. Plunging oscillations of airfoils and wings: progress, opportunities, and challenges. *AIAA Journal*, 57(9):3648–3665, 2019.

S. Heathcote and I. Gursul. Jet switching phenomenon for a periodically plunging airfoil. *Physics of Fluids*, 19(2):027104, 02 2007.

S. Heathcote, Z. Wang, and I. Gursul. Effect of spanwise flexibility on flapping wing propulsion. *Journal of Fluids and Structures*, 24(2):183–199, 2008.

D. Jallas, O. Marquet, and D. Fabre. Linear and nonlinear perturbation analysis of the symmetry breaking in time-periodic propulsive wakes. *Physical Review E*, 95(6):063111, 2017.

K. Jones, C. Dohring, and M. Platzer. Experimental and computational investigation of the Knoller-Betz effect. *AIAA journal*, 36(7):1240–1246, 1998.

S. Kumar, S. K. Singh, S. Jha, K. Baskaran, K. Srinivasan, and S. Narayanan. On the reductions of airfoil broadband noise through circular dimples. *Applied Acoustics*, 217:109819, 2024.

N. Lagopoulos, G. Weymouth, and B. Ganapathisubramani. Universal scaling law for drag-to-thrust wake transition in flapping foils. *Journal of Fluid Mechanics*, 872:R1, 2019.

J. Li, M. Tsubokura, and M. Tsunoda. Numerical investigation of the flow around a golf ball at around the critical Reynolds number and its comparison with a smooth sphere. *Flow, Turbulence and Combustion*, 95:415–436, 2015.

A. Mackowski and C. Williamson. Direct measurement of thrust and efficiency of an airfoil undergoing pure pitching. *Journal of Fluid Mechanics*, 765:524–543, 2015.

S. R. Mallah, P. Sooraj, A. Sharma, and A. Agrawal. Effect of superhydrophobicity on the wake of a pitching foil across various Strouhal numbers. *Physics of Fluids*, 33(11):111905, 11 2021.

J. Massey, B. Ganapathisubramani, and G. D. Weymouth. A systematic investigation into the effect of roughness on self-propelled swimming plates. *Journal of Fluid Mechanics*, 971:A39, 2023.

J. Oeffner and G. V. Lauder. The hydrodynamic function of shark skin and two biomimetic applications. *Journal of Experimental Biology*, 215(5):785–795, 2012.

D. B. Quinn, G. V. Lauder, and A. J. Smits. Scaling the propulsive performance of heaving flexible panels. *Journal of Fluid Mechanics*, 738:250–267, 2014.

M. Samimy and S. Lele. Motion of particles with inertia in a compressible free shear layer. *Physics of Fluids A: Fluid Dynamics*, 3(8):1915–1923, 1991.

References

O. T. Schmidt and T. Colonius. Guide to spectral proper orthogonal decomposition. *AIAA Journal*, 58(3):1023–1033, 2020.

A. Sciacchitano. Uncertainty quantification in particle image velocimetry. *Measurement Science and Technology*, 30(9):092001, July 2019.

S. Y. Shinde and J. H. Arakeri. Jet meandering by a foil pitching in quiescent fluid. *Physics of Fluids*, 25(4), 2013.

S. Y. Shinde and J. H. Arakeri. Flexibility in flapping foil suppresses meandering of induced jet in absence of free stream. *Journal of Fluid Mechanics*, 757:231–250, 2014.

S. Y. Shinde and J. H. Arakeri. Physics of unsteady thrust and flow generation by a flexible surface flapping in the absence of a free stream. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2218):20180519, 2018.

M. van Nesselrooij, L. Veldhuis, B. Van Oudheusden, and F. Schrijer. Drag reduction by means of dimpled surfaces in turbulent boundary layers. *Experiments in Fluids*, 57(9):142, 2016.

L. Veldhuis and E. Vervoort. Drag effect of a dented surface in a turbulent flow. In *27th AIAA Applied Aerodynamics Conference*, page 3950, 2009.

R. Vilumbrales-Garcia, M. Kurt, G. D. Weymouth, and B. Ganapathisubramani. Effects of surface roughness on the propulsive performance of pitching foils. *Journal of Fluid Mechanics*, 982:A1, 2024.

L. Wen, J. C. Weaver, and G. V. Lauder. Biomimetic shark skin: design, fabrication and hydrodynamic function. *Journal of Experimental Biology*, 217(10):1656–1666, 2014.

B. Wieneke. PIV uncertainty quantification from correlation statistics. *Measurement Science and Technology*, 26(7):074002, June 2015.

J. G. Wong and D. E. Rival. Determining the relative stability of leading-edge vortices on nominally two-dimensional flapping profiles. *Journal of Fluid Mechanics*, 766:611–625, 2015.

Q. Zhong and D. B. Quinn. Predicting the slowly converging dynamics of asymmetric vortex wakes. *Physical Review Fluids*, 9(6):064702, 2024.

A. N. Zurman-Nasution, B. Ganapathisubramani, and G. D. Weymouth. Fin sweep angle does not determine flapping propulsive performance. *Journal of the Royal Society Interface*, 18(178):20210174, 2021.

Dynamic Modeling of Cavitation Inception in the Subgrid Scale of Large Eddy Simulation

MEHEDI H. BAPPY & KRISHNAN MAHESH

We present a unified Large Eddy Simulation (LES) framework that dynamically computes all variables critical to cavitation inception, combining high-fidelity turbulence modeling and subgrid-scale (SGS) physics. Cavitation onset is determined not only by resolved flow features and their interaction, but also by rare, extreme low-pressure events driven by unresolved turbulence. Recent work by Madabhushi and Mahesh (2025) demonstrates that capturing the pressure minima responsible for inception requires extremely fine computational grids; conventional LES at coarser resolutions cannot resolve these intermittent, intense pressure dips due to inherent spatial averaging. Our approach addresses this limitation by incorporating a dynamic k -equation SGS closure, a statistical SGS cavitation inception model that utilizes local pressure fluctuations, and a vapor transport equation for dynamic variation in nuclei density. This integrated methodology enables robust and physically consistent prediction of cavitation inception in complex turbulent flows, dynamically reflecting local flow conditions and their fluctuations. Application to vortex interaction in a canonical twin hydrofoil configuration shows that the framework reliably predicts both inception location and frequency, even when resolved pressures alone would underestimate inception probability. The framework's successful validation across a range of angles of attack (AOA) further demonstrates its robustness and suitability for predictive engineering simulation of cavitation onset in turbulent flows at practical computational cost.

Keywords: cavitation inception, SGS modeling, vortex-vortex interaction

7.1 Introduction

The phenomenon of cavitation, characterized by the formation and subsequent dynamic evolution of vapor-filled cavities within a liquid subjected to sufficiently low local pressures, is a critical area of study in fluid mechanics. Its occurrence is pervasive across a multitude of engineering disciplines, most notably in naval architecture and hydromachinery, where it can precipitate significant deleterious effects on system performance, structural integrity, and operational longevity (Brennen, 2014; Knapp et al., 1970). In marine applications, for instance, cavitation manifesting on propeller blades or control surfaces can lead to a marked reduction in propulsive efficiency, induce severe erosive damage, generate unwanted vibrations, and significantly amplify radiated underwater noise, thereby compromising the acoustic stealth and overall effectiveness of naval vessels (Smith et al., 2024). Similarly, within the domain of nuclear engineering, cavitation poses substantial risks to reactor safety and operational efficiency by instigating material erosion, inducing detrimental pressure fluctuations, and impairing heat transfer capabilities in critical components such as primary coolant pumps, heat exchangers, and condensers (Li et al., 2023). Consequently, a profound understanding of the conditions leading to cavitation inception—the initial formation of these vaporous bubbles—is paramount for the design of robust and reliable hydraulic systems.

The fundamental physics governing cavitation inception hinges on the interplay between the local fluid pressure, p , the liquid's vapor pressure, p_v , and the availability of nucleation sites. Cavitation typically occurs when the local pressure in the liquid drops below its vapor pressure at a given temperature. This pressure reduction can be induced by high local velocities, such as those encountered in regions of flow acceleration around hydrofoils, within pump impellers, or in the cores of intense vortical structures. The susceptibility of a flow to cavitation is often quantified by the cavitation number, σ , defined as:

$$\sigma = \frac{p_\infty - p_v}{\frac{1}{2}\rho U_\infty^2} \quad (7.1)$$

where p_∞ is the reference far-field pressure, U_∞ is a reference velocity, and ρ is the liquid density (Brennen, 2014). A lower cavitation number indicates a greater propensity for cavitation. However, the simple criterion $p < p_v$ is often insufficient, as the tensile strength of pure liquids is remarkably high. In practice, cavitation almost invariably occurs heterogeneously, originating from microscopic gas nuclei, impurities, or crevices on solid surfaces that act as preferential sites for vapor formation (Knapp et al., 1970; Brennen, 2014). The size distribution and spatial density of these nuclei, N_d , play a crucial role in determining the actual inception pressure and the type of cavitation observed (Bappy et al., 2020).

In many engineering flows of practical importance, particularly those involving lifting surfaces or shear layers, cavitation inception is intimately linked to the dynamics of vortical structures. Shear-dominated flows, such as those developing in the wake of hydrofoils, around propellers, or within mixing layers, are inherently characterized by the formation, evolution, and interaction of vortices. These vortices, defined by regions of concentrated vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ (where \mathbf{u} is the velocity vector), often exhibit significantly reduced pressures within their cores due to the centripetal acceleration required to maintain the swirling motion. For an idealized Rankine vortex, the pressure in the core can be substantially lower than the ambient pressure, providing a natural locus for cavitation inception (Saffman, 1992). The work of Harvey et al. (1947) highlighted the importance of pre-existing gas nuclei stabilized in crevices for heterogeneous nucleation. Subsequent experimental investigations by numerous researchers further elucidated the link between flow separation, vortex formation, and cavitation inception on various body shapes, see Morch (2000).

A particularly complex and challenging scenario arises from the three-dimensional interaction of multiple vortical structures. As highlighted by extensive literature, in flows such as those around marine propellers and control surfaces, cavitation inception is frequently observed not within the primary, often spanwise-oriented, shedding vortices, but rather within weaker, typically streamwise or obliquely oriented, secondary vortical structures. The prevailing hypothesis, substantiated by both experimental observations and numerical simulations, posits that the intense strain field imposed by stronger, primary vortices can induce significant stretching of these weaker, secondary vortices (Knister, 2024; Leasca et al., 2025). This vortex stretching mechanism is a fundamental aspect of three-dimensional turbulent flows, described by the vorticity transport equation:

$$\frac{D\boldsymbol{\omega}}{Dt} = (\boldsymbol{\omega} \cdot \nabla) \mathbf{u} - \boldsymbol{\omega}(\nabla \cdot \mathbf{u}) + \nu \nabla^2 \boldsymbol{\omega} + \frac{1}{\rho^2} \nabla \rho \times \nabla p \quad (7.2)$$

where D/Dt is the material derivative, ν is the kinematic viscosity, and the term $(\boldsymbol{\omega} \cdot \nabla) \mathbf{u}$ represents the vortex stretching and tilting. For incompressible flow, $\nabla \cdot \mathbf{u} = 0$, and the baroclinic term vanishes if density is constant. The stretching process, by conserving angular momentum within the vortex core (or more formally, by Kelvin's circulation theorem under ideal conditions), leads to an amplification of vorticity and a concurrent reduction in core size, resulting in a dramatic

decrease in core pressure, often sufficient to trigger cavitation even when the ambient cavitation number is relatively high (Tennekes and Lumley, 1972). Furthermore, the inherent instabilities present in turbulent shear flows can exacerbate these vortex interactions and transient pressure minima. Long-wavelength instabilities, such as the Crow instability observed in systems of counter-rotating vortices (Crow, 1970), can lead to significant mutual distortion, stretching, and eventual reconnection or merging of vortical structures. These dynamic events create localized, transient regions of exceptionally low pressure within the cores of the interacting, often weaker, vortices, making them prime candidates for cavitation inception. Observations by Chang et al. (2012) and recent advanced experimental investigations, such as those by Knister et al. (2020); Knister (2024), employing sophisticated techniques like stereoscopic Particle Image Velocimetry (PIV) coupled with high-speed cavitation visualization, have begun to unravel the intricate three-dimensional kinematics and dynamics of these interacting vortex systems and their direct correlation with incipient cavitation events.

Despite these advances, predictive modeling of cavitation inception, particularly in complex three-dimensional and turbulent flow fields, remains a formidable challenge. Cavitation inception is inherently a multi-scale process, driven by transient, extreme pressure minima that occur within the smallest turbulent eddies. These critical pressure fluctuations are often orders of magnitude smaller in spatial and temporal scales than the overarching flow features. Accurately capturing these fleeting, localized events poses a significant challenge for numerical simulations, as demonstrated by Madabhushi and Mahesh (2025), which highlights that reliable prediction of cavitation inception in complex flow configurations, such as twin vortex interactions undergoing Crow instability, fundamentally necessitates simulations with very fine spatial and temporal resolution. The stochastic and intermittent nature of these extreme low-pressure events further compounds the difficulty in their precise numerical detection and traditional computational fluid dynamics (CFD) methodologies face inherent limitations in addressing this multi-scale nature. Direct Numerical Simulation (DNS), while capable of resolving all relevant turbulent scales and the associated pressure fluctuations, remains computationally prohibitive for the high Reynolds number flows characteristic of practical engineering applications. Consequently, DNS is typically restricted to lower Reynolds numbers or canonical flow problems. Reynolds-Averaged Navier-Stokes (RANS) models, widely employed for their computational efficiency, fundamentally rely on time-averaging the flow equations. This averaging process inherently smooths out the transient, small-scale pressure fluctuations that are the primary triggers for cavitation inception, leading to substantial inaccuracies in predicting the onset conditions and locations of cavitation.

LES offers a more sophisticated alternative by explicitly resolving the large, energy-containing turbulent eddies while modeling only the smaller, SGS motions (Pope, 2001). While LES has been increasingly applied to cavitating flows, demonstrating its potential to capture the unsteady dynamics of sheet, cloud, and vortex cavitation (Gnanaskandan and Mahesh, 2015), accurate prediction of inception remains a challenge (Bappy, 2022; Madabhushi and Mahesh, 2025). The most critical extreme pressure minima frequently manifest at the unresolved subgrid scales, which are effectively ignored or inadequately represented by standard LES closures. This deficiency creates a "missing physics" problem, where the very phenomena responsible for initiating cavitation are filtered out or averaged. The filtered pressure field obtained from LES, \bar{p} , may not accurately represent the instantaneous local pressure, p , experienced by a cavitation nucleus. This necessitates the development and implementation of a dynamic LES framework that aims to account for the influence of unresolved pressure fluctuations, $p'_{\text{sgs}} = p - \bar{p}$, and their impact on the activation of the cavitation nuclei for cavitation.

Overview of the proposed LES framework

To overcome the computational and physical modeling challenges, as mentioned above, this paper introduces a novel LES framework specifically tailored for accurate and computationally feasible prediction of cavitation inception. This framework integrates three advanced modeling components, each addressing a critical aspect of the multiscale and multiphysics nature of cavitation:

1. Dynamic SGS Turbulent Kinetic Energy Model (Chai and Mahesh, 2012): Building upon the foundational concepts of SGS modeling, this dynamic one-equation model solves a transport equation for the SGS turbulent kinetic energy, k_{sgs} . This provides a dynamic and physically consistent measure of unresolved turbulence and its fluctuations, which are essential inputs for the SGS cavitation model. This approach extends previous efforts by allowing for a more physically realistic representation of the intensity of subgrid pressure fluctuations, which are highly dependent on the local resolved flow dynamics.
2. SGS Cavitation Event Rate Model (Bappy et al., 2022): This model statistically accounts for unresolved pressure fluctuations at the subgrid level, effectively bypassing the strict requirement for extremely high spatial and temporal resolution. It allows for accurate cavitation inception predictions on coarser computational grids by relating the probability of an inception event to the local subgrid flow conditions.
3. Dynamic Nuclei Transport Model (Brandao and Mahesh, 2022): Another novel aspect of the current research is the utilization of a vapor transport model, which enables the dynamic computation of void fraction and the rigorous handling of polydisperse nuclei size distributions. This component is vital for accurately representing the sensitive dependence of cavitation inception on water quality and local nuclei concentration in highly turbulent flows, as nuclei can be convected, trapped in vortical structures, or generated by previous cavitation events.

This integrated framework holistically addresses the multi-physics and multi-scale nature of cavitation inception. Each component fills a critical gap in traditional LES, allowing for a more comprehensive and physically consistent prediction of inception in complex turbulent flows without incurring the prohibitive computational costs associated with DNS. The synergy between these models is central to achieving both accuracy and computational efficiency. The dynamic SGS turbulent kinetic energy provides the local turbulent environment, which then informs the statistical cavitation model, while the dynamic nuclei transport ensures the correct "seed" population for cavitation. This multi-faceted approach is designed to provide more accurate estimations of the cavitation inception pressure, potentially identifying inception events within secondary or highly strained vortex cores even when the resolved local mean pressure is considerably above the vapor pressure.

By addressing these challenges, this work seeks to advance the modeling of vortex-induced cavitation inception. The insights gained are anticipated to be vital for improving the design methodologies and operational envelopes of naval systems, such as propellers and rudders, where cavitation remains a persistent and critical challenge. Moreover, advances in the modeling approach, particularly those concerning the interplay between resolved turbulent structures, SGS flow, and Lagrangian bubble dynamics, are expected to offer strategic benefits for predictive modeling and risk mitigation of cavitation across a broader spectrum of engineering applications. The remainder of this paper is organized as follows. Section 7.2 details the governing equations, the LES methodology, the specifics of the SGS cavitation inception model including the dynamic k_{sgs} and nuclei transport components, and the numerical setup for the investigated flow configurations. Section

7.3 presents the results of our simulations, focusing on the vortex dynamics, pressure field characteristics, and predicted cavitation inception behavior under various conditions, including analyses of grid and nuclei sensitivity.

7.2 Methodology

The numerical simulations in this study are performed using a LES framework. The LES approach directly resolves the large, energy-containing turbulent motions while modeling the influence of the smaller, unresolved subgrid scales. This section details the governing equations, the SGS closure for the momentum equations, the statistical model employed for predicting cavitation inception, and the dynamic model for nuclei concentration.

Governing equations for LES

The simulations are performed using an unstructured overset grid method (Horne and Mahesh, 2021), which allows for increased flexibility by enabling meshes to overlap and move relative to each other. This approach provides fine resolution in key areas near the propeller blades while coarsening away from the body, which significantly reduces the total cell count and computational cost. The incompressible Navier-Stokes equations are solved in an Arbitrary Lagrangian-Eulerian (ALE) formulation, where the grid velocity is included in the convection term to avoid tracking multiple reference frames. The governing equations for the resolved-scale fluid motion are given by the filtered continuity equation:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \quad (7.3)$$

and the filtered momentum equation:

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial(\bar{u}_i \bar{u}_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j} \quad (7.4)$$

Here, x_i represents the Cartesian coordinates ($i = 1, 2, 3$), t is time, u_i is the i -th component of the filtered velocity vector, ρ is the constant fluid density, \bar{p} is the filtered pressure, and ν is the kinematic viscosity of the fluid. The overbar ($\bar{\cdot}$) denotes the spatial filtering operation, typically a convolution with a filter kernel of characteristic width Δ . The term τ_{ij} in Equation (7.4) is the SGS stress tensor, which arises from the filtering operation and represents the effect of the unresolved small-scale motions on the resolved large scales. This tensor is unclosed and requires modeling, it is defined as:

$$\tau_{ij} = \bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j \quad (7.5)$$

Equation (7.4) is solved using a finite volume segregated predictor-corrector scheme. The velocities are first predicted by solving the momentum equation and then corrected using the pressure Poisson equation to satisfy the continuity equation. The implicit Crank-Nicolson scheme is used for the temporal advancement of the solution. The interpolation between centroid and face velocities is performed in a manner that emphasizes discrete kinetic energy conservation.

Dynamic k-equation SGS model

To close Equation (7.4), the SGS stress tensor τ_{ij} modeled using an eddy viscosity approach. The deviatoric part of the SGS stress tensor, $\tau_{ij}^d = \tau_{ij} - \frac{1}{3} \tau_{kk} \delta_{ij}$ (where τ_{kk} is the trace of the SGS stress

tensor, representing twice the SGS turbulent kinetic energy k_{sgs} , and δ_{ij} is the Kronecker delta), is related to the resolved rate-of-strain tensor \bar{S}_{ij} :

$$\tau_{ij}^d = -2\nu_{\text{sgs}}\bar{S}_{ij} = -2\nu_{\text{sgs}} \left(\frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \right) \quad (7.6)$$

where ν_{sgs} is the SGS eddy viscosity.

In this work, ν_{sgs} is determined using a dynamic one-equation model based on the transport of SGS turbulent kinetic energy, $k_{\text{sgs}} = \frac{1}{2}\tau_{kk}$. We employ the model proposed by Chai and Mahesh (2012), which has demonstrated improved predictions of high-wavenumber energy spectra and kinetic energy evolution compared to algebraic SGS models, particularly in complex non-equilibrium flows. The SGS stress tensor is defined by the following relation, where \sqrt{k} is chosen as the velocity scale instead of $\Delta|\bar{S}|$:

$$\tau_{ij} - \frac{2}{3}\rho k \delta_{ij} = -2C_S \Delta \rho \sqrt{k} \tilde{S}_{ij}^*. \quad (7.7)$$

Here, C_S is the model coefficient for SGS stress and \tilde{S}_{ij}^* is the traceless strain-rate tensor defined as -

$$\tilde{S}_{ij}^* = \tilde{S}_{ij} - \frac{1}{3}\delta_{ij}\tilde{S}_{kk} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \frac{1}{3} \frac{\partial \bar{u}_k}{\partial x_k} \delta_{ij}. \quad (7.8)$$

A key feature of the model used is the dynamic determination of model coefficients in the k_{sgs} -equation. This dynamic procedure adapts the model to the local state of the flow, alleviating the need for ad-hoc tuning of constants. The approach typically involves the application of a second, test filter (with width $\hat{\Delta} > \Delta$) to the resolved flow field, allowing for the evaluation of Germano-type identities for k_{sgs} or related quantities (Germano et al., 1991; Lilly, 1992). The Lagrangian time scale is dynamically computed based on a surrogate-correlation of the Germano-identity error (Verma et al., 2013). This methodology aims to provide a more robust estimation of the SGS dissipation and viscosity.

SGS cavitation inception model

Cavitation inception for a nucleus of radius R_0 occurs when instantaneous pressure (p) falls below a critical value,

$$p_{\text{crit}}(R_0) = p_v - \frac{2S}{R_0} + p_g(R_0), \quad (7.9)$$

where p_v is vapor pressure, S is the surface tension, and p_g accounts for non-condensable gas in the nucleus. LES resolves only the filtered (large-scale) pressure, \bar{p} , and cavitation inception prediction will not be accurate until p accounts for the unresolved SGS turbulence. Therefore, LES needs to resolve the finest scales of the flow, comparable to the size of the nuclei population, to predict proper instantaneous pressure for computing cavitation inception. In such cases, p'_{sgs} would be negligible and can be ignored. However, in the absence of such highly-resolved LES, we need to model the SGS pressure fluctuations and their effects on nuclei to properly capture the inception of cavitation. Central to this approach is the modeling of SGS velocity fluctuation, u'_{sgs} , obtained from the dynamic k -equation model:

$$u'_{\text{sgs}} = \sqrt{\frac{2}{3}k_{\text{sgs}}}, \quad (7.10)$$

where k_{sgs} evolves dynamically according to local production, transport, and dissipation of SGS kinetic energy. This velocity fluctuation sets the scale for unresolved, rapid variations in the pressure field.

To link SGS velocity to pressure, we can assume that the unresolved pressure fluctuation, $p'_{\text{sgs}} = p - \bar{p}$, follows a probability density function (PDF) shaped by local SGS energy:

$$f(p'_{\text{sgs}}) = \frac{1}{\sqrt{2\pi\sigma_p^2}} \exp\left(-\frac{(p'_{\text{sgs}})^2}{2\sigma_p^2}\right), \quad (7.11)$$

with variance σ_p^2 typically related to k_{sgs} as $\sigma_p = C_p \rho k_{\text{sgs}}$, where C_p is a function of unresolved turbulence level defined by Re_{sgs} . This probabilistic construction is necessary because p'_{sgs} is not explicitly resolved. This framework currently utilizes the modeled effects of these fluctuations statistically, drawing on local SGS turbulence quantities. The probability that a nucleus in a given cell incents is thus

$$\mathbb{P}_{\text{inc}}(R_0) = \int_{-\infty}^{p_{\text{crit}}(R_0) - \bar{p}} f(p'_{\text{sgs}}) dp'_{\text{sgs}}, \quad (7.12)$$

which quantifies the likelihood that subgrid-scale pressure excursions will trigger cavitation. To get a 'successful' or 'detectable' cavitation event rates (when nuclei radius reached a certain size e.g. $250 \mu\text{m}$ or $10R_0$), a temporal scale of cavitation (the time it takes for the nucleus to reach successfull cavitation) needs to be considered as well. Bappy et al. (2022) provided a database of cavitation event frequency per unit time per unit volume, f_{cav} , as function of Re_{sgs} , ϵ_{sgs} , R_0 , and p_{abs} . The absolute value of pressure (p_{abs}) is calculated from the resolved LES at every grid cell for some arbitrary reference pressures, p_{ref} , that represents the reference static pressure of the experiment.

$$p_{\text{abs}} = p_{\text{ref}} + \rho U_{\infty}^2 \cdot \bar{p}. \quad (7.13)$$

The SGS variables are computed as

$$\epsilon_{\text{sgs}} = (C_S \bar{\Delta})^2 S^3, \quad (7.14)$$

$$Re_{\text{sgs}} = \frac{2.58 k_{\text{sgs}}}{\sqrt{\nu \epsilon_{\text{sgs}}}}, \quad (7.15)$$

where, C_S is the dynamically computed coefficient using Germano's identity (Chai and Mahesh, 2012).

The total cavitation event rate is then obtained by integrating f_{cav} for the nuclei size distribution $N(R_0)$ over the region of interest for the total simulation time:

$$F_{\text{cav}} = \int_0^T \int_{\mathbf{V}} f_{\text{cav}}(k_{\text{sgs}}, \epsilon_{\text{sgs}}, N(R_0), p_{\text{abs}}) dV dt. \quad (7.16)$$

The average cavitation event rate or frequency is simply -

$$\bar{F}_{\text{cav}} = \frac{F_{\text{cav}}}{T}. \quad (7.17)$$

Dynamic nuclei concentration model

Accurate modeling of cavitation inception depends on a realistic and dynamically evolving description of both the concentration and size distribution of cavitation nuclei. In turbulent flows, the number and size of nuclei are rarely spatially uniform or temporally constant, as they are influenced by advection, turbulent diffusion, local recirculation, and pressure fluctuations (Katz, 1984). To address these complexities, we employ a dynamic model that evolves nuclei concentration as a function of the local flow environment.

The nuclei population is characterized by a size distribution $N(R_0, \vec{x}, t)$, where $N(R_0)$ denotes the number density of nuclei (per unit volume per unit radius interval) at location \vec{x} and time t . The total local vapor volume fraction (α) occupied by nuclei is

$$\alpha(\vec{x}, t) = \int_{R_{min}}^{R_{max}} \frac{4}{3} \pi R_0^3 N(R_0, \vec{x}, t) dR_0. \quad (7.18)$$

The vapor concentration C is related to the local vapor volume fraction such that

$$C(\vec{x}, t) = \alpha(\vec{x}, t) \cdot \rho_v. \quad (7.19)$$

For practical implementation within LES, we solve a scalar transport equation for the vapor concentration:

$$\frac{\partial C}{\partial t} + \nabla \cdot (\bar{\mathbf{u}} C) = \nabla \cdot (D_C \nabla C) + S_C^{source} - S_C^{sink}, \quad (7.20)$$

where $\bar{\mathbf{u}}$ is the resolved flow velocity and D_C is the nuclei diffusivity (which may depend on local turbulence, e.g., through a turbulent Schmidt number $Sc_t = \nu_{sgs}/D_C$). S_C^{source} and S_C^{sink} represent nuclei production and loss mechanisms, including generation by surface entrainment or bubble fragmentation, and removal by dissolution, coalescence, or trapping.

The dynamic vapor concentration field $C(\vec{x}, t)$ then serves as a locally resolved input for the SGS cavitation inception model, ensuring that predictions of cavitation onset account for the true, inhomogeneous distribution of potential nucleation sites in the flow. In a fully polydisperse framework, Eq. (7.20) can be generalized and solved for each nuclei size class to track $N(R_0, \vec{x}, t)$, providing maximum resolution for multi-scale or population balance applications. For the monodisperse-approach presented here, $C(R_0) \propto N(R_0)$, and $N(R_0)_{t=t_n} = \frac{N(R_0)_{t=0} \cdot C(R_0)_{t=t_n}}{C(R_0)_{t=0}}$.

7.3 Numerical setup

This study investigates cavitation inception due to the interaction of vortices generated by a twin hydrofoil flow configuration. It involves a pair of hydrofoils, fixed at certain AOAs, designed to generate unequal strength vortex pairs that interact with each other downstream of the foils, providing a canonical setup for studying vortex-induced cavitation inception mechanisms. This configuration is designed to create a controlled environment in which a weaker stream-wise vortex is subjected to the strain field of a stronger primary vortex, leading to stretching and intensification of the weaker vortex core.

The hydrofoils employed are based on a modified NACA 66 section, similar to those used in previous studies of vortex interactions and cavitation (Shen and Dimotakis, 1989). Both hydrofoils share a root chord length $c = 167$ mm. The primary (bottom) hydrofoil is a rectangular planform with the modified NACA 66 profile, set at an AOA of 6° relative to the incoming flow. This foil is designed to generate a relatively strong tip vortex. The secondary (top) hydrofoil, responsible for generating the weaker vortex that is targeted for cavitation inception, features the same modified

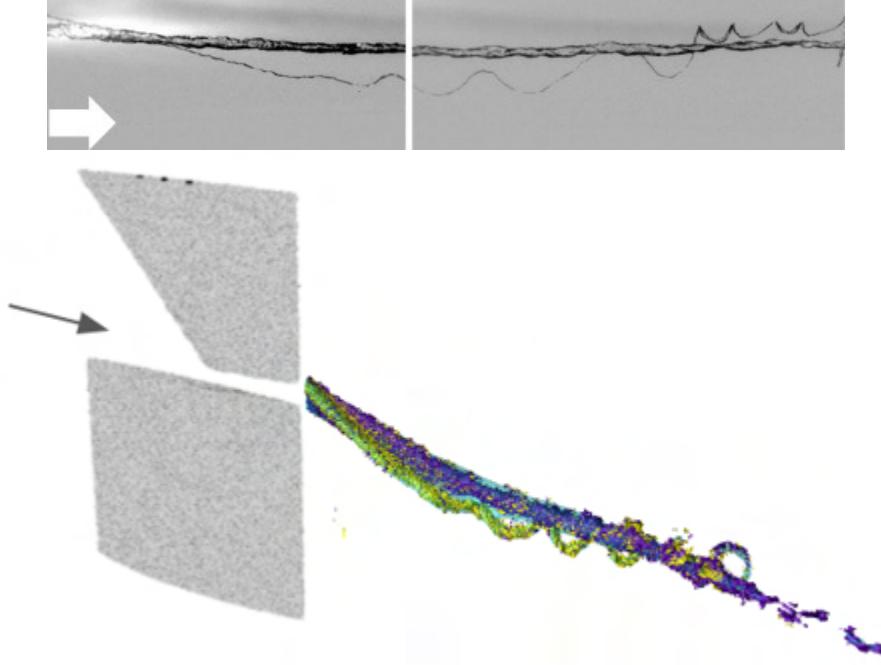


Figure 7.1: Numerical setup for the twin hydrofoil configuration. Top: Experimental image showing two initially non-interacting vortices of different strength interacting downstream (From Knister (2024)). Bottom: Numerical hydrofoils with counter-rotating vortices from one instantaneous flow solution of the medium grid.

NACA 66 profile but is tapered along its span and incorporates a semicircular tip. This geometry promotes the formation of a more concentrated, tightly rolled-up tip vortex. This foil is mounted upstream and above the primary foil and is set at different AOAs of $0^\circ, -2^\circ, -4^\circ$. The relative positioning of the foils is critical to achieve the desired vortex interaction. Figure 7.1 illustrates the numerical setup, detailing the hydrofoil geometries, their relative spacing, and orientation within the computational domain.

The computational domain is a rectangular channel designed to simulate an effectively unbounded environment around the hydrofoils, similar to a large water tunnel test section. The leading edge is at $x = 0$ and the trailing edge is at $x = 1.0$. The domain extends $3c$ upstream of the leading edge of the primary hydrofoil, $7c$ downstream to allow full development and observation of the interacting vortices, and $0.5c$ in both the vertical and spanwise directions from the center of the hydrofoil system. A uniform velocity profile $\mathbf{u} = (U_\infty, 0, 0)$ with $U_\infty = 10\text{m/s}$ is specified at the inlet while an outflow boundary condition with zero-gradient velocity and pressure applied at the outlet. No-slip wall conditions are applied on the top, bottom, and side boundaries to represent the experimental channel section. Finally, no-slip boundary conditions are applied on the surfaces of both hydrofoils.

To assess the sensitivity of the simulation results to spatial resolution, particularly for predicting pressure minima and the efficacy of the SGS cavitation model, two different grid resolutions were used (see Table 7.1) which were compared to the experimental study of Knister (2024).

The meshes are hybrid with prismatic layers near the leading and trailing edges of the hydrofoil surfaces. Boundary layers are resolved with $y^+ < 1$ for the first cell height and hexahedral cells in the far-field and vortex interaction regions. Significant mesh refinement is applied in the regions where the tip vortices are expected to form and interact, based on preliminary simulations and theoretical estimations of vortex trajectories.

| Grid | Coarse | Medium | Fine |
|-----------------------------------|--------|--------|------|
| Number of cells ($\times 10^6$) | 110 | 180 | 250 |
| Cells in vortex core | 6 | 12 | 25 |
| Timestep ($\times 10^{-4}$) | 5 | 2 | 1 |
| Number of Processors | 1250 | 2070 | 3072 |

Table 7.1: Spatial & temporal resolution with computational resources used for simulation of each numerical grid.

Unless otherwise specified, all the quantities presented are nondimensionalized using the reference chord length c , the reference velocity U_∞ , and the reference density ρ . The reference quantities for this configuration are the undisturbed upstream velocity $U_\infty = 10\text{m/s}$, the chord length of the primary hydrofoil $c = 0.167\text{m}$, and the density of freshwater $\rho = 998\text{kg/m}^3$. This yields a chord-based Reynolds number $Re_c = U_\infty c / \nu \approx 1.67 \times 10^6$.

7.4 Results

This section presents the key findings from the large-eddy simulations (LES) and the application of the SGS cavitation inception model. The results focus on the complex flow structures generated by the interacting hydrofoil wakes, the resulting pressure fluctuations, the characteristics of both resolved and unresolved turbulence, the role of cavitation nuclei, and the predicted cavitation event rates. Comparisons are made to available experimental data where possible, and the sensitivity of the predictions to numerical parameters like grid resolution is investigated.

Flow field

The flow field downstream of the hydrofoils is characterized by the generation and evolution of complex vortex structures originating from the leading edges. As illustrated by isosurfaces of pressure in Fig. 7.1 (e.g., blue isosurfaces at $\bar{p}/\rho U_\infty^2 = -0.12$), two distinct vortices of differing strengths are observed shed from each hydrofoil. One vortex is typically weaker and rotates counter-clockwise (CCW), while the other is stronger and rotates clockwise (CW). These vortices propagate downstream, drawing closer to each other. The weaker, secondary vortex is observed to wrap around the stronger, primary vortex, initiating a complex interaction process. Further downstream, this interaction leads to the onset of the Crow instability, resulting in the formation of an intricate array of spanwise and streamwise vortex structures, as shown in Fig. 1. This process is consistent with observations of interacting vortex pairs in other studies, such as those detailed by (Leweke et al., 2016). There are three distinct regions: undisturbed vortex pair, undulating secondary vortex with undisturbed primary vortex, and mutually interacting vortices.

During their interaction and subsequent evolution, the primary and secondary vortices, along with the smaller breakup vortices resulting from the instability, are subjected to significant shear and turbulent fluctuations indicated by high levels of turbulent kinetic energy (TKE). Prior to the strong interaction, the minimum pressure is typically located within the core of the stronger primary vortex, aligning with expectations for isolated or weakly interacting vortices where the pressure minimum scales with the square of the tangential velocity and vortex core radius. However, following the merging and interaction phase, the dynamics become more complex. The secondary vortex undergoes significant vortex stretching, a process known to reduce the vortex core size. This

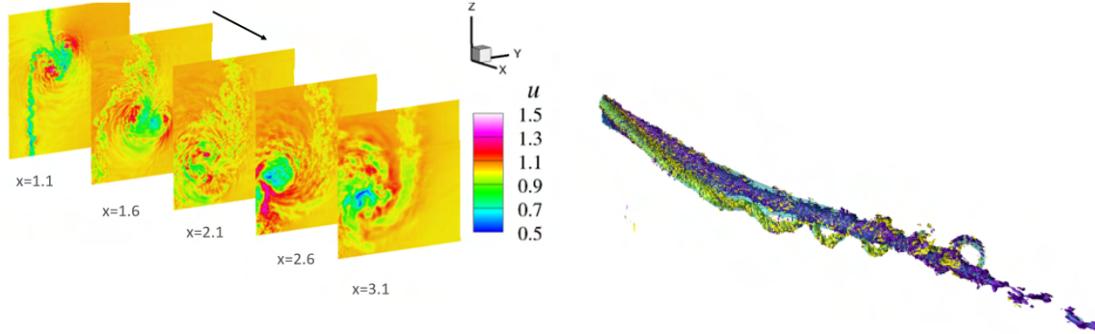


Figure 7.2: (a) Contour plot of a instantaneous solution for streamwise velocity u illustrating the overall flow pattern across stations at different downstream locations in the wake of the hydrofoils. (b) Isosurfaces of instantaneous vorticity ω at the same instant visualizing the instantaneous three-dimensional vortical structures in the flow.

reduction in core size leads to an intensification of tangential velocities and a further decrease in pressure within the core. Additionally, the development of axial flow or 'jetting' within the vortex cores can also contribute to a reduction in central pressure. The combination of these dynamic effects, particularly the intense stretching experienced by the secondary vortex, results in notable pressure drops predominantly affecting its core. This explains why the secondary vortex core is often observed to be the preferential site for cavitation inception, even under conditions where the ambient fluid pressure is relatively high (Knister, 2024), as shown in Fig. 2 which presents instantaneous streamwise velocity and pressure contours at various downstream stations. The velocity fluctuations are negligible far from the vortex cores, moderate within the cores, but have higher values around the vortex peripheries. Further analysis of instantaneous pressure contours (Figure 7.3) confirms these localized low-pressure regions within the evolving vortex structures.

| | Location of Instability | b | r_p | r_s |
|-------------------------------|-------------------------|-------|--------|-------|
| Experiment | 2.1-2.5 | 0.083 | 0.0335 | 0.021 |
| Coarse | 2.0-2.6 | 0.08 | 0.037 | 0.026 |
| Medium | 2.1-2.6 | 0.085 | 0.035 | 0.025 |
| Fine | 2.15-2.65 | 0.088 | 0.034 | 0.025 |
| Madabhushi and Mahesh (2025)* | 2.1-2.6 | 0.09 | 0.033 | 0.024 |

Table 7.2: Comparison of instability location, and values of distance between vortices (b), radii of primary and secondary vortices at $x = 1.68$. All quantities are measured similar to Knister (2024). Madabhushi and Mahesh (2025) simulation was performed at an AOA of -1.5° .

Resolved pressure field and fluctuations

The nondimensional pressure field resolved in the LES is presented in Fig. 7.3. The measurement region spans from 0.0 to $2.0c$ (measured from trailing edge) in the streamwise direction, $0.6c$ to $1.2c$ in the wall-normal direction, and $-0.3c$ to $0.3c$ in the spanwise direction. Fig. 7.3(b) has the iso-surface of $\bar{p} = -0.12$ showing the relative size of the vortices as they move downstream and interact. From station-wise instantaneous pressure contours in Fig. 7.3(a), we can see that the pressure is initially very low within the cores of the vortices; but in the downstream region, these low pressures

within the cores (more for the secondary vortex) do not sustain as the vortices gradually lose their strength. Therefore, cavitation is less probable beyond $0.5c$ downstream of the trailing edge of the foils. However, due to the instabilities through vortex-vortex interaction, the secondary vortex experiences stretching and the local pressure within the secondary vortex drops significantly to initiate cavitation there. The instantaneous pressure is utilized for the prediction of cavitation inception.

For the medium and coarse grids, the minimum pressure histories are presented in Fig. 7.3(c) and their PDFs are in Fig. 7.3(d). Over the entire simulation duration, the minimum instantaneous pressure recorded within this region is approximately $-1.05 \times \rho U_\infty^2$ for the medium grid at $AOA = -4^\circ$.

The time history of the average pressure within this box exhibits a mean value of approximately $-0.1 \times \rho U^2$. The fluctuations around this mean are observed to be within approximately 5% of ρU_∞^2 . Focusing on the resolved minimum pressure history, the mean value is substantially lower, around $-0.35 \times \rho U_\infty^2$. The time history of the minimum pressure reveals numerous events where the pressure drops below $-0.6 \times \rho U_\infty^2$. Assuming that the resolved lowest pressure persists for a sufficient duration to allow for the growth of a cavitation nucleus into a detectable event, the cavitation inception pressure (p_{inc}) can be estimated directly from the minimum resolved pressure (p_{min}) using the relationship $p_{inc} = -p_{min} \times \rho U^2 + p_{vap}$, where p_{vap} is the vapor pressure of the fluid. Using this criterion, the predicted cavitation inception pressure is 103.91 kPa, resulting in a cavitation inception number $\sigma_i = 1.05$. The inception number is defined as $\sigma_i = \frac{p_{inc} - p_{vap}}{\frac{1}{2} \rho U_\infty^2}$, where p_{ref} is the reference pressure, taken upstream of the foils to obtain undisturbed freestream pressure.

Resolved turbulence

Accurate representation of turbulence is paramount in predicting cavitation inception, as turbulent fluctuations contribute significantly to local pressure variations. In LES, turbulence is partitioned into resolved and SGS components. The resolved turbulence, which comprises the large, energy-containing eddies, is directly computed by the simulation. The spatial distribution and intensity of resolved turbulence provide insights into the flow regions most susceptible to low-pressure events.

As shown in Fig. 7.4, contours of resolved TKE reveal regions of high velocity fluctuations downstream of the hydrofoils. These high-TKE regions are primarily concentrated within and around the interacting vortex structures and in the wake shear layers. The elevated resolved TKE indicates significant resolved velocity fluctuations, which are directly linked to resolved pressure fluctuations through the incompressible Navier-Stokes equations. Regions of high resolved TKE are therefore potential sites for instantaneous low pressure events, especially when coupled with the deterministic pressure drop due to vortex dynamics. The spatial extent and magnitude of the resolved turbulence observed here are qualitatively consistent with experimental measurements of streamwise-vorticity (Table 7.2).

Unresolved turbulence

A critical component for accurately modeling cavitation inception in LES, particularly for grids that do not fully resolve the Kolmogorov scale, is the estimation of unresolved or SGS turbulence. While the resolved field captures the large-scale pressure fluctuations, the small-scale turbulent eddies, though unresolved, can induce significant localized pressure drops that trigger cavitation inception. The SGS inception model specifically accounts for the influence of these unresolved scales on cavitation nuclei through modeling their cavitation event rates.

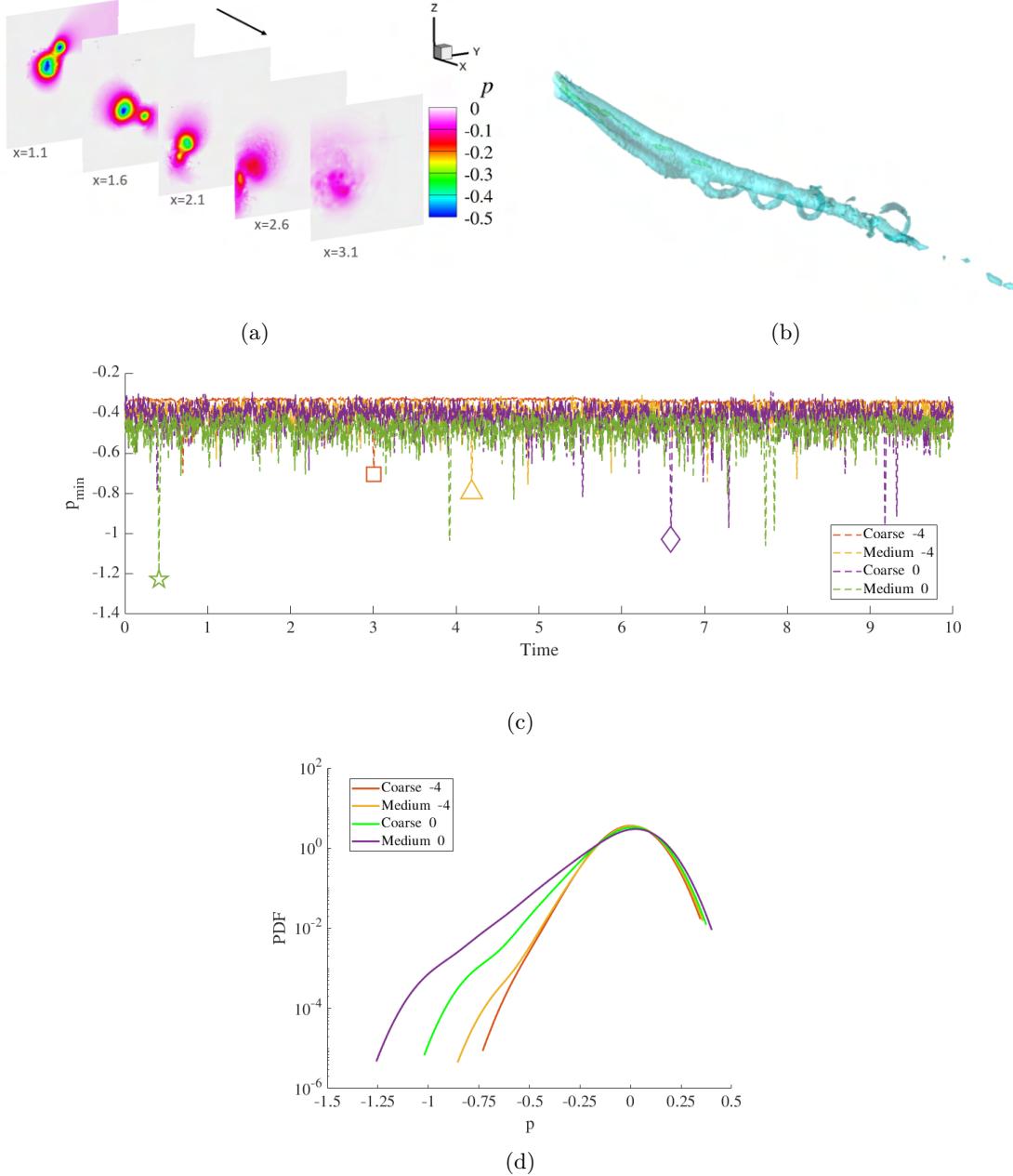


Figure 7.3: (a) Contour plot of the time-averaged pressure coefficient $C_p = (\bar{p} - \bar{p}_{\text{ref}})/(0.5\rho U_{\infty}^2)$, showing the spatial distribution of the mean pressure field downstream of the hydrofoils relative to the reference pressure \bar{p}_{ref} . (b) Isosurfaces of pressure (\bar{p}) at -0.12 (Blue) and -0.25 (green) showing the low pressure within the vortices. (c) Minimum pressure histories for coarse and medium grids at -4° and 0° AOAs, the symbols represent the minimum pressure within the specified time. (d) and pdfs of resolved pressure within the defined region.

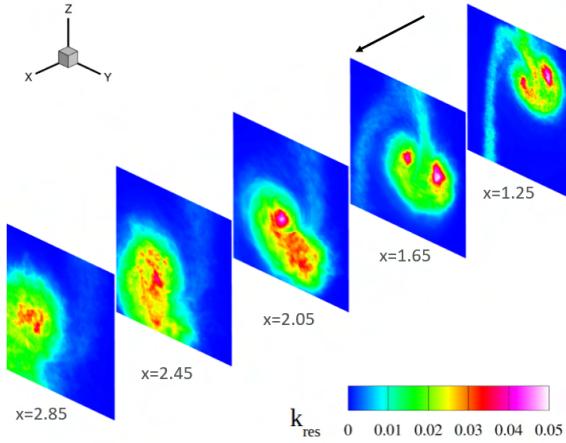


Figure 7.4: Contour plot of the time-averaged resolved turbulent kinetic energy k_{res} , showing the spatial distribution of kinetic energy contained in the resolved turbulent scales.

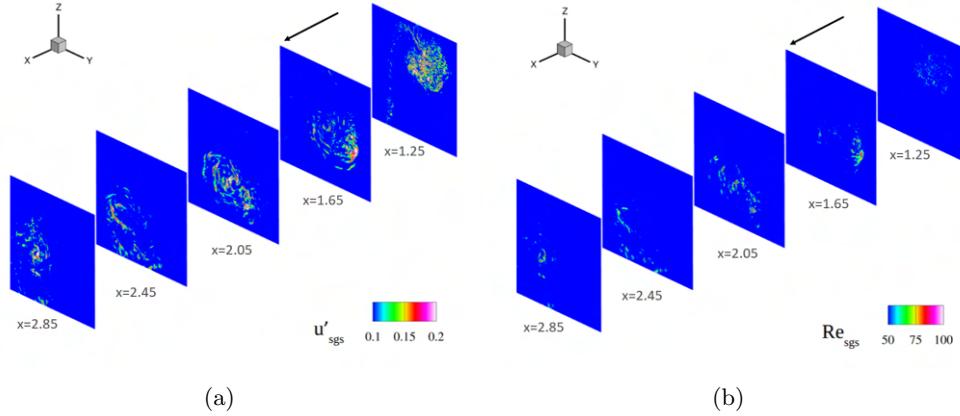


Figure 7.5: SGS (a) velocity fluctuation scale u'_{sgs} and (b) Taylor-scale Reynolds number Re_{sgs} at a representative time, indicating the spatial distribution of the velocity scale of the unresolved turbulence as estimated by the SGS model.

The model framework relies on estimating key SGS turbulence parameters based on the resolved flow field and the local grid characteristics. These parameters include the SGS dissipation rate (ε_{sgs}), the SGS velocity fluctuations (u'_{sgs}), and the SGS Taylor-scale Reynolds number (Re_{sgs}). These quantities are then used to statistically determine the pressure fluctuations experienced by a spherical nucleus of initial radius R_0 at a local mean pressure P_0 . Higher strain rates and larger grid sizes lead to increased SGS turbulence estimation. Such estimated SGS turbulence, in turn, predicts higher SGS pressure fluctuations that augment the resolved low pressures. Figure 7.5 illustrates contours of instantaneous u'_{sgs} and Re_{sgs} highlighting regions where the SGS model is most active. The contour plot of Re_{sgs} indicates the level of SGS turbulence downstream of the hydrofoils. Generally, these values are relatively low, suggesting that the grid is well resolved and the contribution of SGS pressure fluctuations to the total pressure experienced by a nucleus should be modest compared to cases with coarser grids. Outside the vortex cores, Re_{sgs} is very low (< 10), but within cores and around the core boundaries, Re_{sgs} can reach values greater than 250. This spatial variation underscores the importance of considering local turbulence levels when assessing cavitation potential. The magnitude of the SGS velocity fluctuations (u'_{sgs}) ranges from 0 to 18% of

the freestream velocity (U_∞) for -4° AOA case, with an average value of about 5% in the vicinity of the vortices. From the perspective of the SGS inception model, these values, particularly in the localized high-activity regions, are substantial enough to introduce a notable correction to the inception pressure predicted solely by the resolved field. Based on the model's formulation and the typical nuclei size distribution, these SGS fluctuations can contribute an additional pressure drop of 10-40 kPa for inception, depending on the specific nucleus size (R_0) and local absolute pressure (p_{abs}). This demonstrates the importance of the SGS model in capturing the cumulative effect of unresolved turbulence on cavitation susceptibility. Further analysis could involve examining the correlation between high Re_{sgs} regions and the spatial location of resolved low-pressure events to better understand the interplay between resolved and unresolved scales in driving cavitation inception.

Nuclei distribution and transport

Cavitation inception is fundamentally dependent on the presence of cavitation nuclei, which are small bubbles or cavities within the liquid that serve as preferential sites for vapor bubble growth. The spatial distribution and transport of these nuclei within the flow domain significantly influence where and when inception occurs. In complex turbulent flows, nuclei are not passively advected but can be influenced by pressure gradients, turbulent diffusion, and potentially even centripetal forces within vortex cores.

To assess the impact of nucleus distribution, two distinct approaches were investigated: one assuming a constant, uniform nuclei concentration throughout the flow domain ($CN(R_0)$), and another employing a dynamic nuclei concentration model $DN(R_0)$. For both methods, the initial nuclei size distribution is estimated using the suggested nuclei size distribution by (Li and Carrica, 2021) with a nuclei density of 1×10^7 nuclei/m³. The dynamic model accounts for their advection by the resolved flow, turbulent diffusion based on SGS quantities and dispersion.

Modeled SGS cavitation events

Leveraging the information from the resolved flow, the estimated unresolved turbulence, and the nucleus distribution, the SGS inception model predicts the volumetric rate of cavitation events (f_{cav}) within the flow. This model processes a set of local independent variables, including the SGS Taylor-scale Reynolds number (Re_{sgs}), the SGS dissipation rate (ε_{sgs}), the local mean pressure (P_0), and the initial nucleus radius (R_0), to determine the probability of a nucleus experiencing a sufficient pressure drop and duration to trigger cavitation.

Fig. 7.6 presents instantaneous cavitation event rates at five different downstream stations, assuming a dynamic nuclei density. These plots reveal that the model predicts higher event rates predominantly in the far downstream region, where the vortex structures are undergoing intense interaction and breakup. Interestingly, the model does not predict high event rates in the near downstream of the hydrofoils, despite the presence of relatively low resolved pressures there. This suggests that the combination of pressure magnitude, resolved and unresolved turbulence, and potentially the residence time of nuclei in low-pressure regions are all critical factors for inception, and the model well captures this interplay. Furthermore, the contours clearly indicate that the higher event rates are concentrated within the cores of the secondary vortices, which aligns well with experimental observations of cavitation initiation locations. This spatial agreement provides strong support for the SGS inception model's ability to correctly utilize the information about unresolved scales and local conditions to identify cavitation hotspots.

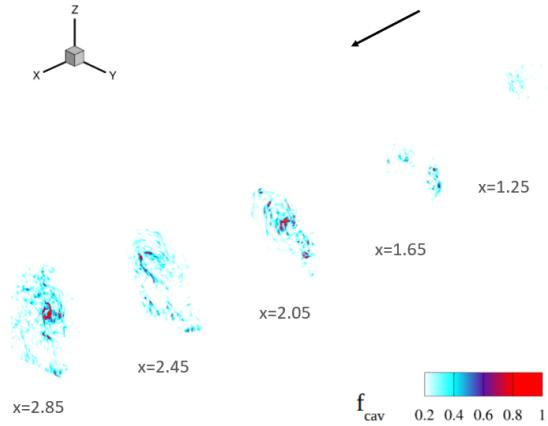


Figure 7.6: Contour plot of the instantaneous cavitation event frequency f_{cav} , showing the locations with high probability of having cavitation inception.

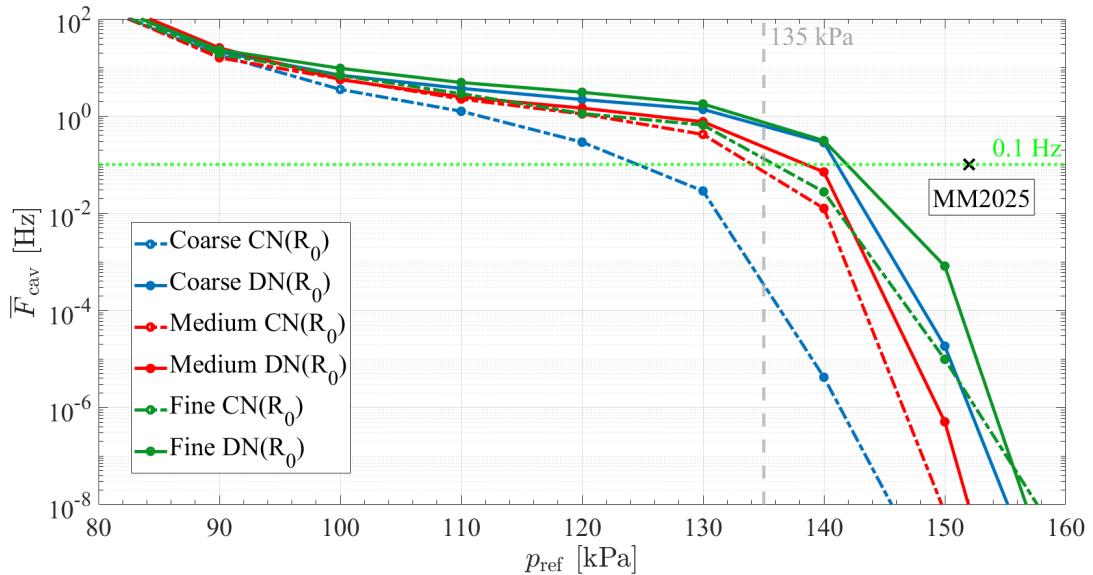


Figure 7.7: Average cavitation event frequency for Medium & Coarse grids with and without the dynamic nuclei size distribution. MM2025: Madabhushi and Mahesh (2025).

The instantaneous cavitation event rates can be averaged spatially and temporally at different reference pressures to generate a cavitation event rate curve, as shown in Figure 7.7. These curves, typically plotting f_{cav} against the reference pressure, provide a means to compare the simulation results with experimental data, which often report cavitation event frequency as a function of test section pressure. The curves presented in Figure 7.7 indicate that small cavitation activities are predicted even at relatively high reference pressures, such as 150 kPa. This is particularly notable considering that the local resolved minimum pressure in the simulation might only reach around 103 kPa. The ability of the SGS model to predict inception at higher reference pressures is a direct consequence of accounting for the additional pressure fluctuations induced by the unresolved turbulence, which can momentarily drive the total pressure experienced by a nucleus below the vapor pressure even when the resolved pressure is higher. The time-averaged f_{cav} contours also suggest strong cavitation activity localized between $1.2c$ and $1.8c$ downstream of the hydrofoils (where c is the hydrofoil chord length), which is consistent with experimental observations of the primary cavitation zone. Therefore, both the predicted inception pressure (as indicated by the onset of significant event rates on the curves) and the spatial location of inception show reasonable agreement with experimental data (see Tables 7.3 and 7.4).

Effect of grid resolution

The spatial resolution of the LES grid exerts a profound influence on the ability to capture the fine-scale, intermittent turbulent motions that govern the local pressure minima responsible for cavitation inception. In LES, grid spacing sets a lower bound for the smallest scales of motion that are explicitly resolved, meaning smaller, faster pressure fluctuations—those most likely to transiently drive local pressures below the vapor threshold—are filtered out in coarser simulations. This results in a fundamental trade-off: while finer grids increase fidelity by resolving a greater portion of the turbulence spectrum, they also elevate computational cost to levels that are often prohibitive for practical engineering applications.

Our results, supported by the findings of Madabhushi and Mahesh (2025), highlight the limitations of using only the minimum resolved pressure (p_{min}) as a criterion for inception. On coarse grids (e.g., -4° AOA), predicted inception pressures based on resolved p_{min} are as low as 97.29 kPa, dramatically below the experimentally observed range. Even with our finest grid, the minimum resolved pressure rises only to 115.39 kPa—still notably under the experimental inception threshold of approximately 135–155 kPa (for high DO conditions). In contrast, Madabhushi and Mahesh (2025), using an exceptionally fine LES grid (and at -1.5° AOA), report an inception pressure near 152.83 kPa, which closely matches experimental results. This demonstrates that reliably capturing the most extreme, short-lived low-pressure events using minimum-pressure criteria alone effectively requires grids that may be computationally infeasible option for complex flows.

This strong dependence on grid resolution arises because coarse meshes inherently average out, or suppress, the fine-scale turbulence that feeds the statistical tails of the pressure distribution. In turbulent flows, inception-relevant pressure dips—often associated with intensifying, stretched vortex cores or sharp local straining—can be highly localized both in space and time. If these spots occur below the filtering scale, the resolved LES solution simply cannot predict their occurrence, thereby underestimating cavitation risk.

To overcome this intrinsic limitation, our framework integrates a subgrid-scale (SGS) inception model that statistically reconstructs the effect of unresolved turbulence on the pressure field. By leveraging local SGS kinetic energy and dissipation, the model predicts the frequency and probability of pressure excursions beyond what the resolved field can provide. Even on coarse and medium grids, this statistical SGS approach yields a marked correction: with a constant nuclei distribu-

tion ($CN(R_0)$)), inception thresholds increase to 123.07 kPa (coarse) and 134.33 kPa (medium), capturing a major portion of the gap between resolved LES predictions and experiment.

The predictive fidelity is further enhanced by adding a dynamic nuclei transport model ($DN(R_0)$), which accounts for spatial nuclei clustering in low-pressure or high-shear flow regions—especially within vortex cores, where inception is most likely. Results are also interpreted using two inception detection thresholds: $f_{\text{det}} = 0.1$ kHz (visual: sensitive only to frequent, macroscopic events) and $f_{\text{det}} = 0.001$ kHz (acoustic: sensitive to rare, isolated events). This distinction is critical for interpreting both simulation and experiment, as cavitation events detected acoustically may occur at much lower frequencies and higher thresholds than those observable by eye or camera.

With this integrated modeling approach, the predicted inception pressures, with f_{det} of 0.1 & 0.001 Hz, increases to 141.24 kPa & 146.4 kPa, respectively, on the coarse grid. On the medium grid, it increases to 138.39 kPa & 144.15 kPa. Fine grid results show further convergence with the experimental window, reaching 142.87 kPa & 149.21 kPa. For comparison, Madabhushi and Mahesh (2025)'s DNS-like result of 152.83 kPa (for -1.5° AOA) confirms the physical upper bound for inception pressure at nearly full resolution. All model predictions for acoustic detection ($f_{\text{det}} = 0.001$ Hz) cluster tightly around or within the experimentally measured inception range of 135–155 kPa.

Table 7.3 compiles these results and affirms the critical advantages of the dynamic SGS and nuclei transport models: they bridge the gap between computationally practical LES and the physics of intermittent, high-risk cavitation inception. By explicitly reconstructing the effect of subgrid-scale turbulence and locally evolving nuclei populations, this approach provides accurate inception risk prediction at grid resolutions far coarser than DNS, making it a robust, reliable, and computationally tractable tool for engineering simulation.

| | Resolved LES | $CN(R_0)$ | $DN(R_0)$ | $DN(R_0)$ |
|-------------------------------|--------------|-----------|-----------|-----------|
| | | | 0.1Hz | 0.001Hz |
| Coarse | 97.29 | 123.07 | 141.24 | 146.4 |
| Medium | 103.91 | 134.33 | 138.39 | 144.15 |
| Fine | 115.39 | 135.58 | 142.87 | 149.21 |
| Madabhushi and Mahesh (2025)* | 152.83 | - | - | - |

Table 7.3: Inception pressure as predicted by LES and the inception model (with and without dynamic nuclei distributions) at different grid resolutions for -4° AOA. Values are in kPa.* Madabhushi and Mahesh (2025) had a different AOA of -1.5° .

Effect of angle of attack

AOA is a primary determinant of lift and drag on a hydrofoil, and it directly influences the characteristics of the vortices shed from its tips. For a single hydrofoil, increasing the AOA generally increases the lift, which, by the Kutta-Joukowski theorem, is directly related to the circulation (Γ) of the shed vortex. A higher AOA leads to stronger vortices with greater circulation. This increased circulation results in a larger pressure drop within the vortex core, as described by the relationship between circulation and pressure in an idealized vortex model. Consequently, a higher AOA makes a single vortex more susceptible to cavitation inception.

In a twin hydrofoil configuration, the interaction between the two vortices is a complex, three-dimensional process that is highly dependent on their initial strength, size, and relative positions.

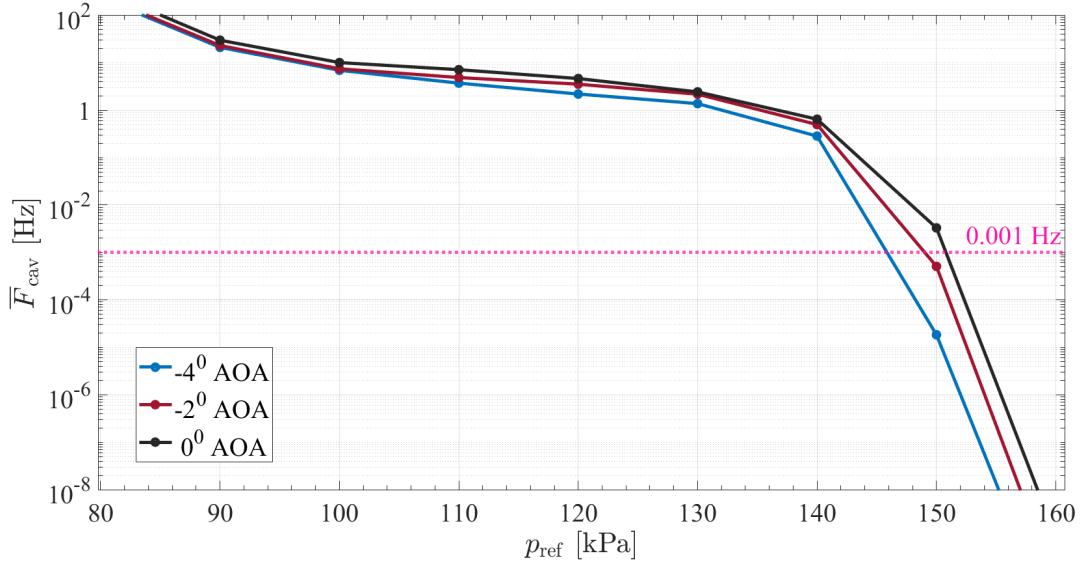


Figure 7.8: Average cavitaion event frequency for different AOA with the dynamic nuclei size distribution.

As the angle of attack becomes more negative, the flow faces increased resistance, resulting in higher shear to initiate the vortices. this also leads to a larger initial separation between the two vortices. While this greater distance may initially seem to imply a weaker interaction, it also provides a longer streamwise distance for the Crow instability to develop. This extended evolution allows for a more prolonged and intense process of vortex stretching and tilting, which significantly amplifies the vorticity and causes a dramatic reduction in core pressure.

The effect of AOA was assessed only using the Coarse grid. Table 7.4 outlines how cavitation inception pressure changes with the AOA. With resolved pressure from LES, AOA of 0° gives the maximum inception pressure while AOA of -4° gives the minimum inception pressure. This trend is in strong agreement with experimental observations, which indicate higher inception pressures under conditions that lead to more vigorous vortex interactions. SGS inception model with dynamic nuclei distributions accurately captures this underlying physics. It predicts a continuous increase in inception pressure as the AOA becomes more negative. For the $f_{\text{det}} = 0.001$ Hz criterion, it corrects for about ~ 50 kPa in the SGS, from an inception pressure with resolved LES of 97.29 kPa to 146.4 kPa at -4° AOA. Similar trend is observed for the -2° and 0° AOA with the model predicted inception pressure of 148.71 and 150.95 kPa, respectively. The significant correction provided by the SGS model shows its value in providing accurate, physically-grounded predictions on coarser grids. By correctly identifying that a more negative AOA leads to more intense downstream interactions and a greater propensity for inception, the model provides valuable insights that are not accessible through traditional methods.

7.5 Discussion

Predicting cavitation inception in complex turbulent flows has long challenged fluid mechanics due to the intricate and multiscale interplay among coherent vortical structures, unresolved small-scale turbulence, and the evolving distribution of vapor nuclei. Traditional modeling approaches, including Reynolds-averaged and even LES, often struggle because they do not capture the intermittent extreme pressure fluctuations responsible for vapor nucleation or neglect the effects of local nuclei

| AOA | Resolved LES | $DN(R_0) : 0.001Hz$ | Experimental (1 Hz - 0.1 Hz) |
|------------|--------------|---------------------|------------------------------|
| 0° | 105.13 | 150.95 | 140 – 160 |
| -2° | 99.53 | 148.71 | 110 – 145 |
| -4° | 97.29 | 146.4 | 95 – 135 |

Table 7.4: Inception pressure as predicted by LES and the inception model (with dynamic nuclei distributions) at different AOAs for the coarse grid. Values are in kPa.

transport and accumulation. Recognizing these limitations, we have developed a unified LES framework that integrates a dynamic k -equation SGS turbulence closure, a statistical cavitation inception model based on probability of pressure fluctuations, and a dynamic nuclei transport equation. This novel combination addresses all key physics underlying inception: it captures both the intensity and the statistical frequency of extreme pressure excursions at subgrid scales, while simultaneously tracking the spatially and temporally varying concentration of nuclei that act as cavitation seeds.

The effectiveness of this dynamic LES framework is demonstrated in the canonical twin hydrofoil problem - a configuration in which two vortices of differing strengths, generated by staggered foils, interact and stretch downstream in a way representative of many engineering flows prone to cavitation. The model resolves large-scale coherent vortex structures and their interactions, and reconstructs the crucial, unresolved turbulence-driven pressure reductions using SGS statistics. By directly simulating the advection and turbulent dispersion of nuclei, the framework identifies where they cluster within low-pressure vortex cores, which are revealed as focal points for cavitation onset. This approach avoids the limitations associated with uniform nuclei assumptions or minimum-pressure criteria, which often poorly estimate inception thresholds and fail to localize likely event sites.

Validation against experimental data illustrates the dynamic and robust capability of the framework. Across a range of practical grid resolutions, the unified LES-SGS approach consistently reproduces observed inception pressures and event locations, bridging the gap between computational feasibility and physical fidelity. Even when using coarser grids that underestimate the true range of pressure fluctuations, the SGS model reconstructs the frequency and spatial location of cavitation inception with notable accuracy. The inclusion of a dynamic nuclei transport model adds further realism, capturing the way nuclei accumulate in stretched vortex regions under both turbulent diffusion and local pressure gradients. This integrated modeling strategy not only achieves good agreement with experiments, but also exposes the underlying flow physics: stretches and interactions of primary and secondary vortices, enhanced by the Crow instability, generate the transient and acute pressure minima that govern when and where cavitation emerges. The framework's consistent prediction of inception across different AOA proves its robustness. In summary, by explicitly capturing both resolved and subgrid turbulence, dynamically evolving nuclei fields, and their statistical coupling, the proposed LES framework shows promise in providing accurate, efficient, and physically consistent predictions for complex engineering-relevant flows.

References

M. H. Bappy. *Statistical modeling of cavitation inception*. PhD thesis, The University of Iowa, 2022.

M. H. Bappy, P. M. Carrica, A. Vela-Martín, L. S. Freire, and G. C. Buscaglia. Pressure statistics of gas nuclei in homogeneous isotropic turbulence with an application to cavitation inception. *Physics of Fluids*, 32(9), 2020.

References

M. H. Bappy, P. M. Carrica, J. Li, J. E. Martin, A. Vela-Martin, L. S. Freire, and G. C. Buscaglia. A sub-grid scale cavitation inception model. *Physics of Fluids*, 34(3), 2022.

F. L. Brandao and K. Mahesh. Large-eddy simulation of cavitation inception in a shear flow. *International Journal of Multiphase Flow*, 146:103865, 2022.

C. E. Brennen. *Cavitation and Bubble Dynamics*. Cambridge University Press, 2014.

X. Chai and K. Mahesh. Dynamic-equation model for large-eddy simulation of compressible flows. *Journal of Fluid Mechanics*, 699:385–413, 2012.

N. A. Chang, J. Choi, R. Yakushiji, and S. L. Ceccio. Cavitation inception during the interaction of a pair of counter-rotating vortices. *Physics of Fluids*, 24(1), 2012.

S. C. Crow. Stability theory for a pair of trailing vortices. *AIAA Journal*, 8(12):2172–2179, 1970.

M. Germano, U. Piomelli, P. Moin, and W. H. Cabot. A dynamic subgrid-scale eddy viscosity model. *Physics of fluids a: Fluid dynamics*, 3(7):1760–1765, 1991.

A. Gnanaskandan and K. Mahesh. A numerical method to simulate turbulent cavitating flows. *International Journal of Multiphase Flow*, 70:22–34, 2015.

E. N. Harvey, W. D. McElroy, and A. Whiteley. On cavity formation in water. *Journal of Applied Physics*, 18(2):162–172, 1947.

W. J. Horne and K. Mahesh. A hardware accelerated unstructured overset method to simulate turbulent fluid flow. *Journal of Computational Physics*, 444:110574, 2021.

J. Katz. Cavitation phenomena within regions of flow separation. *Journal of Fluid Mechanics*, 140:397–436, 1984.

R. T. Knapp, J. W. Daily, and F. G. Hammitt. *Cavitation*. McGraw-Hill, 1970.

D. Knister. *Experimental Study of Cavitation Inception During Vortex-Vortex Interaction and Investigation of Constrained Volumetric Velocimetry Methods*. PhD thesis, The University of Michigan, 2024.

D. Knister, E. Callison, H. Ganesh, and S. L. Ceccio. Experimental study of cavitation inception during the interaction of vortices. In *Proc. 33rd Symp. on Naval Hydrodynamics*, Osaka, Japan, 2020.

T. J. Leasca, T. B. Kroll, and K. Mahesh. Large-eddy simulation of the tip vortex flow in a ducted propulsor. *Journal of Fluid Mechanics*, 1010:A51, 2025.

T. Leweke, S. Le Dizès, and C. H. Williamson. Dynamics and instabilities of vortex pairs. *Annual Review of Fluid Mechanics*, 48(1):507–541, 2016.

J. Li and P. M. Carrica. A population balance cavitation model. *International Journal of Multiphase Flow*, 138:103617, 2021.

Q. Li, S. Zhang, M. Wang, C. Zong, X. Li, and X. Song. Numerical and experimental analysis of cavitation characteristics in safety valves of the nuclear power second circuit using a modified cavitation model. *Engineering Applications of Computational Fluid Mechanics*, 17(1):2251546, 2023.

D. K. Lilly. A proposed modification of the Germano subgrid-scale closure method. *Physics of Fluids A: Fluid Dynamics*, 4(3):633–635, 1992.

A. Madabhushi and K. Mahesh. Cavitation inception mechanisms during the interaction between a pair of counter-rotating vortices. *Journal of Fluid Mechanics (Submitted)*, 2025.

K. A. Morsch. Cavitation nuclei and bubble formation—a dynamic liquid-solid interface problem. *Journal of Fluids Engineering*, 122(3):494–498, 2000.

S. B. Pope. Turbulent flows. *Measurement Science and Technology*, 12(11):2020–2021, 2001.

P. G. Saffman. Vortex dynamics. In *Theoretical Approaches to Turbulence*, pages 263–277. Springer, 1992.

References

Y. Shen and P. Dimotakis. Viscous and nuclei effects on hydrodynamic loadings and cavitation of a NACA 66 (MOD) foil section. *Journal of Fluids Engineering*, 111(3):306–316, 1989.

T. Smith, A. Grech La Rosa, G. Piggott, J. Gaivota, and S. McMorran. An experimental study of underwater radiated noise from a small vessel with damaged and fouled propellers. *JASA Express Letters*, 4(12), 2024.

H. Tennekes and J. L. Lumley. *A first course in turbulence*. MIT press, 1972.

A. Verma, N. Park, and K. Mahesh. A hybrid subgrid-scale model constrained by reynolds stress. *Physics of Fluids*, 25(11), 2013.

Understanding the Drag Production Mechanisms of Soft Biofilms using Live Species and Synthetic Replicates

RODRIGO VILUMBRALES-GARCIA & HARISH GANESH

Predicting drag penalties due to soft fouling on vehicles of naval interest remains a significant challenge. Several factors, such as types and distributions (coverage) of species, the growth duration, the nature of the resulting fouling morphology (presence of streamers, compliance, etc.), and sloughing with time, need to be considered to develop models. However, these factors make fundamental studies challenging in a controlled laboratory setting. Our effort encompasses laboratory studies of drag production mechanisms and growth of both live fouling at NRL-Key West and synthetic surfaces developed at the University of Michigan in collaboration with Prof. Jon Estrada. Specifically, the relationship between artificial surface properties and coverage is studied to understand the effect of non-uniform growth. Measurements using live fouling are used to understand the relationship between coverage, species concentration, and drag. The results of both aspects of the study will be combined to improve our understanding of the drag produced by the soft films.

Keywords: biofouling, skin-friction drag, fouling growth

8.1 Introduction and motivation

The soft and hard biofilms that form in marine vessels and vehicles produce undesirable effects such as increased skin friction drag, and degradation of material and performance, as seen in previous studies (Schultz and Swain, 2000; Andrewartha et al., 2010; Schultz et al., 2015; Hartenberger et al., 2020). Conducting experiments to analyze the drag production of surfaces covered with biofouling presents many challenges. The time-varying nature of the problem, which involves the growth of live specimens, implies a high degree of uncertainty, making repeatable experiments difficult.

A previous study by Hartenberger et al. (2020) discusses the challenges associated with studying drag produced by soft fouling in a laboratory setting, such as nonuniform growth and sloughing. For example, figure 8.1 shows various stages of drag measurements involving a flat plate covered with algae (Hartenberger et al., 2020). Several plates were grown for various durations of time, and images were acquired before and after the experiments were conducted. The after cases show a noticeable amount of sloughing, caused by detachment of the biofilm due to the shearing of the flow. Although this is an expected phenomena, it limits any forces or drag measurements to one trial since sloughing results in a different roughness configuration. This results in greatly limiting the repeatability of the study and hence preventing a general understanding of the causes behind drag production. Similarly, experiments conducted at NRL-Key West (Chapter 9) reveal a large spread in drag production of live biofilms grown under similar conditions. This variance in the drag measurement and growth of live biofilms under theoretically similar conditions makes the development of a general understanding of the mechanisms linking biofilms and drag, and the development of models able to predict it challenging. The main objective of the present study is

Understanding the Drag Production Mechanisms of Soft Biofilms using Live Species and Synthetic Replicates

to address the challenges presented above, specifically identifying factors that contribute to the observed drag measurement spread of live fouling, by considering three aspects of the problem. These are,

-First, we identify several synthetic materials to mimic live fouling. By using components derived from actual biofouling, such as Carrageenan or Alginate, or compliant textures like PDMS, we study the impact of mechanical properties (compliance) on drag production for various fouling-like materials.

-Second, we address the variation in drag measurements due to coverage and distribution. By varying the arrangement of synthetic biofouling test substrates, we study the effect of the total coverage area and distribution.

-Third, we leverage our current collaboration with NRL-Key West to perform growth and drag studies on live fouling found in Key West, Florida. We accomplish our goal by using a turbulent channel flow setup designed for live fouling and by performing controlled growth studies using a facility designed to study the effect of shear, temperature, sunlight, and nutrients.



Figure 8.1: Photographs of biofilms grown on acrylic test surfaces are recorded before, during, and after experimental trials at the skin friction flow facility (SF3). The four image pairs are overhead photos of biofilms grown for three different nominal incubation times (three, five and ten weeks) recorded before and after a trial. Two different three week panels are shown, displaying the large variety of growth for the shortest incubation time studied. The panels are 1.14 m in length and flow moved from left to right during growth and trials. From Hartenberger et al. (2020).

8.2 Developing synthetic materials to replicate biofouling

As discussed before, experiments involving the drag production of biofilms are challenging, mainly due the lack of repeatability caused by non-uniform growth at nominally equal conditions and sloughing. To overcome this, we develop synthetic materials capable of replicating the effects of the live specimens. The samples have been developed in collaboration with Prof. Jon Estrada at the University of Michigan. The advantage of this approach relies on the capability to carefully control

Understanding the Drag Production Mechanisms of Soft Biofilms using Live Species and Synthetic Replicates

the mechanical properties of the artificial samples, and their distribution over the surface (both arrangement and coverage area). An expected outcome is to isolate the effect of material properties and/or distribution on the drag measurement spread observed in live fouling experiments at the same conditions. Three materials have been explored for the artificial surfaces: carrageenan-based hydrogels, alginate-based hydrogels and polydimethylsiloxane (PDMS).

Carrageenan-based hydrogels

Carrageenan is a natural component extracted from seaweed that can produce water-based hydrogels (Therkelsen, 1993). By controlling the concentration of the material in the dissolution and immersion time in water, hydrogels with various strengths can be extracted. Figure 8.2 presents a set of mechanical properties tests conducted for carrageenan-based hydrogels of concentration ranging from 1 to 2.5% and submersion time ranging from 0 to 48h. The samples are made by mixing water and carrageenan at constant stirring (300 rpm) and 90°C for 3 hours, with an extra hour under no stirring for the mixture to set (while keeping the temperature). Then, the samples are cast into molds and left for a day at 4°C. The tests are conducted using dynamic mechanical analysis (DMA), and consist of steady and dynamic cases. The steady set involves compression at a constant velocity, where the deformation is recorded (left column, Figure 8.2). The second set corresponds to a dynamic amplitude sweep test, where the deformation is applied in a sinusoidal form (compression-relaxation), and the amplitude is increased under a constant frequency. The results suggest that larger concentrations of the sample provide stronger hydrogels, as observed by the steepness of the deformation versus stress curve. Water submersion time weakens the samples and reduces their mechanical strength.

Although repeatable samples were obtained, testing at the flow tunnel facility showed a large degradation of the hydrogels at low velocities, with the samples losing more than half of the volume after less than 15 minutes of continuous flow velocity at less than 2.5 m/s. High-speed recorded videos suggest that the samples are being peeled off at a constant rate by the forces exerted by the flow. This undesired degradation prevents us from using carrageenan as a candidate for repeatable drag studies, but opens the door to sloughing analysis, since the degradation can be controlled by tuning the properties of the original synthetic biofouling.

Alginate-based hydrogels

Alginates are polysaccharides isolated from brown algae such as *Laminaria hyperborea* and *Lessonia* found in coastal waters around the globe (Augst et al., 2006). Our samples are developed using the double crosslinking method proposed by Ji et al. (2022), as depicted in figure 8.3. First, a weak hydrogel is formed with a mixture of sodium alginate with a total weight concentration of 2% and a CaSO_4 solution of 0.3%. The mixture is stirred at 300 rpm at room temperature for a duration of 3 hours and then placed in vacuum to remove the air. This procedure helps to hold the shape of the hydrogel, and creates an initial weak specimen, with a semi-transparent appearance, as depicted in figure 8.3. After curing at 4°C for one day, the samples are immersed in a CaCl_2 bath with a molarity of 0.1M for another day to enhance the mechanical strength of the samples. The differences after the cross-linking method are highlighted by the amount of calcium in the overall hydrogel, represented by the white color, in contrast to the transparent appearance of the original surface. The external crosslinking bath greatly enhances the mechanical capabilities of the hydrogel (Ji et al., 2022), and preliminary testing has shown that the material is able to stay intact for more than 30 minutes when exposed to constant flow velocity of 5 m/s. We plan on performing

Understanding the Drag Production Mechanisms of Soft Biofilms using Live Species and Synthetic Replicates

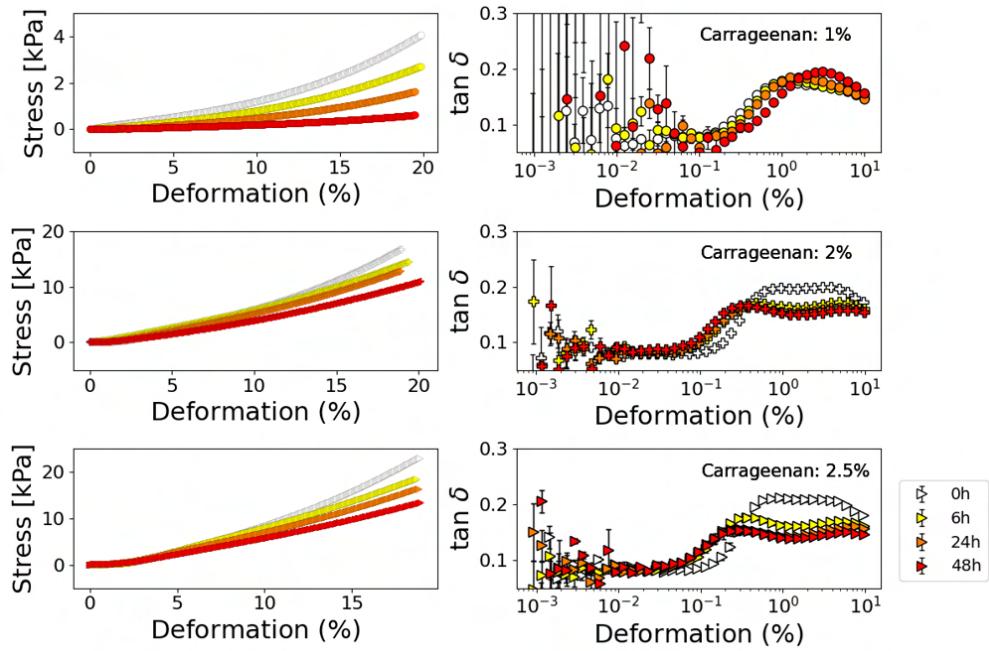


Figure 8.2: Mechanical properties analysis of the carrageenan-based samples. Left column presents the steady testing, and right column introduces the dynamic sweeps.



Figure 8.3: Procedure followed to make alginate synthetic biofilms. Top left shows the first step of the process, an internal crosslinking of sodium alginate with CaSO_4 . Top right displays the results after the internal crosslinking with a CaCl_2 bath. Bottom figure shows the samples already installed in the flat plate.

skin-friction measurements for a range of material properties and distributions in the upcoming months.

PDMS

The last artificial component developed is based on Polydimethylsiloxane (PDMS). This type of component has been used to replicate flexible and compliant materials, and by properly tuning its components, can provide a vast range of mechanical properties (Palchesko et al., 2012; Wang et al., 2020). Although this element does not have an organic base, its compliant-like properties allow us to use it for our study. Several blends of the material will be used, and the effects of various mechanical properties will be analyzed.

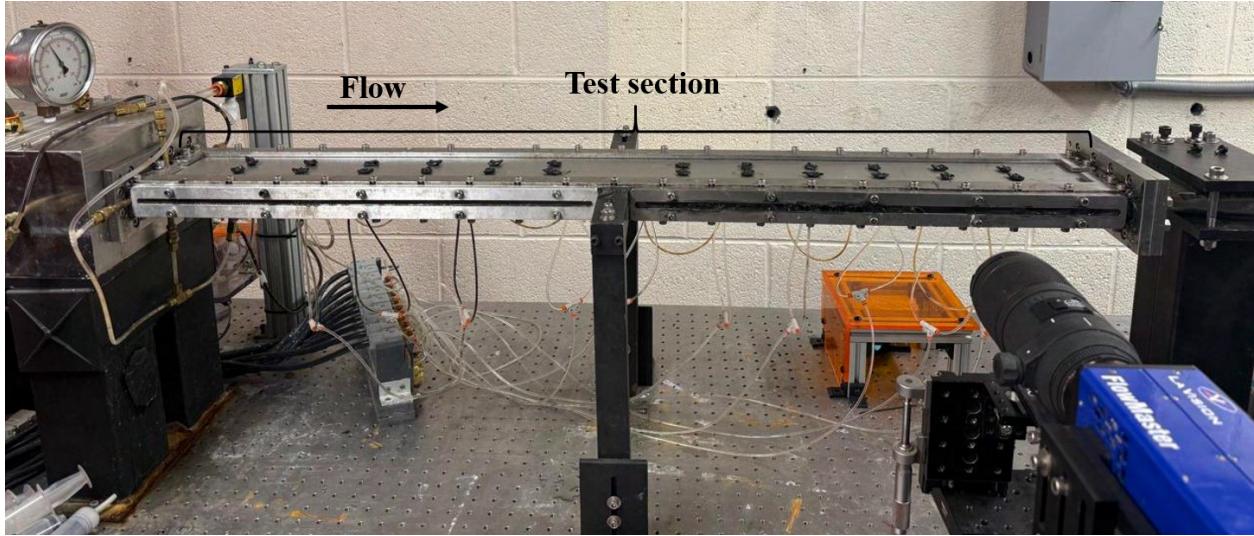


Figure 8.4: The Skin Friction Flow Facility (SF3).

8.3 Effect of surface distribution and area coverage

Measuring or estimating the drag produced by biofouling is a challenging problem. Live organisms tend to grow in heterogeneous arrangements, with varying spatial and height distribution. In such cases, classical methods of predicting the drag caused by a surface with homogeneous roughness, such as the equivalent sand grain roughness, need to be revised. Although work has been done to account for heterogeneity, both in the spanwise and streamwise axis (Hutchins et al., 2023; Frohnapfel et al., 2024), the vast majority of the previous studies consider hard biofouling-like components, such as sand paper. When it comes to soft biofouling, the compliant nature of the specimens, or the presence of filaments, can make classical approaches not reliable.

To better advance our understanding of the latter, we propose a systematic approach that will use the synthetic materials developed before to build plates mimicking the growth of fouling (informed by experiments at Key West) in a controlled manner. By using artificial samples, we can tackle the two main problems presented before, repeatability and growth heterogeneity. Here, we start by conducting drag measurements on flat plates with spanwise heterogeneity made of strips of sandpaper. We analyze the effects of varying the area coverage, and the distribution arrangement around the plate. Next, we will present preliminary results involving the use of compliant, soft materials.

Skin-friction flow facility (SF3)

All the experiments presented here have been conducted at the University of Michigan Skin Friction Flow Facility (SF3). The SF3, shown in figure 8.4 is a high-aspect ratio, closed channel flow facility with fully developed turbulent flow in the test-section, capable of reaching $Re_H \sim 150,000$. The SF3 channel gap height, H_0 , is nominally 7 mm, span is 101.6 mm, and has a downstream length of 1.14 m. The SF3 measures the skin-friction C_f of surfaces based on the pressure drop across the 10 pressure ports (located at 20H, 30H, 50H, 70H, 84H, 98H, 112H, 126H, 140H, and 150H) along the test-section as shown in figure 8.4. The flow is tripped after passing through honeycombs and mesh screens at the entrance to the channel using an 80 grit sandpaper strip. Only measurements acquired in the fully developed turbulent flow region (between 70H and 150H) are used for the pressure gradient estimation. The volumetric flow rate, Q , was measured using a Siemens Sitrans

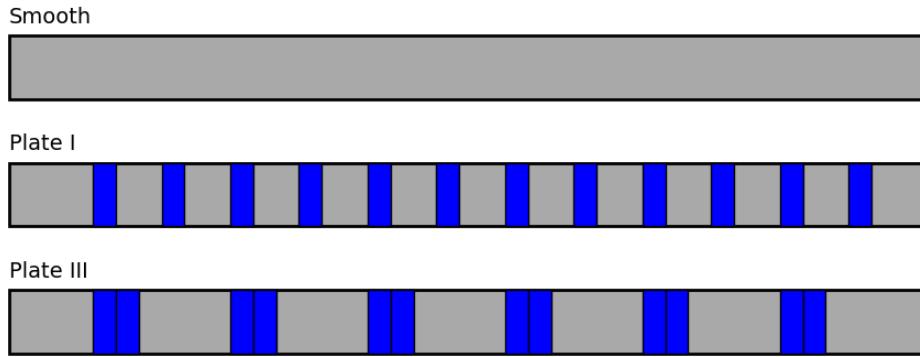


Figure 8.5: Plate configurations considered in this study. Plate I and Plate III are examples of two arrangements with a similar coverage area. Blue areas represent the location of the roughness strips.

FM MAG5100 magnetic flow meter. The temperature is also measured using an RTD and logged during data acquisition. Density and kinematic viscosity are obtained as functions of the logged temperature following the ITTC freshwater tables. The measurements at each step are averaged after trimming 5 seconds post and pre transition to another step. The averaged measurements along with the known streamwise port location are used to estimate the pressure gradient dp/dx used in the estimation of the shear stress at the wall. An accurate estimation of the shear stress relies upon the correct dp/dx , and height measurements. The bottom window of the SF3 used in conjunction with the fur surfaces is a hydrodynamically smooth panel with an optically clear fused quartz insert for PIV laser illumination. Simultaneous PIV measurements are acquired during the pressure data acquisition. The velocity profiles enable the extraction of the shear stress on the smooth panel as well as the correct channel height using the method detailed below.

As discussed before, arrangement and coverage of rough and biofilm-like elements on a flat plate can play a major role on its drag performance. To further analyze it, we start by placing hard-like biofouling in the shape of sand paper on a flat plate under different arrangement, in the configurations depicted in figure 8.5.

Preliminary results: Two configurations have been studied so far. The first is Plate I, where strips of 80-grid sandpaper are distributed perpendicular to the flow, and equally spaced in the streamwise direction. The coverage is varied from 0% (control case, or smooth), to 100%. We observe a non-linear evolution of the drag increase, with a peak penalty being reached for a coverage of around 20%. Since the strips are protruding into the flow, we hypothesize that the constant obstacles encountered by the water as it advances across the plate can be responsible for such behavior. To evaluate the arrangement effects, we prepare a second plate, where the thickness of the strips is doubled. By doing so, we reduce the number of strips by half for a given coverage area, hence diminishing the number of obstacles encountered by the flow. This new configuration promotes a drag increase for a comparable coverage ratio of 30% less than Plate I, highlighting the relevance of distribution on drag performance. To further assess these effects, further studies will analyze similar configurations, but with the strips being flushed with the surface.

When it comes to soft biofouling, preliminary data presented in figure 8.7 shows that a flat plate covered with circular PDMS samples flush with the surface, for a total coverage area of 10%, increases the drag with respect to a flat plate by around 7%. Although only one test plate has been studied so far, it suggests that the presence of compliant-like patches on a flat plate can increase its drag, even if the mentioned biofilm-like surfaces are flushed with the surface. Further studies will involve the use of several blends of PDMS (to ensure various levels of compliance), and alginate biofilms under different homogeneous and heterogeneous arrangements around the plate, coverage

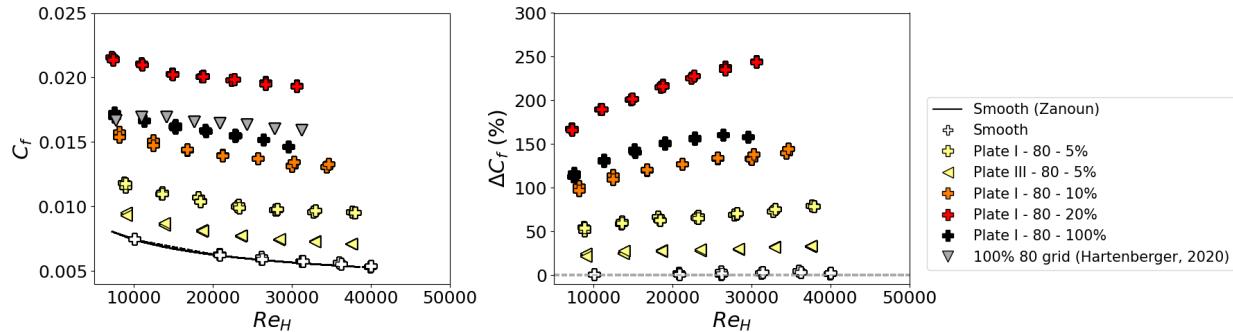


Figure 8.6: Coefficient of friction vs Re_H and drag penalty for plates with strip-like roughness obstacles.

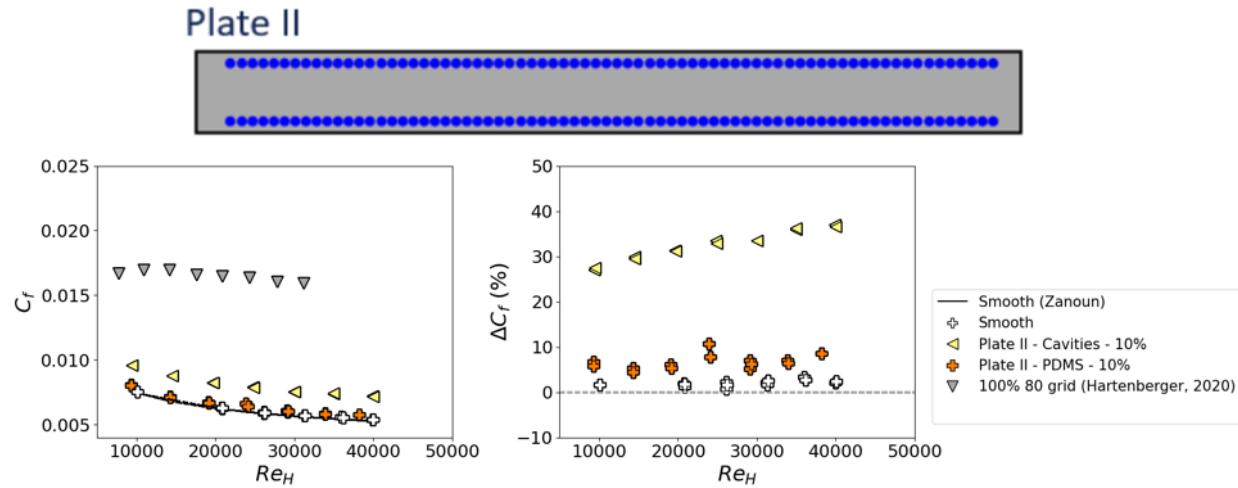


Figure 8.7: Schematic of Plate II covered with soft-like biofilms, along with coefficient of friction and drag penalty vs Re_H for plates with artificial biofouling with PDMS base.

areas, and protrusion heights.

8.4 Controlled growth of live biofilms

Environmental and flow conditions are critical for the growth of biofilms. Aspects such as light, temperature, or flow conditions can largely impact the type and density of the biofouling community that grows on a given surface. To further study this aspect, a growth-loop has been designed, built, and installed at the Naval Research Facility in Key West. The design of the facility is depicted in figure 8.8.

The growth facility intakes water from the ocean in an open-loop configuration and provides enough flow to feed three chambers. Each of those chambers is independent, and its growing conditions can be fully controlled without affecting the others. This opens the possibility of conducting a parametric study to assess the impact of aspects such as sunlight or flow velocity have on the growth of fouled surfaces. The velocity at each of the modules is controlled by a flow meter, and can be varied using a set of inlet valves. Each of the growth chambers is able to hold 4 flat plates, and is designed with a transparent acrylic cover to allow sunlight into them.

Preliminary findings shown in figure 8.8 show a substantial growth of algae after a period of 8 days, with constant velocity and sunlight exposure. Further work will include an analysis of

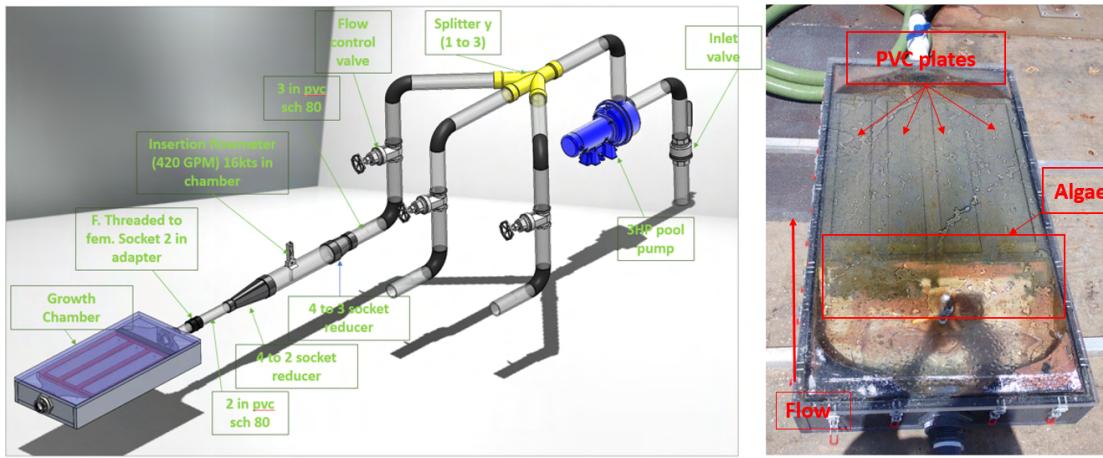


Figure 8.8: Growth loop design and schematics (left). Detail of one of the growth modules installed at NRL Key West after 8 days of growing time. Algae communities are observed both in the module and on the plates.

the colonies on the plate via DNA sequencing, 3D scans of the surface to assess arrangement and coverage, and drag studies to better understand the mechanisms of drag-producing biofilms.

Acknowledgments

The study at NRL-KW (growth characteristics, section 8.4) is supported by ONR under Grant N000142512263. NRL-KW is supported by the Naval Research Laboratory under Contract No. N00173-23-C-2018. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Research Laboratory.

References

- J. Andrewartha, K. Perkins, J. Sargison, J. Osborn, G. Walker, A. Henderson, and G. Hallegraeff. Drag force and surface roughness measurements on freshwater biofouled surfaces. *Biofouling*, 26(4):487–496, 2010.
- A. D. Augst, H. J. Kong, and D. J. Mooney. Alginate hydrogels as biomaterials. *Macromolecular Bioscience*, 6(8):623–633, 2006.
- B. Frohnapfel, L. von Deyn, J. Yang, J. Neuhauser, A. Stroh, R. Örlü, and D. Gatti. Flow resistance over heterogeneous roughness made of spanwise-alternating sandpaper strips. *Journal of Fluid Mechanics*, 980:A31, 2024.
- J. D. Hartenberger, E. G. Callison, J. W. Gose, M. Perlin, and S. L. Ceccio. Drag production mechanisms of filamentous biofilms. *Biofouling*, 36(6):736–752, 2020.
- N. Hutchins, B. Ganapathisubramani, M. P. Schultz, and D. I. Pullin. Defining an equivalent homogeneous roughness length for turbulent boundary layers developing over patchy or heterogeneous surfaces. *Ocean Engineering*, 271:113454, 2023.
- D. Ji, J. M. Park, M. S. Oh, T. L. Nguyen, H. Shin, J. S. Kim, D. Kim, H. S. Park, and J. Kim. Superstrong, superstiff, and conductive alginate hydrogels. *Nature Communications*, 13(1):3019, 2022.
- R. N. Palchesko, L. Zhang, Y. Sun, and A. W. Feinberg. Development of polydimethylsiloxane substrates with tunable elastic modulus to study cell mechanobiology in muscle and nerve. *Public Library of Science One*, 7(12):e51499, 2012.
- M. P. Schultz and G. W. Swain. The influence of biofilms on skin friction drag. *Biofouling*, 15(1-3):129–139, 2000.

References

M. P. Schultz, J. M. Walker, C. N. Steppe, and K. A. Flack. Impact of diatomaceous biofilms on the frictional drag of fouling-release coatings. *Biofouling*, 31(9-10):759–773, 2015.

G. H. Therkelsen. Carrageenan. In *Industrial Gums*, pages 145–180. Elsevier, 1993.

J. Wang, S. S. Koley, and J. Katz. On the interaction of a compliant wall with a turbulent boundary layer. *Journal of Fluid Mechanics*, 899:A20, 2020.

Understanding the Drag of Soft Biofilms: A Multidisciplinary Experimental Approach

MOHAMMAD A. ELSOUHT¹, RODRIGO VILUMBRALES-GARCIA, VANESSA MOLINA²,
MATTHEW FIRST³, JASON LEE⁴ & HARISH GANESH

Biofilms forming on marine vehicles produce undesirable effects such as increased skin-friction drag, and material and performance degradation. In this study, we analyze the drag production and biological composition of native soft fouling species found in ambient water at the Naval Research Laboratory in Key West (NRL-KW). Drag measurements highlight a large drag production disparity for test cases with biofouling developed under similar conditions and growth times, which could be caused by growth heterogeneity. To further understand this, we develop synthetic materials to replicate and study biofouling in a controlled and repeatable way. By varying the distribution of these materials in flat plates, both in terms of area coverage and distribution, and the mechanical properties of the biofilm-like surfaces, we aim to obtain a better understanding on the key aspects of drag produced by biofilms.

9.1 Introduction

Soft and hard biofilms forming on marine vessels and vehicles produce undesirable effects such as increased skin-friction drag, and material and performance degradation, as seen in previous studies (Schultz and Swain, 2000; Schultz et al., 2015; Hartenberger et al., 2020; Andrewartha et al., 2010). Understanding the growth and the drag produced by soft films in a marine environment is a challenging multidisciplinary problem difficult to study in a laboratory setting. The interdisciplinary nature of the problem stems from the interaction of several features including the type and the distribution of species (diatoms, bacteria, etc.), their growth dependence on environmental factors (growth duration, flow conditions, temperature, salinity, substrate properties, etc.), the nature of the resulting fouling morphology (presence of streamers, compliance, etc.), and the drag produced by the resulting fouled surface. In addition to these challenges, growth nonuniformities and sloughing make repeatable drag measurements challenging thereby resulting in a lack of understanding of drag production mechanisms. Here, we present recent work toward understanding the growth and drag production mechanisms of live, soft biofilms and artificial biofilm-like surfaces. The drag and biological composition of native soft fouling species found in ambient water at the Naval Research Laboratory in Key West (NRL-KW) for a growth duration of two, eight, and twelve weeks will be presented. Trends, challenges, and lessons learned from measuring the drag of live fouling using a one-of-a-kind turbulent channel flow will be discussed. Findings from the drag measurement of several artificial fur-like and hydrogel-compliant surfaces will be presented and compared with live biofouling behavior. Finally, a strategy based on the findings of the two studies to understand the growth and drag of live fouling will be presented.

¹University of Michigan, Ann Arbor, MI

²Precise Systems Inc., 22290 Exploration Dr., Lexington Park, MD 20653

³Naval Research Laboratory, 4555 Overlook Avenue SW, Washington, DC 20375

⁴Naval Research Laboratory, 1009 Balch Blvd, Stennis Space Center, MS 39529

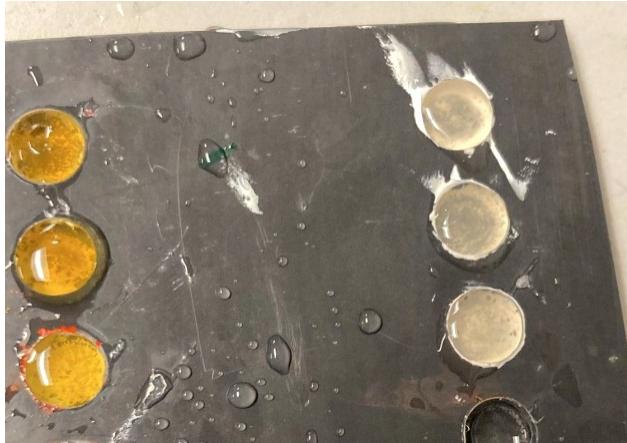


Figure 9.1: Example of the artificial biofilms. Flat plate with carrageenan hydrogel samples at different concentrations. Samples are colored for visualization purposes.

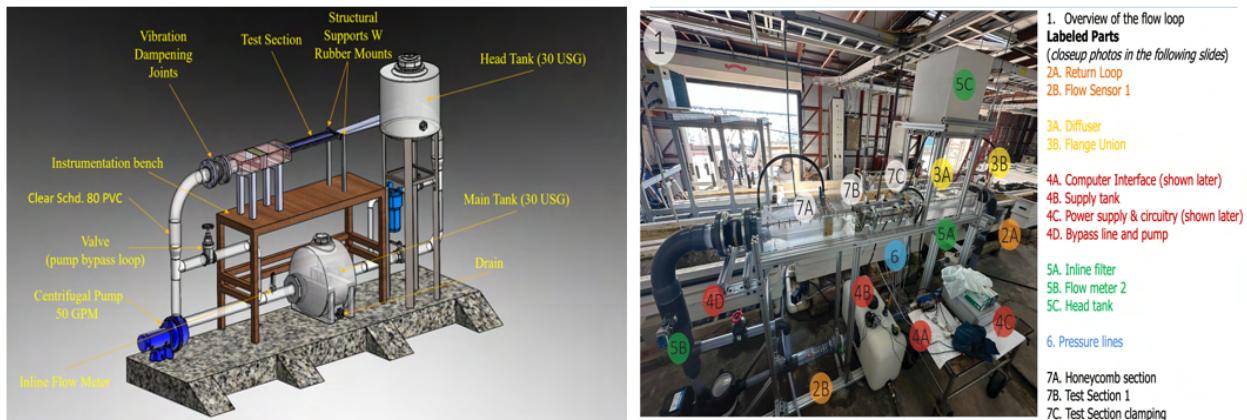


Figure 9.2: A CAD rendering of the BSVF illustrating various components. (b) Picture of the BSVF installed at NRL-KW. Several minor differences exist between the CAD model and the actual flow loop at NRL. The loop is constructed using transparent materials to facilitate unwanted fouling growth monitoring during and after testing. The loop is designed to handle freshwater and seawater with temperature control (cooling tank shown as 4B).

9.2 Approach

Flat plates in a flow-through ambient seawater trough were submerged for 12 weeks, and measurements were conducted at 3, 8, and 12 weeks to analyze the effects of biofilms on drag production. First, to analyze the effects of soft films on drag production, we developed hydrogels with Carrageenan and Calcium Alginate base at different concentrations. The mechanical properties of the samples were acquired using a Dynamic Mechanical Analyzer under steady and dynamic compression tests at several amplitudes and frequencies, and both in dry and wet conditions. Once the samples were properly characterized, we placed them on a flat plate under different surface distributions and coverages, aiming to analyze the effects of those parameters, together with mechanical properties and protruding height, on the drag characteristics of the elements. The final results will analyze surface arrangement, both homogeneous and heterogeneous, in increasing coverage ratios in an effort to replicate the growing effects of live biofilms. We will also force sloughing to better understand the real conditions experienced when conducting experiments with live colonies. An example of a plate with several carrageenan based artificial fouling samples is presented in Figure 9.1.

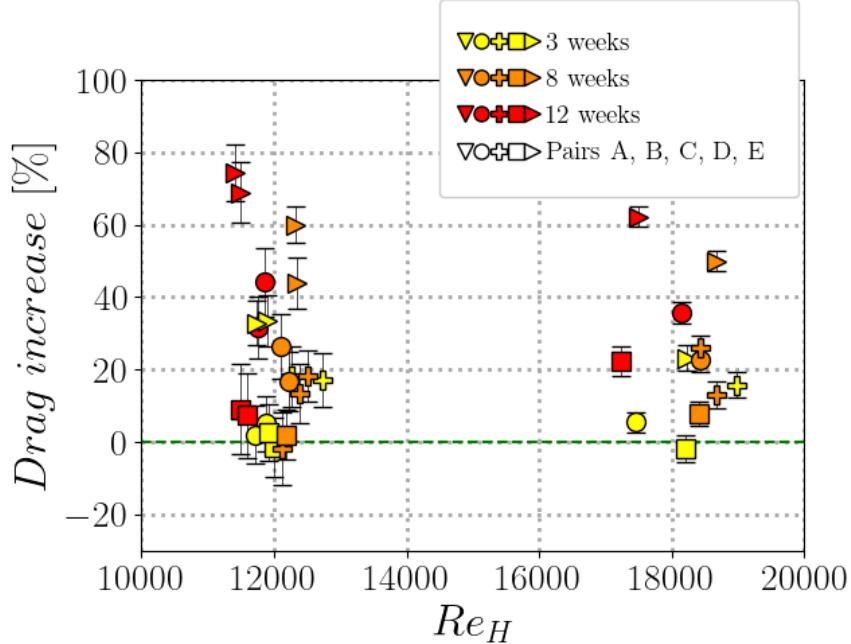


Figure 9.3: Drag increase vs Re (height based) for live biofilms at three growth stages. 3 Weeks (yellow), 8 weeks (orange), and 12 weeks (red). 5 different pairs of plates are tested for each growth to analyze repeatability.

A recirculating flow channel (BSVF at NRL-KW, and SF3 at University of Michigan) is used to assess the drag characteristics of live and artificial biofilms. The SF3 and BSVF measures pressure drop along the channel at eight locations in the streamwise direction: four in the development region and four in the developed region to estimate wall shear stress. Bulk velocity is estimated using a flow meter and in situ channel height measurements using a micrometer. Currently, the BSVF estimates C_f (skin-friction coefficient) of surfaces (smooth, rough, fouled) mounted on both the top and bottom sides of the fully developed channel flow. Using the pressure drop measurement, channel height, and mean velocity, the wall shear stress and the skin friction are estimated. Drag measurements from the BSVF have been validated against the expected drag of a smooth plate in a turbulent channel flow reported from the literature.

9.3 Results

Preliminary results of the drag measurements conducted on live biofilms are presented in Figure 9.3. We observe that growing time tends to increase the drag penalty (with respect to a smooth plate), but a large spread is observed between different pairs of plates grown for the same amount of weeks. For the highest Re considered, Pair D produced a drag increase of around 20%, while Pair E showed a much larger drag increase of 60%, highlighting the challenges of biofilm testing. Biological and genetic testing of the plates will be presented to provide addition information to this disparity, which could be caused from different surfaces arrangements or biological differences across the pairs of plates studied. We will also introduce the results obtained for the artificial biofilms and compare them with natural biofilms tested at NRL-KW.

Acknowledgments

The authors would like to acknowledge the DARPA- Arcadia program for the resources provided to perform this study. We would also like to thank the ONR-funded Center for Naval Research and Education (CNRE) at the University of Michigan for supporting Dr. Vilumbrales-Garcia.

References

J. Andrewartha, K. Perkins, J. Sargison, J. Osborn, G. Walker, A. Henderson, and G. Hallegraeff. Drag force and surface roughness measurements on freshwater biofouled surfaces. *Biofouling*, 26(4):487–496, 2010.

J. D. Hartenberger, E. G. Callison, J. W. Gose, M. Perlin, and S. L. Ceccio. Drag production mechanisms of filamentous biofilms. *Biofouling*, 36(6):736–752, 2020.

M. P. Schultz and G. W. Swain. The influence of biofilms on skin friction drag. *Biofouling*, 15(1-3):129–139, 2000.

M. P. Schultz, J. M. Walker, C. N. Steppe, and K. A. Flack. Impact of diatomaceous biofilms on the frictional drag of fouling-release coatings. *Biofouling*, 31(9-10):759–773, 2015.

Two-Way Coupled Simulation of Biofouling-Inspired Dynamic Roughness in Turbulent Channel Flow

DIVYANSHU GOLA, SOHAM PRAJAPATI¹ & KRISHNAN MAHESH

Simulation results are presented from a two-way coupled fluid–structure interaction (FSI) framework developed to examine the influence of wall-mounted deformable elements on turbulent channel flow, as a model for biofouled surfaces. The governing equations of the one-continuum formulation are solved using a finite volume-based, Euler–Euler approach, enabling simultaneous treatment of the fluid and hyperelastic solid phases. The channel walls are populated with a neo-Hookean cuboidal element of two different heights, one extending into the logarithmic region and the other into the outer layer of a turbulent channel flow at friction Reynolds number $Re_\tau = 180$. The simulations resolve the coupled dynamics between turbulent eddies and the deformable structures, highlighting modifications to near-wall vortical motions and element deformation. The results demonstrate the capability of the present framework to capture key mechanisms of turbulence–structure interaction in turbulent channel flows and provide a foundation for systematic extensions to arrays of elements and realistic biofouled wall conditions.

Keywords: fluid–structure interaction, turbulent channel flow, biofouling, neo-Hookean elements, one-continuum formulation

10.1 Introduction

Vessels that remain at sea for extended periods frequently experience the accumulation of a slimy layer on their hulls, which reduces performance by increasing drag and decreasing propulsion efficiency. This phenomenon, known as biofouling, can be broadly classified into hard (calcareous) and soft types. While hard biofouling has been extensively investigated through both experiments and simulations (Kaminaris and Balaras, 2024), soft biofouling presents greater challenges. These arise from the diversity of viscoelastic properties and constitutive models used to describe biofilms (Wilking et al., 2011; Estrada et al., 2018; Peterson et al., 2015), as well as the difficulties in producing laboratory analogues or in controlling the growth of real biofilms under experimental conditions (Tsagkari and Sloan, 2018). From a modeling perspective, two-way fluid–structure interaction (FSI) between fluids and viscoelastic solids emerges as a promising approach to address this problem.

Numerous studies have attempted to characterize the mechanical properties of biofilms and their response in shear flows. Biofilms are often modeled as viscoelastic solids consisting of bacterial organisms embedded in an extracellular polymeric matrix, with shear modulus (G) and viscosity (μ) as key parameters. These can be measured experimentally using oscillatory shear to extract storage and loss moduli (Stoodley et al., 1999). Theoretical frameworks often represent biofilms using spring–dashpot systems arranged in series and/or parallel, depending on the deformation

¹University of Minnesota, Minneapolis, MN

regime of interest (Peterson et al., 2015). Experiments further demonstrate that biofilms, being living systems, can actively adapt to applied shear, leading to phenomena such as strain stiffening and hysteresis (Rupp et al., 2005; Araújo et al., 2016).

In recent years, Eulerian-Eulerian two-way FSI methods have been developed in which the fluid and solid are treated as a single continuum, with the phases distinguished by a solid volume fraction. Within this framework, density, viscosity, stress, and velocity fields are defined monolithically and solved in a coupled manner (Prajapati et al., 2025). The Eulerian nature of the method is maintained by introducing additional advection equations for the solid fraction and the solid deformation tensor. Since these equations must resolve sharp phase interfaces, specialized numerical schemes such as geometric volume-of-fluid methods (Fakhreddine et al., 2024) and Weighted Essentially Non-Oscillatory (WENO) reconstructions (Liu et al., 1994; Jiang and Shu, 1996) are often employed. Using this approach, Sugiyama et al. (2011) validated different viscoelastic material models, while more recent studies have applied it to compliant-wall problems in turbulent channel flows (Esteghamatian et al., 2022; Prajapati et al., 2025).

Despite advances in numerical modeling of biofilms (García et al., 2020) and roughness effects associated with hard biofouling (Jensen and Forooghi, 2025), studies directly addressing the coupled interaction of turbulent flow with soft, biofilm-like structures remain limited. Experimental efforts in related directions include investigations of canopy-turbulence interactions by Park and Nepf (2025) as well as interaction of compliant wall with a turbulent boundary layer by Wang et al. (2020), but the fluid–biofilm problem remains largely unexplored.

The present study extends two-way FSI simulations toward realistic modeling of soft biofouling and aims to be the first step towards modeling turbulent flow over complex biofouled surfaces. The Euler-Euler method offers an accurate way to not only simulate the fluid-solid interaction, but also to analyse various quantitative metrics related to the solid itself, thereby presenting a strong case for its use in advanced FSI problems (multiple elements, real biofilm models, etc.). The results presented here are a brief testament to that potential. Specifically, we investigate the compliance of a cuboidal viscoelastic element attached to the wall of a turbulent channel flow at a friction Reynolds number $Re_\tau = 180$. The objectives of this work are to determine: (i) how structural compliance modifies the wake dynamics, (ii) the additional resistance or drag exerted by a compliant element, and (iii) how wall-mounted compliant obstructions alter the outer turbulent flow. The problem setup and methodology are presented in Section 10.2, results in Section 10.3, and discussion and outlook in Section 10.4.

10.2 Methodology

Problem setup

Figure 10.1 shows a schematic of the computational setup. A cuboidal element of height h and square cross-section side s is mounted on the bottom wall of a channel of height $2H$. The streamwise direction is denoted by x , vertical by y and spanwise by z . The channel Reynolds number is defined as

$$Re_{ch} = \frac{\rho^f U_b H}{\mu^f},$$

where ρ^f and μ^f are the fluid density and dynamic viscosity, and U_b is the bulk velocity obtained from a precursor periodic channel-flow simulation used to generate inflow conditions (see Domain and Boundary Conditions).

Each case is simulated in two stages: (i) a non-FSI stage, where the solid is rigid and the flow is allowed to develop until reaching statistical stationarity, and (ii) an FSI stage, where structural

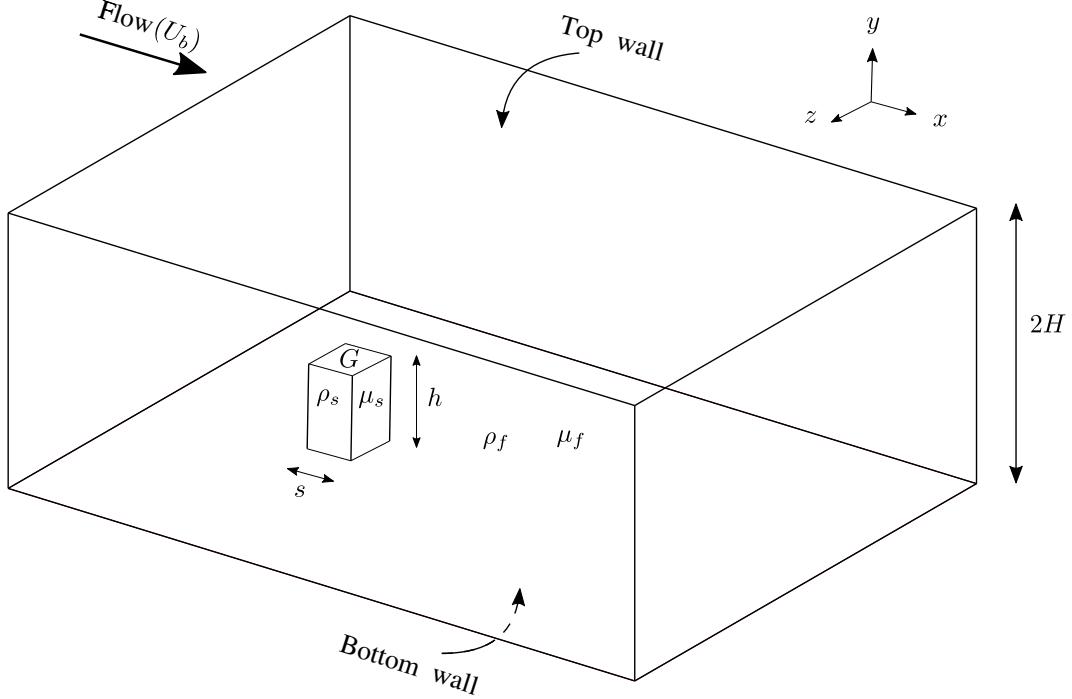


Figure 10.1: Schematic of the simulation setup.

compliance is activated. In the FSI stage, the solid is modeled as a neo-Hookean hyperelastic material (Sugiyama et al., 2011) characterized by shear modulus G , density ρ^s , and viscosity μ^s .

Governing equations

For non-FSI simulations, the solid region is masked from the fluid. For FSI simulations, a one-continuum formulation is employed by introducing a solid volume fraction $\phi \in [0, 1]$. The governing equations are the incompressible Navier–Stokes equations written for the mixture field:

$$\frac{\partial u_i}{\partial t} + \frac{\partial(u_i u_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \frac{1}{\rho} \frac{\partial \sigma_{ij}}{\partial x_j} + g_i, \quad (10.1)$$

with the incompressibility condition

$$\frac{\partial u_i}{\partial x_i} = 0. \quad (10.2)$$

The mixture density, viscosity, and velocity are defined as

$$\rho = \phi \rho^s + (1 - \phi) \rho^f, \quad \mu = \phi \mu^s + (1 - \phi) \mu^f, \quad u_i = \phi u_i^s + (1 - \phi) u_i^f.$$

The stress tensor is expressed as

$$\sigma_{ij} = \phi \sigma_{ij}^s + (1 - \phi) \sigma_{ij}^f, \quad (10.3)$$

where the fluid stress is purely viscous,

$$\sigma_{ij}^f = \mu^f \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad (10.4)$$

and the solid stress includes both viscous and hyperelastic contributions,

$$\sigma_{ij}^s = \mu^s \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + G (B_{ij} - \delta_{ij}). \quad (10.5)$$

Here, B_{ij} is the left Cauchy–Green deformation tensor, which evolves as (Hameduddin et al., 2018)

$$\frac{\partial B_{ij}}{\partial t} + u_k \frac{\partial B_{ij}}{\partial x_k} = B_{ik} \frac{\partial u_j}{\partial x_k} + B_{kj} \frac{\partial u_i}{\partial x_k}. \quad (10.6)$$

The advection of the solid volume fraction is given by

$$\frac{\partial \phi}{\partial t} + u_i \frac{\partial \phi}{\partial x_i} = 0. \quad (10.7)$$

Further details on the formulation and validation can be found in Prajapati et al. (2025).

Domain and boundary conditions

Two element aspect ratios are considered: $h/s = 2$ and 5. The domain dimensions are identical across the two cases:

$$0 < \frac{x}{s} < 30, \quad 0 < \frac{y}{s} < 20, \quad 0 < \frac{z}{s} < 10\pi,$$

with $H/s = 10$. Thus, the element tips extend respectively into the inner layer, log layer, and outer layer of the turbulent channel flow. The element is centered at $(x/s, z/s) = (10, 5\pi)$.

The channel Reynolds number is $Re_{ch} = 2800$, corresponding to an element Reynolds number $Re_s = 280$. The density ratio is fixed at $\rho^s/\rho^f = 1$, while the viscosity ratio is $\mu^s/\mu^f = 1000$ (Zhang and Bishop, 1994; Peterson et al., 2015). The normalized shear modulus is $G/(\rho^f U_b^2) = 10$.

A nonuniform grid is used in x and z and a uniform grid in y . The uniform section spans $6 < x/s < 14$ and $5\pi - 2 < z/s < 5\pi + 2$, with mild stretching ($\approx 0.4\%$) elsewhere. The minimum grid spacing is $\Delta x_{\min}/s = \Delta y/s = \Delta z_{\min}/s = 0.03125$, yielding a grid size of $678 \times 640 \times 628$.

The inflow boundary condition at $x = 0$ is obtained from a precursor periodic turbulent channel simulation at $Re_{ch} = 2800$ ($Re_\tau = u_\tau H/\nu = 180$, where $u_\tau = \sqrt{\tau_w/\rho}$ is the friction velocity calculated using the mean wall shear stress τ_w), with the same domain size in y and z but coarser resolution (286×164 in y - z) and temporal step $\Delta t U_b/H = 4 \times 10^{-3}$. For the present simulations, bilinear spatial interpolation and linear temporal interpolation are applied to transfer the inflow fields.

At the outlet, a convective boundary condition matching the channel mass flux is imposed. No-slip conditions are enforced at the top and bottom walls, and periodicity is applied in z . In the non-FSI cases, the solid region is masked with no-slip boundaries. In the FSI cases, Neumann boundary conditions are applied for ϕ , and extrapolation is used for B at the inlet, outlet, and walls.

A uniform time step of $\Delta t U_b/H = 10^{-4}$ is used. The momentum equations are advanced using the second-order Crank–Nicolson scheme. The advection of ϕ is handled by a geometric volume-of-fluid method (Fakhreddine et al., 2024), while B is advanced using a fifth-order weighted essentially non-oscillatory (WENO-5) scheme (Dominguez et al., 2024).

10.3 Results

Element deformation and stress

Figure 10.2 shows the temporal evolution of the solid elements shortly after transitioning from the non-FSI to the FSI stage. The interface is visualized using the iso-surface $\phi = 0.5$.

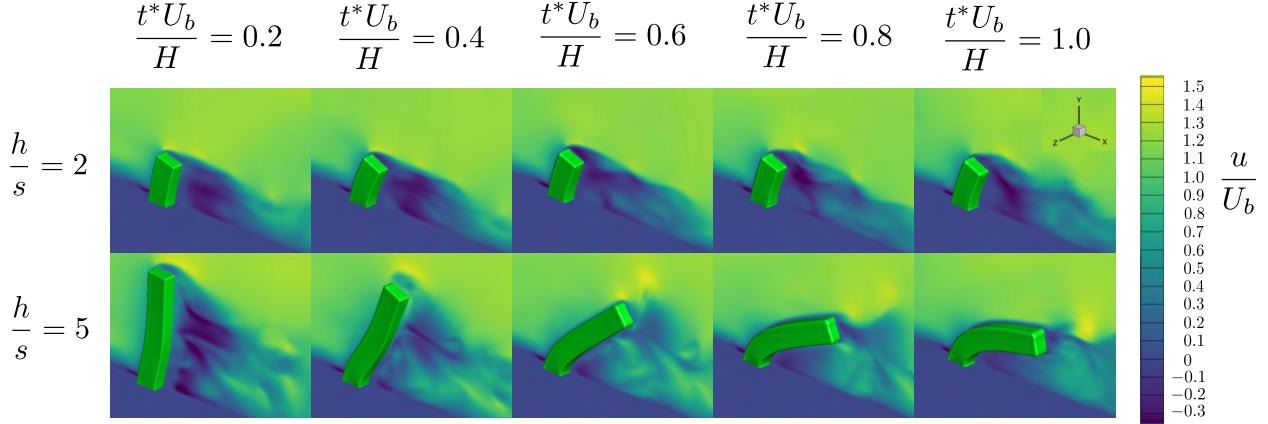


Figure 10.2: Element orientation following the switch from non-FSI to FSI simulation.

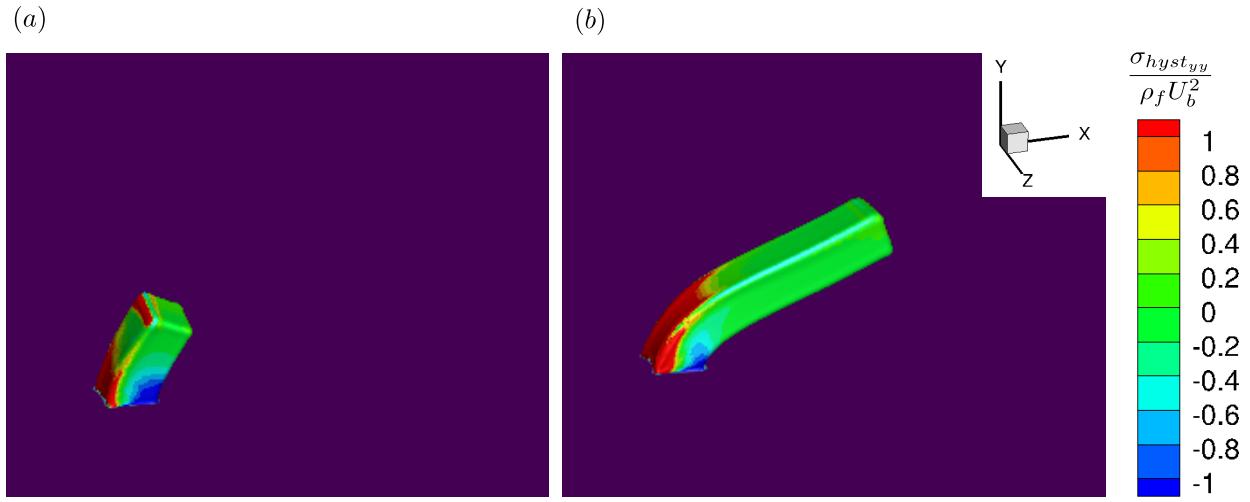


Figure 10.3: Instantaneous contours of the yy -component of the hyperelastic stress on the solid surface for (a) $h/s = 2$ and (b) $h/s = 5$.

The $h/s = 2$ element undergoes noticeable bending, reaching its mean deflected orientation within approximately 0.4 time units after enabling FSI. The $h/s = 5$ element shows the largest deviation from vertical, owing to the larger moment induced at its root. This case requires roughly one time unit to settle into a mean equilibrium shape. The resulting wake differs markedly from the $h/s = 2$ case, and instead resembles that of a rigid element tilted to a similar inclination. During the initial transient of the $h/s = 5$ case, an inflection point forms in the element shape at $t^*U_b/H = 0.2$ and 0.4. This arises because the near-wall portion of the element experiences higher shear stress than the upper tip. By $t^*U_b/H = 0.6$, the upper portion deforms sufficiently to eliminate the inflection, producing the smoothly curved equilibrium profile observed at $t^*U_b/H = 1$.

Figure 10.3 shows instantaneous distributions of the hyperelastic stress component σ_{yy}^s on the solid surface. Bending moments about the base generate tensile stresses near the windward face and compressive stresses near the leeward face, particularly close to the wall. The higher flexibility of the $h/s = 5$ element results in larger tensile stresses and reduced compressive stresses compared with the stiffer $h/s = 2$ element.

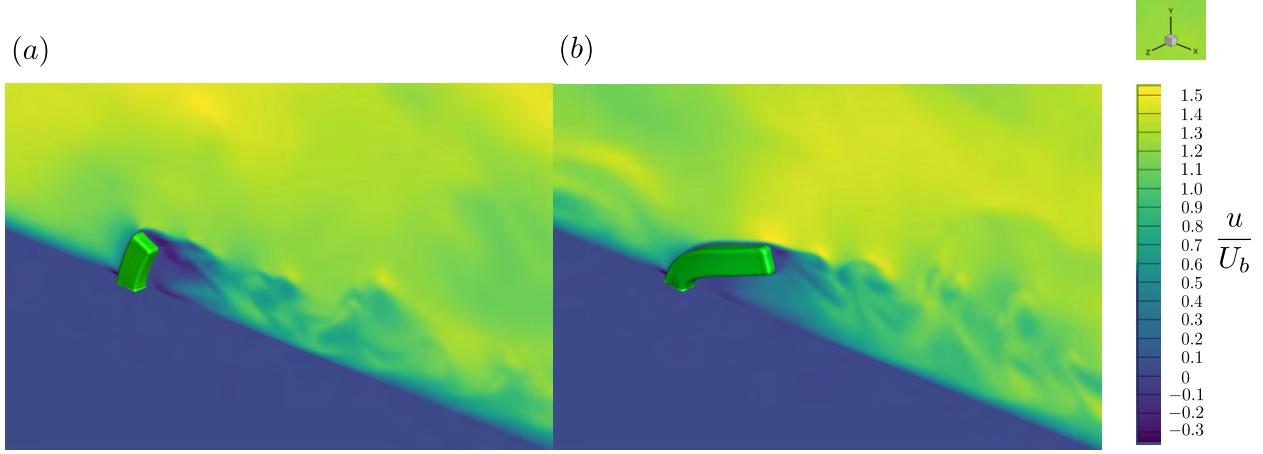


Figure 10.4: Instantaneous contours of streamwise velocity in the downstream wake of the compliant elements for (a) $h/s = 2$ and (b) $h/s = 5$.

Element oscillation and drag

Figure 10.4 shows the instantaneous streamwise velocity contours in the downstream wake of the compliant elements. Both cases exhibit coherent vortex shedding, which is especially pronounced for the $h/s = 5$ element. The vortices are spaced at a length scale comparable to the element height. Because the $h/s = 2$ element undergoes only mild bending, its wake retains a bluff-body character, whereas the strongly deflected $h/s = 5$ element produces a more streamlined wake.

The oscillatory dynamics of the elements are analyzed by tracking the spanwise coordinate of the centroid of the top surface. The element is not constrained in the spanwise direction, which allows for free oscillations in that direction. The displacement time series is analyzed using a discrete Fourier transform (after discarding the first time unit to remove transients). Figure 10.5 shows both the time histories and their corresponding spectra. For $h/s = 2$, the dominant non-dimensional frequency is $fH/U_b = 1.11$, corresponding to $fh/U_b = 0.222$ when normalized by the element height. For $h/s = 5$, the dominant frequency is $fH/U_b = 1.27$, corresponding to $fh/U_b = 0.635$. Table 10.1 compares these frequencies with theoretical values for free vibration of a cantilever beam. It can be observed that the $h/s = 2$ element primarily oscillates in its first mode, while the $h/s = 5$ element exhibits second-mode vibration. The corresponding Strouhal numbers are comparable between the two cases.

| Case | $\frac{h}{s} = 2$ | $\frac{h}{s} = 5$ |
|--|--------------------------|-------------------------|
| Simulation FFT (Dominant fH/U_b) | 1.11 | 1.27 |
| Theoretical ($1^{st}, 2^{nd}, 3^{rd}$ mode) | 1.28, 7.99, 20.69 | 0.51, 1.28, 8.28 |

Table 10.1: Comparison of Strouhal numbers from simulation and linear beam theory.

The coefficients of pressure and viscous drag on each element are defined as

$$C_{p_x} = \frac{2F_{p_x}}{\rho_f U_b^2 h s}, \quad C_{v_x} = \frac{2F_{v_x}}{\rho_f U_b^2 h s}, \quad (10.8)$$

where,

$$F_{p_x} = \hat{\mathbf{i}} \cdot \left(- \oint_{\phi=0.5} p \hat{\mathbf{n}} dA \right), \quad F_{v_x} = \hat{\mathbf{i}} \cdot \left(\oint_{\phi=0.5} \boldsymbol{\sigma}^f \cdot \hat{\mathbf{n}} dA \right) \quad (10.9)$$

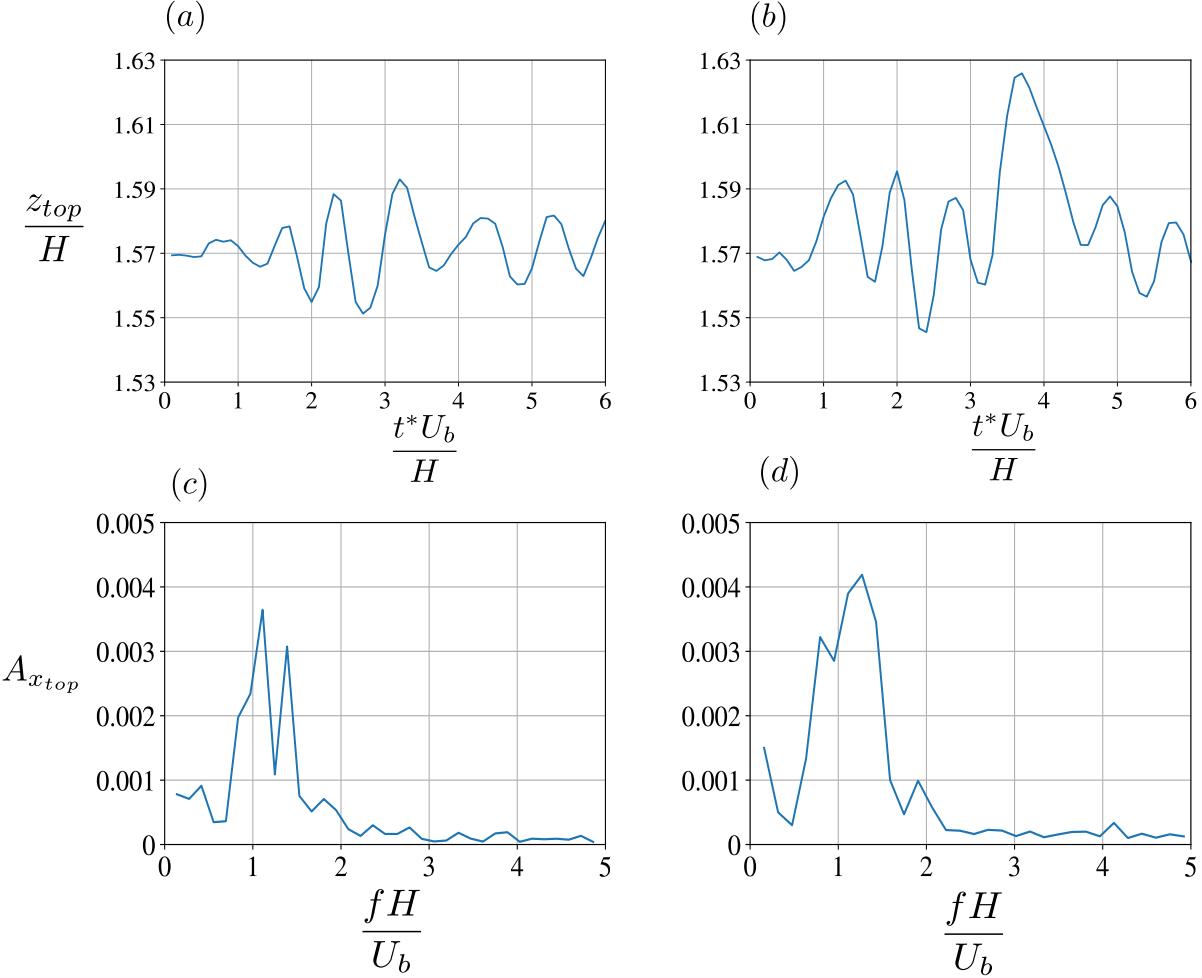


Figure 10.5: Time history of top-surface spanwise coordinate and its Fourier spectrum, respectively, for (a, c) $h/s = 2$ and (b, d) $h/s = 5$.

represent the streamwise components of the pressure and viscous forces, respectively, integrated over the interface $\phi = 0.5$. Here, $\hat{\mathbf{i}}$ and $\hat{\mathbf{n}}$ denote the streamwise unit vector and local surface normal. Figure 10.6 presents the time series of C_{p_x} and C_{v_x} for both cases. Owing to its bluff wake, the $h/s = 2$ element experiences a pressure drag roughly an order of magnitude higher than that of the $h/s = 5$ case. Conversely, the viscous drag is slightly larger for $h/s = 5$. Overall, the total drag is dominated by the pressure contribution, which exceeds the viscous contribution by approximately two orders of magnitude.

Mean velocity and mean-square intensities

Figure 10.7 shows the mean streamwise velocity contours on the $z/s = 5\pi$ ($z/H = \pi/2$) plane. In the near-body region, the bluff nature of the $h/s = 2$ case contrasts with the more streamlined wake observed for $h/s = 5$. The corresponding mean streamwise velocity profiles on the vertical center plane ($z/H = \pi/2$) at three streamwise locations are shown in Figure 10.8. Interestingly, the differences between the two cases do not persist beyond $y/H \approx 0.4$, which lies just above the end of the logarithmic layer in the reference channel flow. This could be attributed to the inclination of the $h/s = 5$ case, which effectively limits its vertical extent to $y/H \approx 0.3$ and therefore makes its

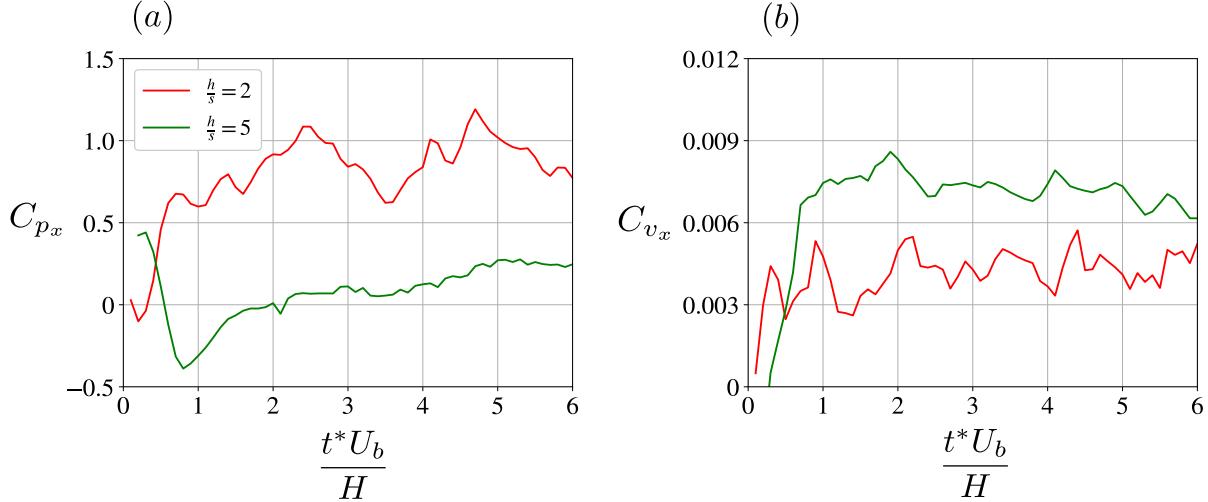


Figure 10.6: (a) Pressure and (b) viscous drag coefficients.

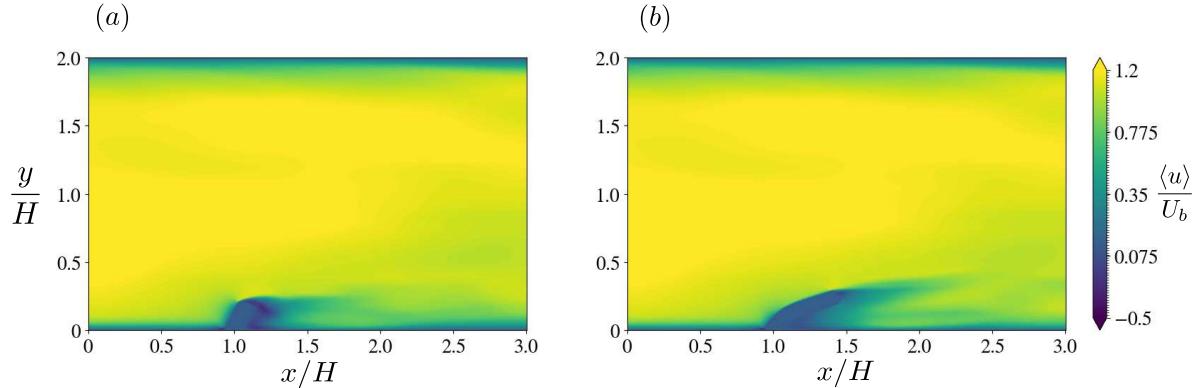


Figure 10.7: Mean streamwise velocity contours on the $z/s = 5\pi$ ($z/H = \pi/2$) plane for (a) $h/s = 2$ and (b) $h/s = 5$.

impact more comparable to that of $h/s = 2$.

Figure 10.9 shows the mean planar velocity field ($\langle u \rangle, \langle v \rangle$) on the plane $z/H = \pi/2$, with streamlines overlaid on contours of the mean velocity magnitude $\sqrt{\langle u \rangle^2 + \langle v \rangle^2}$. Both cases display an upstream recirculating vortex. However, the downstream stagnation point differs in location: for the $h/s = 2$ element, it lies farther from the wall and farther downstream. This happens as a result of a larger recirculating region for the $h/s = 2$ case, thereby pushing the stagnation point further downstream. In contrast, the $h/s = 5$ element generates a more streamlined flow pattern. It is interesting to note that the streamlines for the element with $h/s = 5$ resemble the streamline pattern observed by Muppudi and Mahesh (2005) for a jet in crossflow (see Figure 5 in their paper).

The mean-squared velocity fluctuations ($\langle u'^2 \rangle, \langle v'^2 \rangle$, and $\langle w'^2 \rangle$) on the plane $z/H = \pi/2$ are shown in Figure 10.10. The streamwise component ($\langle u'^2 \rangle$) is dominant in both cases, with the $h/s = 5$ element exhibiting higher overall intensity than the $h/s = 2$ element, especially in the windward boundary layer (Figures 10.10a,b). The bluff geometry of the $h/s = 2$ element, however, deflects the flow laterally around its sides rather than over the top, as occurs for the more inclined $h/s = 5$ case. This lateral deflection, combined with the incoming turbulent fluctuations, leads to enhanced spanwise velocity fluctuations ($\langle w'^2 \rangle$) in the $h/s = 2$ case relative to $h/s = 5$ (Figures 10.10e,f).

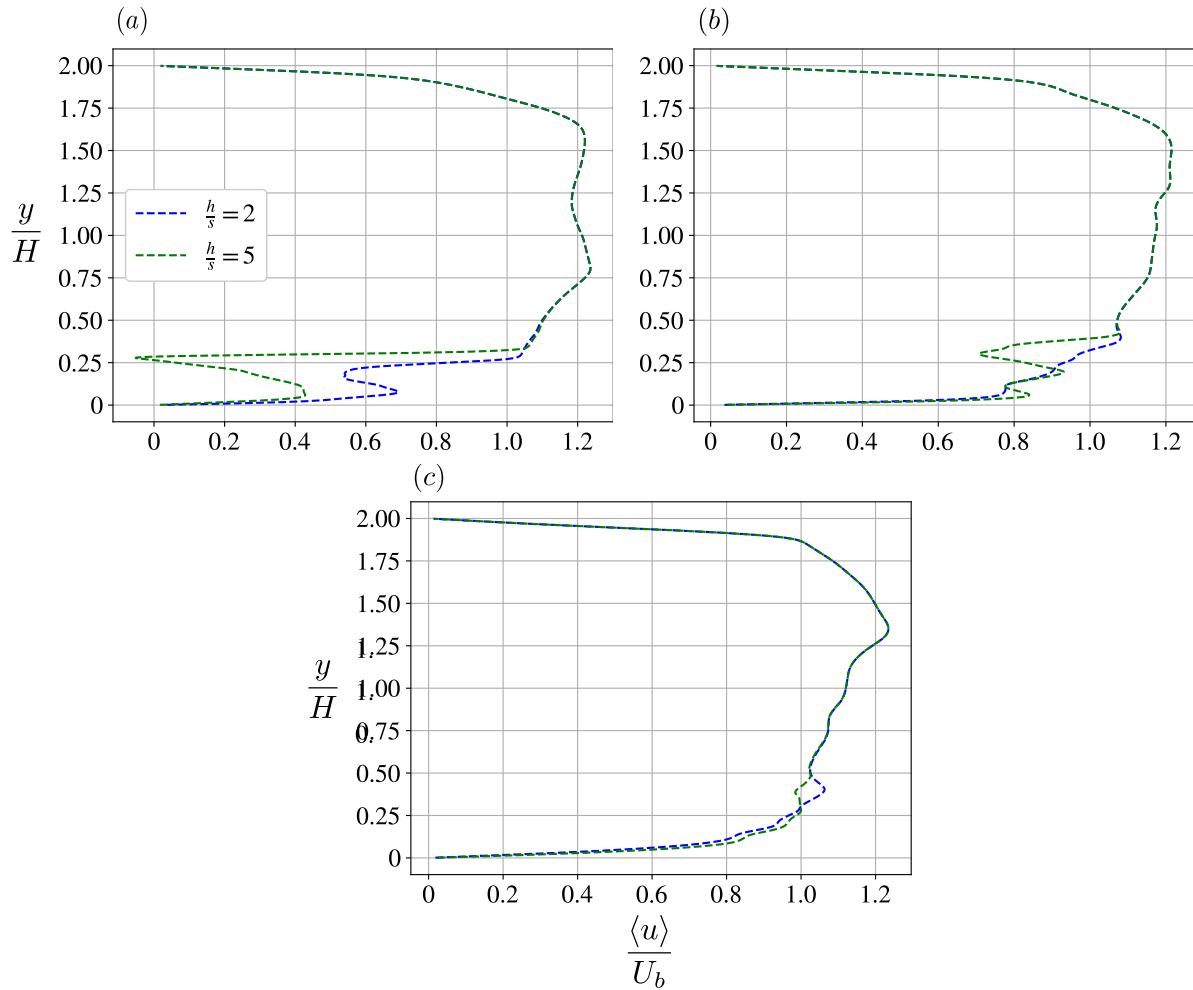


Figure 10.8: Mean streamwise velocity profiles on the vertical center plane ($z/H = \pi/2$) at (a) $x/H = 1.5$, (b) $x/H = 2.0$, and (c) $x/H = 2.5$.

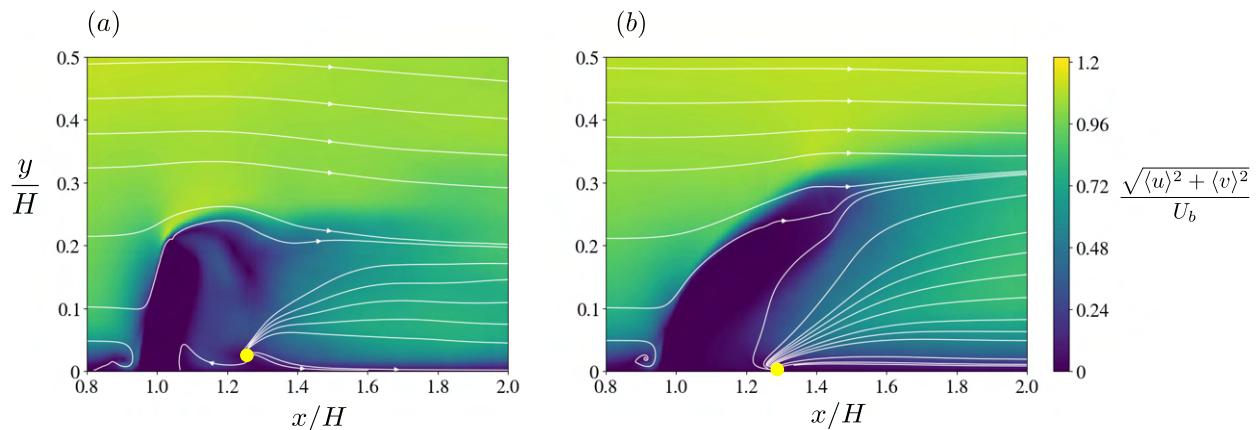


Figure 10.9: Mean streamlines plotted on contours of mean planar velocity magnitude on the vertical center plane ($z/H = \pi/2$) for (a) $h/s = 2$ and (b) $h/s = 5$. Yellow circle represents stagnation point downstream.

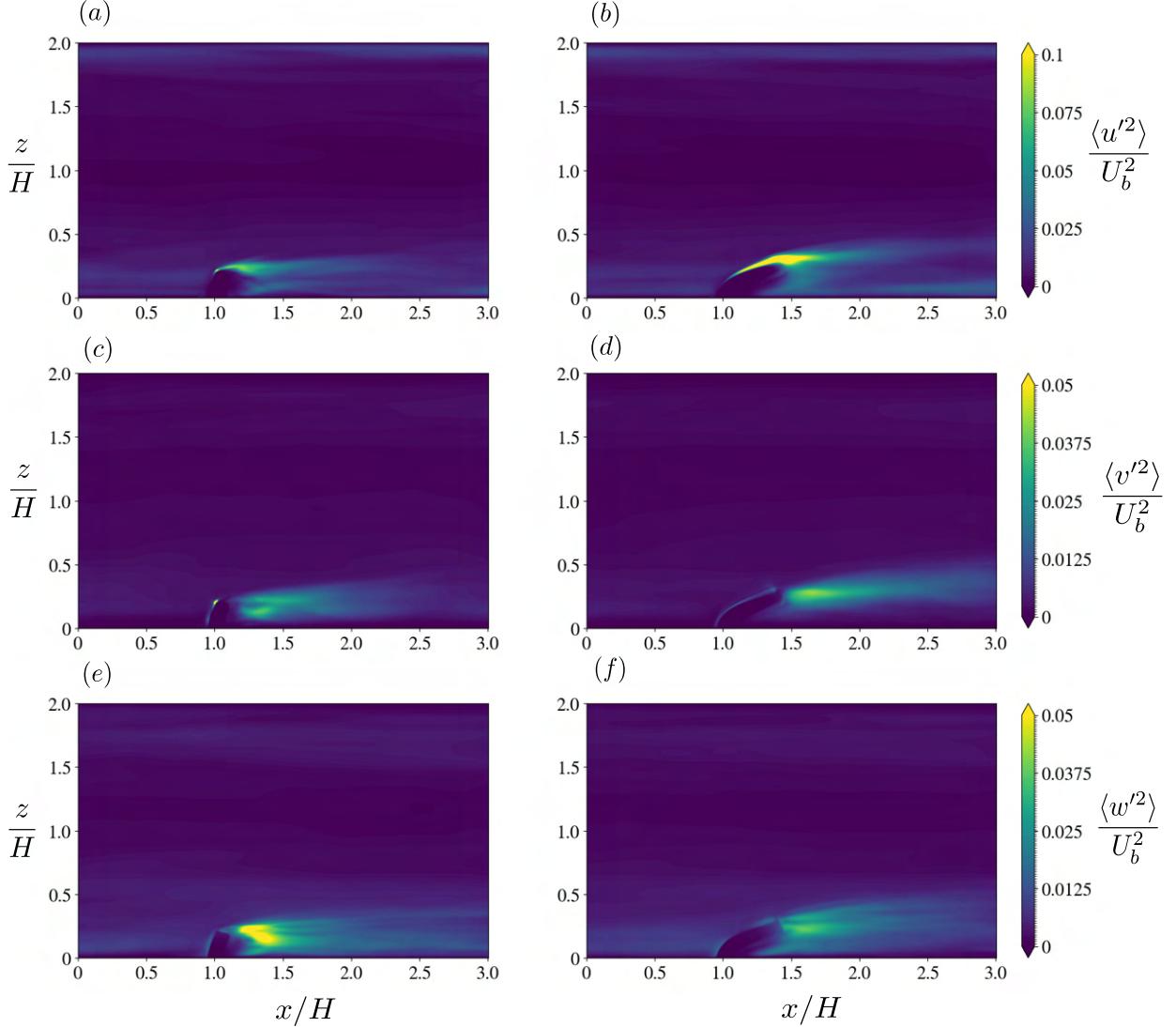


Figure 10.10: Contours for mean-square intensities on the vertical center plane ($z/H = \pi/2$) for (a, c, e) $h/s = 2$ and (b,d,f) $h/s = 5$.

10.4 Discussion

The present simulations highlight the effect of element compliance on the coupled flow-structure interaction in a turbulent channel. The deformation dynamics reveal that element with ($h/s = 2$) undergoes modest bending, leading to an asymmetric but still bluff wake, while the element with $h/s = 5$ exhibits pronounced curvature and large deviations from vertical. The emergence and disappearance of inflection points during the initial transient highlight the competition between near-wall shear stresses and forces in outer layers of the flow in establishing the equilibrium shape.

The stress contours demonstrate that longer elements not only deform higher but also experience different hyperelastic stresses near the base. This observation is consistent with the scaling of bending moments with element height. The pressure drag contributes approximately two orders of magnitude more to the total drag compared to the skin-friction drag. Furthermore, the pressure drag is higher for the shorter element, whereas the skin-friction drag is higher for the longer element. Moreover, the oscillation frequencies of the element vibration in the unconstrained spanwise

References

direction correspond to different modes of free vibration for the two elements.

The mean streamwise velocity contours and profiles indicate that, while the wakes differ significantly in the near-body region, their influence does not extend far beyond the logarithmic region of the boundary layer. This suggests that the element geometry primarily alters the near-field dynamics, with limited impact on the outer layer. The inclination of the $h/s = 5$ element appears to reduce its effective vertical extent, making its influence on the outer flow comparable to that of the stiffer $h/s = 2$ element.

The mean streamline patterns further highlight these differences. Both cases exhibit a similar upstream vortex and outer flow streamlines; however, the downstream stagnation point for the $h/s = 2$ element is located slightly farther downstream and away from the wall. Interestingly, the streamline topology for the $h/s = 5$ case resembles that of a jet in crossflow. The higher streamwise fluctuations observed for the $h/s = 5$ element reflect stronger interaction with the boundary layer on the element, while the enhanced spanwise fluctuations for the $h/s = 2$ case arise from lateral flow diversion around its bluff geometry. These differences suggest that element shape and flexibility not only modify wake structure but can also lead to different distributions of turbulent intensities among velocity components in the near-wall region.

While the present study focuses on deformation, wake dynamics and oscillatory behavior of a single compliant element, it is also a foundation for modeling soft biofouled surfaces in the environment. A natural extension of this work is to introduce multiple deformable elements arranged in arrays or staggered configurations. Such setups would allow us to examine the effect of distribution for the same surface coverage, the ultimate goal being simulating turbulent flow interaction with realistic biofouled surface configurations, obtained from field testing or grown in a lab environment.

References

P. A. Araújo, J. Malheiro, I. Machado, F. Mergulhao, L. Melo, and M. Simoes. Influence of flow velocity on the characteristics of *pseudomonas fluorescens* biofilms. *Journal of Environmental Engineering*, 142(7):04016031, 2016.

H. Dominguez, N. Riel, and P. Lanari. Modelling chemical advection during magma ascent. *Geoscientific Model Development*, 17(16):6105–6122, 2024.

A. Esteghamatian, J. Katz, and T. A. Zaki. Spatiotemporal characterization of turbulent channel flow with a hyperelastic compliant wall. *Journal of Fluid Mechanics*, 942:A35, 2022.

J. B. Estrada, C. Barajas, D. L. Henann, E. Johnsen, and C. Franck. High strain-rate soft material characterization via inertial cavitation. *Journal of the Mechanics and Physics of Solids*, 112:291–317, 2018.

A. Fakhreddine, K. Alamé, and K. Mahesh. Directionally-split volume-of-fluid technique for front propagation under curvature flow. *International Journal for Numerical Methods in Fluids*, 96(9):1517–1554, 2024.

S. García, A. Trueba, D. Boullosa-Falces, H. Islam, and C. G. Soares. Predicting ship frictional resistance due to biofouling using reynolds-averaged navier-stokes simulations. *Applied Ocean Research*, 101:102203, 2020.

I. Hameduddin, C. Meneveau, T. A. Zaki, and D. F. Gayme. Geometric decomposition of the conformation tensor in viscoelastic turbulence. *Journal of Fluid Mechanics*, 842:395–427, 2018.

R. M. Jensen and P. Forooghi. A study of turbulent flow over patchy roughness. *Journal of Fluid Mechanics*, 1006: A28, 2025.

G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted ENO schemes. *Journal of Computational Physics*, 126(1):202–228, 1996.

I. K. Kaminaris and E. Balaras. Direct numerical simulations of turbulent boundary layers developing over synthesised calcareous marine biofouling surfaces. *Ships and Offshore Structures*, pages 1–12, 2024.

References

X.-D. Liu, S. Osher, and T. Chan. Weighted essentially non-oscillatory schemes. *Journal of Computational Physics*, 115(1):200–212, 1994.

S. Muppudi and K. Mahesh. Study of trajectories of jets in crossflow using direct numerical simulations. *Journal of Fluid Mechanics*, 530:81–100, 2005.

H. Park and H. Nepf. Turbulence in a channel with a patchy submerged canopy: the impact of spatial configuration. *Journal of Fluid Mechanics*, 1006:A15, 2025.

B. W. Peterson, Y. He, Y. Ren, A. Zerdoum, M. R. Libera, P. K. Sharma, A.-J. Van Winkelhoff, D. Neut, P. Stoodley, H. C. Van Der Mei, et al. Viscoelasticity of biofilms and their recalcitrance to mechanical and chemical challenges. *FEMS Microbiology Reviews*, 39(2):234–245, 2015.

S. Prajapati, A. Fakhreddine, and K. Mahesh. Monolithic framework to simulate fluid-structure interaction problems using geometric volume-of-fluid method. *arXiv preprint arXiv:2505.22920*, 2025.

C. J. Rupp, C. A. Fux, and P. Stoodley. Viscoelasticity of *staphylococcus aureus* biofilms in response to fluid shear allows resistance to detachment and facilitates rolling migration. *Applied and Environmental Microbiology*, 71(4): 2175–2178, 2005.

P. Stoodley, Z. Lewandowski, J. D. Boyle, and H. M. Lappin-Scott. Structural deformation of bacterial biofilms caused by short-term fluctuations in fluid shear: An in situ investigation of biofilm rheology. *Biotechnology and bioengineering*, 65(1):83–92, 1999.

K. Sugiyama, S. Ii, S. Takeuchi, S. Takagi, and Y. Matsumoto. A full Eulerian finite difference approach for solving fluid–structure coupling problems. *Journal of Computational Physics*, 230(3):596–627, 2011.

E. Tsagkari and W. T. Sloan. Turbulence accelerates the growth of drinking water biofilms. *Bioprocess and Biosystems Engineering*, 41(6):757–770, 2018.

J. Wang, S. S. Koley, and J. Katz. On the interaction of a compliant wall with a turbulent boundary layer. *Journal of Fluid Mechanics*, 899:A20, 2020.

J. N. Wilking, T. E. Angelini, A. Seminara, M. P. Brenner, and D. A. Weitz. Biofilms as complex fluids. *MRS bulletin*, 36(5):385–391, 2011.

T. C. Zhang and P. L. Bishop. Density, porosity, and pore structure of biofilms. *Water Research*, 28(11):2267–2277, 1994.

Statistical Modeling of Electromagnetic Waves Scattering from Random Ocean Surfaces

BEHZAD YEKTAKHAH, RYAN STROHMAN, & KAMAL SARABANDI

Accurate modeling of electromagnetic (EM) wave propagation and scattering in complex random environments is crucial for evaluating the performance of radar and communication systems. In particular, modeling the scattering of EM waves from ocean surfaces plays a vital role in both scientific research and practical applications, including microwave remote sensing, weather forecasting, and the assessment of radar and communication systems' performance in oceanic environments. A robust statistical model that accurately captures the characteristics of EM waves scattered from ocean surfaces can significantly enhance the efficiency of these assessments. Such a model enables the rapid generation of scattered wave data across various conditions and scenarios, reducing the reliance on costly and time-consuming field measurements. This capability is invaluable for advancing both commercial and defense technologies while supporting the broader understanding of wave interactions with complex surfaces. In this project, an accurate and high-fidelity statistical model is presented to describe the random phase and amplitude of waves scattered from ocean surfaces. The model is developed under different conditions, including wind speed, wind direction, incident wave direction, wave frequency, and wave polarization. A method for randomly generating scattered fields from the developed statistical models is also presented.

Keywords: electromagnetic waves, ocean surface, scattering, statistical

11.1 Introduction

As the initial step in developing a statistical model for scattering from ocean surfaces, it is essential to acquire scattered field data under various conditions. These conditions include different wind speeds, angular differences between the incident wave direction and wind direction, frequency, and incident wave polarization. These data can be obtained through either measurements or simulations, with simulations serving as a versatile and efficient tool for this purpose, as utilized in this work. To generate reliable data from simulations, accurate modeling of the random surface of the ocean is critical. The surface model must precisely represent the ocean profile at a resolution comparable to a fraction of the electromagnetic wavelength. This ensures fidelity in capturing the physical characteristics necessary for accurate scattering predictions under diverse environmental scenarios. This work utilizes accurate models to represent gravity and capillary waves, two components of ocean waves, to generate random ocean surfaces at different wind speeds. The models for generating random ocean surfaces are presented in Section 11.2.

Due to the large wavelength of the gravity waves (hundreds of meters), numerical simulations of scattering from the ocean surface involve electrically large problems (typically thousands of wavelengths). The presence of capillary waves with wavelengths smaller than or comparable to those of typical frequencies used in naval scenarios (C-band to Ku-band) prevents the use of fast and efficient methods like physical optics (PO). In this project, the ocean surface is modeled as an impedance boundary condition, which significantly simplifies the numerical calculations. The surface integral

equation representing the scattering from the surface is solved by the multilevel fast multipole method (MLFMM). Details of the electromagnetic simulations are provided in Section 11.3.

Statistical modeling aims to provide a model to generate tangential components of the scattered field on a near-field plane above the ocean surface for a given ocean condition and incident wave attributes, such as incident angle, polarization, and frequency. Knowing the tangential fields on a plane, the scattered waves can be calculated at any point above the plane. Assuming the incident plane wave is decomposed into horizontal (h) and vertical (v) polarizations, it can be shown that the real and imaginary parts of co-polarized components of tangential scattered fields due to incident h and v polarizations are Gaussian random variables with a joint probability density function described by the covariance of real and imaginary parts of co-polarized components. The same is also true for the real and imaginary parts of cross-polarized components. To develop the model, parameters describing the joint PDFs and also the spectral density of one of the co- and cross-polarized components must be derived for different ocean conditions and incident wave attributes. To use the model, random samples of one component of co- and cross-polarization are generated by using its power spectral density, and then the other components are generated knowing the samples of the first component at each location and their joint PDFs. Details of the statistical models are provided in Section 11.4.

11.2 Ocean Surface Model

When the wind blows over the ocean surface, turbulent airflow first creates capillary waves with lengths of a few centimeters and heights of a few millimeters. Capillary waves wrinkle the ocean surface, and the water's surface tension resists these wrinkles. As the wind continues to blow, the capillary waves nonlinearly interact and continue to grow both in length and height, eventually evolving into gravity waves with larger lengths and heights. Gravity waves get energy from the wind by means of direct push by wind on their upwind faces at wave speeds lower than the wind speed or through frictional drag of the air on the ocean surface at any wave speed. Without energy loss and in the presence of continuous wind energy transfer, the gravity waves continue to grow infinitely. However, the energy is transferred through viscosity to shorter wavelengths or lost through wave breaking. As the waves grow in height, they also grow in length at a faster rate. When the height reaches a limit and stops growing, the wavelength may continue to grow (Kinsman, 1984; Long and Ulaby, 2015). Capillary waves are continuously generated and turned into gravity waves. As a result, the ocean surface can be considered as a superposition of short-wavelength capillary waves and large wavelength gravity waves, as illustrated in Figure 11.1. Large waves can propagate over long distances and thus may not be directly related to local wind conditions. On the other hand, capillary waves tend to be in equilibrium with the local wind. Unlike gravity waves, capillary waves quickly disappear as the wind stops.

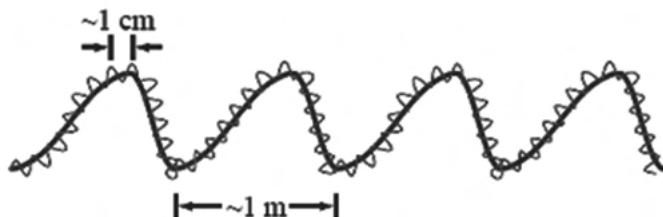


Figure 11.1: Capillary waves with short wavelength and gravity waves with longer wavelength (Long and Ulaby, 2015).

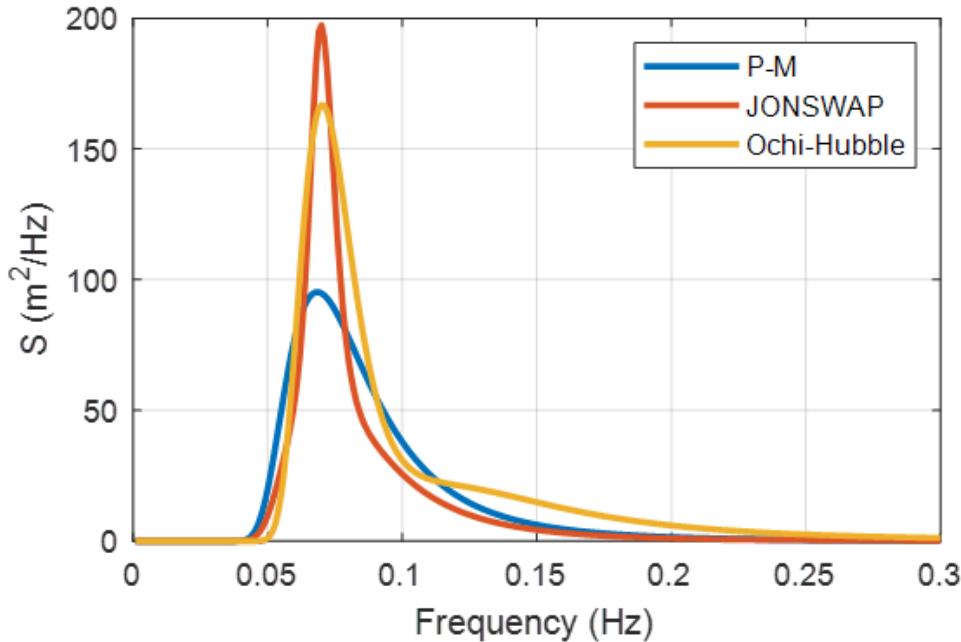


Figure 11.2: Comparison of different wave spectra at $U_{19.5} = 20$ m/s. Fetch for JONWSAP is set to 270 km, and the significant wave height for the Ochi-Hubble spectrum is set to 10 m.

Gravity waves model

Looking at the random gravity waves, it seems that they are made up of countless random waves with different wavelengths and amplitudes. To describe the surface, wave spectra are used. Wave spectrum represents the distribution of wave energy among diverse wave frequencies (or wavelengths). The Pierson-Moskowitz (P-M) (Carter, 1982) and Joint North Sea Wave Project (JONSWAP) (Hasselmann et al., 1973) are two well-known ocean wave spectra. P-M spectrum is only a function of windspeed, while JONSWAP is also a function of “fetch”. Fetch is the distance over which the wind blows in a single direction across the ocean and generates waves. Figure 11.2 shows an example of wave spectra for wind speed of 20 m/s at the height of 19.5 m above the mean ocean surface ($U_{19.5}$) and fetch of 270 km. Both P-M and JONSWAP spectra show a single peak. However, in the measured ocean wave spectra, there are high-frequency components in the spectrum that must be modeled for a more accurate representation of the ocean waves that matters to the radar response. These high-frequency components with smaller wavelengths are also crucial in accurate statistical modeling of scattering from the ocean surface, as they are comparable to the wavelength of the electromagnetic waves. The Ochi-Hubble spectrum is a six-parameter model that provides a double-peaked spectrum with better modeling of the spectrum at higher frequencies. This model includes both remote and local wind-generated waves (Ochi and Hubble, 1976). The six parameters in the Ochi-Hubble spectrum are a function of “significant wave height”. The significant wave height (in m), defined as the average height of the highest one-third of waves, is a function of wind speed, fetch, and wind duration. Figure 11.2 shows the Ochi-Hubble spectrum for the same wind speed and fetch. As it can be observed, it provides a more accurate spectrum for high-frequency waves. In this project, the Ochi-Hubble spectrum is used. Considering N samples of the wave spectrum $S(f)$, the ocean surface height (h) can be randomly generated over time (t) and position (x, y) by superposition of N waves with frequency f_n traveling in discrete directions (θ_n) using the following formula (Jefferys, 1987; Lin et al., 2022)

$$h = \sum_{n=1}^N a_n \cos(2\pi f_n t - k_n(x \cos(\theta_n) + y \sin(\theta_n)) + \phi_n) \quad (11.1)$$

where a_n is the amplitude of the n th wave and is equal to $\sqrt{2S(f_n)\Delta f}$ (S is the wave spectrum and Δf is the frequency difference between the samples). Based on the discrete Fourier transform properties, Δf is inversely proportional to the maximum time/distance where the profile of the wave can be generated (beyond that the profile will be repeated). ϕ_n is randomly generated over 0 – 2π with uniform distribution. k_n is the n th wavenumber defined by

$$\omega_n^2 = k_n g \tanh(k_n d) \quad (11.2)$$

where d is the water depth, $\omega_n = 2\pi f_n$, and g is Earth's gravitational acceleration. For deep water ($k_n d \gg 1$), the formula is simplified to $k_n = (\omega_n^2)/g$. In (11.1), θ_n is randomly generated following the wave direction spectrum (WDS) (Mitsuyasu et al., 1980). A simple model for WDS is

$$f(\theta) = \frac{2}{\pi} \cos^2(\theta) \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}. \quad (11.3)$$

To generate random numbers with a specific spectrum, the method "acceptance and rejection" can be used (Flury, 1990). Figure 11.3 shows samples of a random ocean surface profile (at $t = 0$) generated using the spectra shown in Figure 11.2. As expected, the Ochi-Hubble model exhibits more local variations.

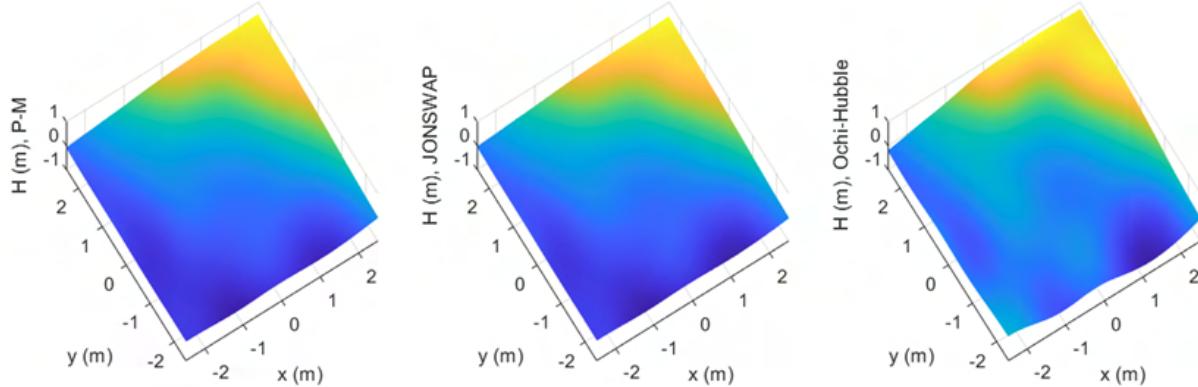


Figure 11.3: Samples of random ocean surface profile generated using the spectra shown in Figure 11.2. *Left:* P-M spectrum. *Center:* JONSWAP spectrum. *Right:* Ochi-Hubble spectrum.

Capillary waves model

To model capillary waves (with wavelength in the centimeter range), the surface autocorrelation function (R) or its Fourier transform, surface spectrum, is used. An approximate representation of the surface autocorrelation function for the capillary waves is given by (Long and Ulaby, 2015)

$$R(r, \phi) = s^2 \exp\left(-\frac{r}{L_t}\right), \quad (11.4)$$

where, s^2 is the surface height variance and is an increasing function of the wind speed and L_t is

$$L_t = L_u \cos^2(\phi) + L_c \sin^2(\phi). \quad (11.5)$$

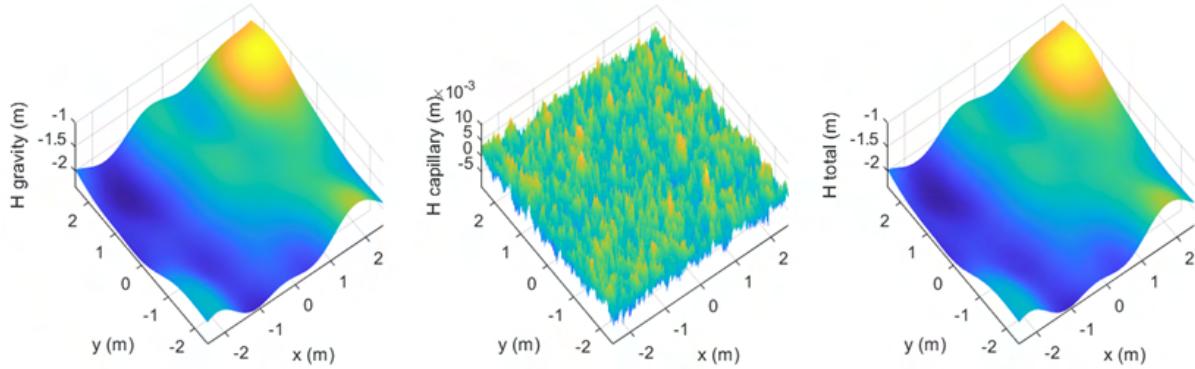


Figure 11.4: Sample of random ocean surface profile at wind speed of 20 m/s (scale of height is different from x and y axes). The spectrum for gravity waves is shown in Figure 11.2 by Ochi-Hubble spectrum. *Left:* Gravity wave. *Center:* Capillary wave. *Right:* Superposition of gravity and capillary waves.

In (11.5), L_u is the correlation length in the upwind direction (along $\phi = 0^\circ$) and L_c is the correlation length in the crosswind direction ($\phi = 90^\circ$). In the oceanographic convention, the wind direction $\phi = 0^\circ$ corresponds to a flow from the south to the north. As the wind direction rotates clockwise as seen above, ϕ increases. In electromagnetic modeling, the difference between the wind direction and the azimuthal angle of the incident electromagnetic wave are important and so the wind direction can always be considered equal to 0° . L_u and L_c are decreasing functions of the wind speed and $L_c > L_u$. A model relating s , L_u , and L_c to the wind speed (U) is (Long and Ulaby, 2015)

$$s = -0.00768 + 0.0237U - 0.00587U^2, \quad (11.6a)$$

$$L_u = 24.25 - 1.56U + 0.0302U^2, \quad (11.6b)$$

$$L_c = 33.89 - 0.225U + 0.0453U^2. \quad (11.6c)$$

To generate two-dimensional (2D) samples of the capillary waves from the correlation function, the following procedure can be followed (Novarini and Caruthers, 1973; Oh, 1993): (1) Generate a set of random, uncorrelated numbers (X) with a Gaussian distribution of zero mean and unit standard deviation. (2) Find the convolution of (filtering process) X with W where W is the square root of the surface spectrum (the 2D Fourier transform of the surface autocorrelation). Figure 11.4 shows a sample of the generated gravity and capillary waves and their superposition (H_{total}). For this surface, the Ochi-Hubble spectrum shown in Figure 11.2 is used. Capillary waves are generated at a wind speed of 20 m/s. In Figure 11.4 (center) the rms height of capillary waves is 0.23 cm. Figure 11.5 shows the autocorrelation of the capillary wave in Figure 11.4 (center) along x and y directions. As observed, the generated wave autocorrelation follows the model.

11.3 Electromagnetic Simulation of Large Ocean Surfaces

Complex dielectric constant of saline water

The most accurate model available for the dielectric constant of water is based on a double-Debye dielectric model (D3M) developed by William Ellison. The model was developed for seawater, and it reduces to a model for pure water when the water salinity is set to zero Long and Ulaby (2015).

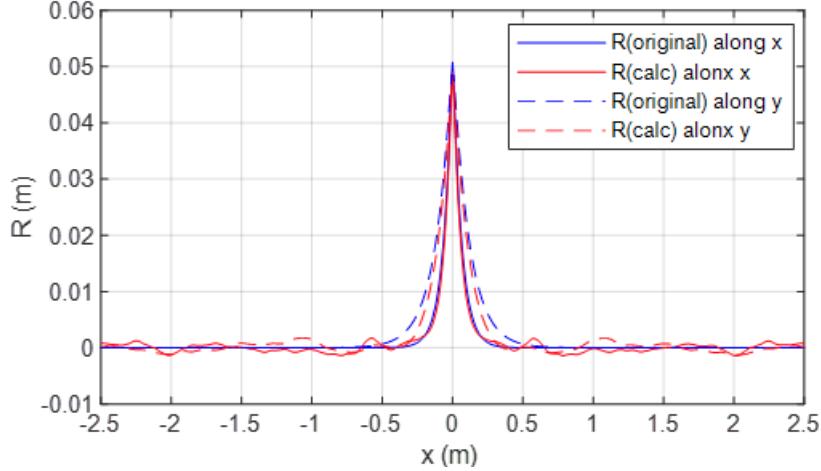


Figure 11.5: Autocorrelation of the generated surface in Figure 11.4 (center). The curves in blue show the original model used to create the surface.

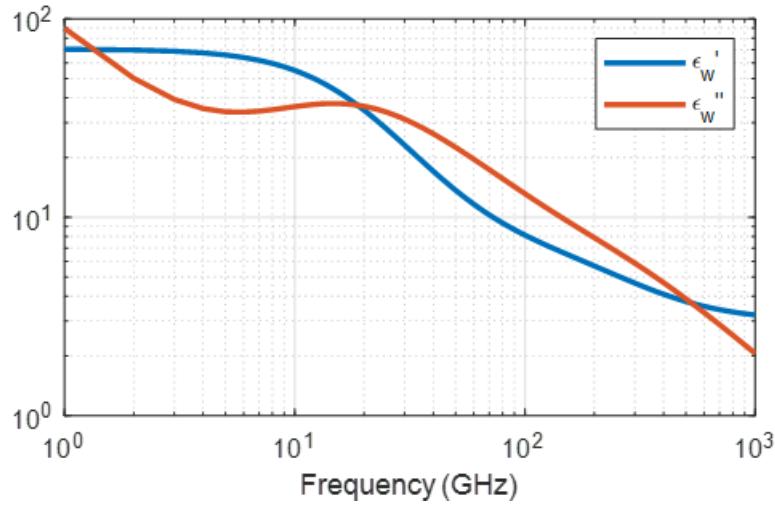


Figure 11.6: Complex permittivity of ocean water at 20° and for $S = 35\%$.

D3M is valid for frequencies up to 1 THz, the temperature range of 0–30°C, and salinity of up to 40% (parts per thousand). The average salinity of ocean water is 35%. Figure 11.6 shows the complex permittivity of ocean water versus frequency at 20°C and for $S = 35\%$.

Ocean surface as an impedance surface

Figure 11.7 shows a planar interface between two dielectric half-spaces with refractive indices of $n_1 = \sqrt{\mu_r \varepsilon_r}$ and $n_2 = \sqrt{\mu_r \varepsilon_r}$ illuminated by a plane wave at an angle of θ_1 from medium 1. It can be shown that the ratio of the tangential electric field to the tangential magnetic field in medium 2 at the interface ($z = 0$) is

$$\frac{E_{x,2}}{H_{y,2}} = \frac{k_2 \cos(\theta_2)}{\omega \varepsilon_0 \varepsilon_r} = Z_{TE} \quad \text{for TE polarization,} \quad (11.7a)$$

$$\frac{E_{y,2}}{H_{x,2}} = \frac{\omega \mu_0 \mu_r}{k_2 \cos(\theta_2)} = Z_{TM} \quad \text{for TM polarization.} \quad (11.7b)$$

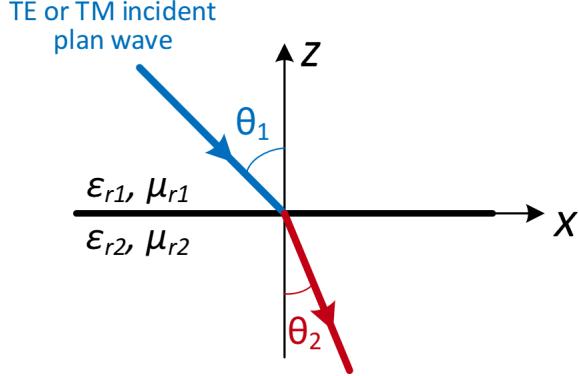
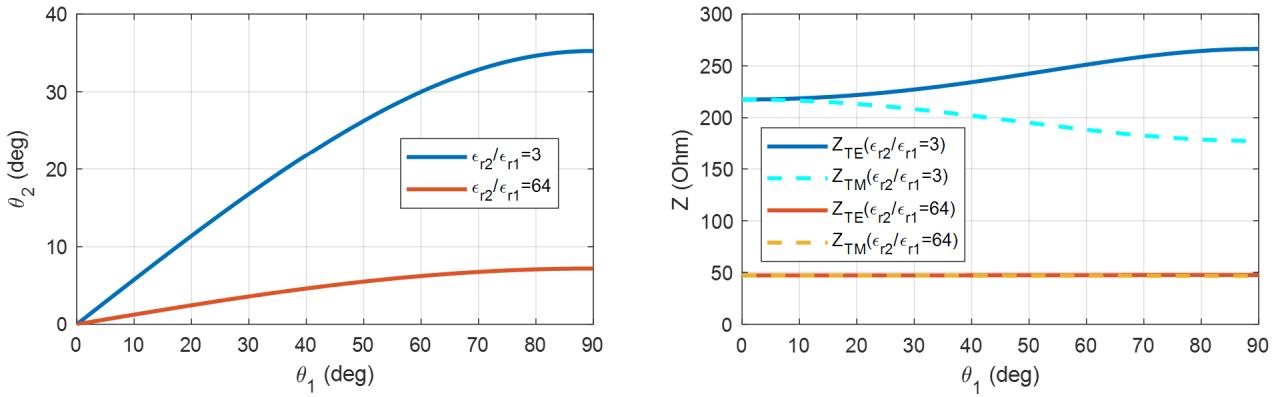


Figure 11.7: Plane wave at the interface of two dielectric half-spaces.


 Figure 11.8: Left: θ_2 and Right: Z_{TE} and Z_{TM} versus θ_1 for $\epsilon_{r2}/\epsilon_{r1} = 3$ and 64 ($n_1/n_2 = 1.73$ and 8).

In (11.7a) and (11.7b), $k_2 = \omega\sqrt{\mu_0\mu_{r2}}$ is the propagation constant in the second medium and θ_2 is related to θ_1 by Snell's law of refraction

$$\sin(\theta_2) = \frac{n_1}{n_2} \sin(\theta_1). \quad (11.8)$$

According to (11.8), if $n_2 \gg n_1$ and $\theta_2 \approx 0$, (11.7a) and (11.7b) can be written as

$$\frac{E_{x,2}}{H_{y,2}} = \frac{k_2}{\omega\epsilon_0\epsilon_{r2}} = \sqrt{\frac{\mu_0\mu_{r2}}{\epsilon_0\epsilon_{r2}}} = \eta_2 \quad \text{for TE polarization,} \quad (11.9a)$$

$$\frac{E_{y,2}}{H_{x,2}} = \frac{\omega\epsilon_0\epsilon_{r2}}{k_2} = \sqrt{\frac{\mu_0\mu_{r2}}{\epsilon_0\epsilon_{r2}}} = \eta_2 \quad \text{for TM polarization,} \quad (11.9b)$$

indicating that the relation between tangential electric and magnetic fields at the interface is independent of polarization and angle of incidence, and is described by surface impedance $Z_s = \eta_2$. In general, the impedance boundary condition for electric (\bar{E}) and magnetic fields (\bar{H}) just above the surface is given by Sarabandi (2022) as

$$\hat{n} \times (\hat{n} \times \bar{E}) = -Z_s \hat{n} \times \bar{H}. \quad (11.10)$$

In (11.10), Z_s can be complex in general and \hat{n} is normal to the interface ($= \hat{z}$ in Figure 11.7). With an impedance boundary condition, the surface impedance can replace medium 2, thereby simplifying the calculation of scattering from the surface for any arbitrary incident wave. Figure

11.8 shows θ_2 , Z_{TE} , and Z_{TM} versus θ_1 for $\varepsilon_{r2}/\varepsilon_{r1} = 3$ and 64 ($n_2/n_1 = 1.73$ and 8). As observed for all θ_1 in the case $\varepsilon_{r2}/\varepsilon_{r1} = 64$, θ_2 is less than 8° and Z_{TE} and Z_{TM} are very close and independent of θ_1 . According to Figure 11.6, the real and imaginary parts of the complex relative permittivity of ocean water are larger than 55 and 34, respectively, at frequencies less than 10 GHz, and thus, it satisfies the condition $n_2 \gg n_1$ ($n_1 = 1$) and the impedance boundary condition is applicable.

Simulation of scattering from the ocean surface

Due to the large wavelength of the gravity waves (hundreds of meters), numerical simulations of scattering from the ocean surface involve electrically large problems. As an example, an ocean surface area of $80 \text{ m} \times 80 \text{ m}$ is $1600\lambda_0 \times 1600\lambda_0$ (λ_0 is free-space wavelength) at 6 GHz (C-band). The solution of this huge problem requires a fast and efficient numerical method. The presence of capillary waves with wavelengths smaller than or comparable to those of typical frequencies used in ocean remote sensing (C-band to Ku-band) prevents the use of efficient methods like physical optics (PO). In this work, the multi-level fast multipole method (MLFMM) is used for full-wave and accurate analysis of scattering from large ocean surfaces. MLFMM is an alternative formulation of the method of moments (MoM) that reduces the computational complexity (CPU time and memory) of MoM from $O(N^2)$ to $O(N \log_2 N)$ and so it is applicable to electrically large problems. In this work, the ocean surface is modeled as an impedance boundary condition (IBC) as described above, and the surface integral equation (SIE) representing the scattering from the surface is solved by MLFMM.

Truncation of the solution region

Figure 11.9 (top left) shows the electric surface current density distribution on a large flat perfect electric conductor (PEC) surface ($10\lambda_0 \times 10\lambda_0$ at 6 GHz). The source is an x -polarized plane wave normally incident on the surface. The results are obtained by the MLFMM module in FEKO. Though the incident field is uniform and constant over the surface, the induced currents are not uniform and fluctuate over the surface. This fluctuation that produces error in finding the scattering from the surface originates from the effect of a singularity near the edges of the truncated surface. To suppress the singular behavior of the current at the edges, a resistive sheet with tapered impedance is added to the edges of the surface. The tapered impedance gradually changes from the surface impedance to the free-space impedance (Oh and Sarabandi, 1994). In Figure 11.9 (top right), a tapered impedance sheet is added to the edges of the plate. The resistive sheet is divided into eight regions, each with a width of $\lambda_0/8$. The surface impedance gradually increased from 0Ω to $Z_0 = 377 \Omega$ (free-space impedance). The surface impedance of the i th region is set to

$$Z_{s,i}(d) = Z_0 \left(\frac{d_i}{D} \right)^\alpha, \quad (11.11)$$

where, D is the thickness of the impedance sheet, d_i is the distance from the edge of the original surface to the center of the i th region, and α determine the profile of the tapered impedance. Figure 11.9 (top right) is shown for $D = \lambda_0$ and $\alpha = 3$, which exhibits much smaller variation in surface current amplitude compared to Figure 11.9 (top left). Figure 11.9 (bottom) shows the current distribution along the dashed line in Figure 11.9 (top right) for different α s and $D = \lambda_0$. According to the results, $\alpha = 3$ is an optimum value that results in the least variation in the surface current amplitude.

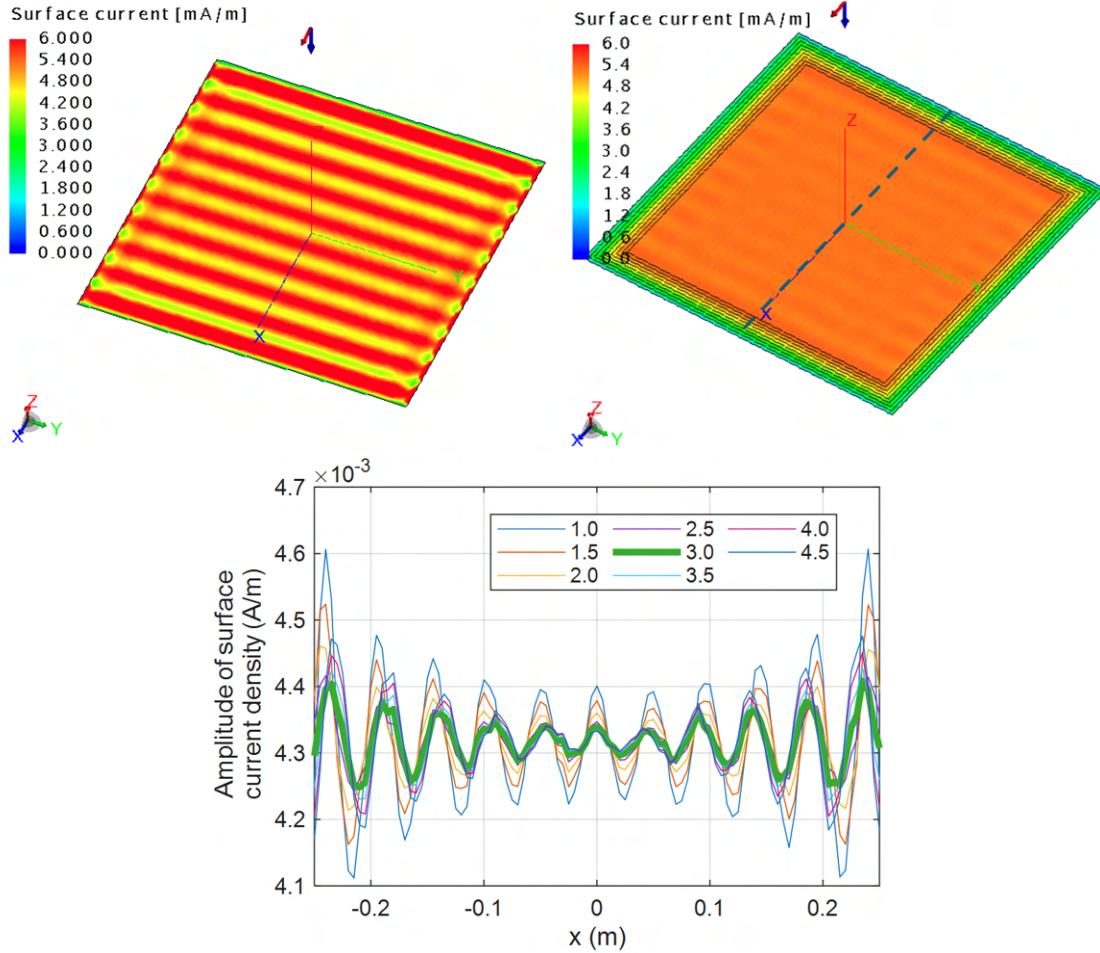


Figure 11.9: *Top left*: Electric surface current density on a $10\lambda_0 \times 10\lambda_0$ PEC plate illuminated by an x -polarized plane wave normally incident on the surface. *Top right*: Tapered impedance sheet ($\alpha = 3$ and $D = \lambda_0$) is added to the edges of the surface in *Top left*. *Bottom*: Amplitude of electric surface current density along the line in *Top right* for different α and $D = \lambda_0$.

Simulation of scattering from the ocean surface

For a selected wind speed, the ocean surface profile can be generated using the Ochi-Hubble model and the model presented for the capillary waves. For the generated surface, a resistive sheet with tapered impedance is added to the edges of the surface. The surface is then simulated using the MLFMM module of FEKO. Then the resulting scattered fields due to the incident plane wave (with a specific frequency, polarization, and angle of incidence) are recorded on a surface just above the surface. The recorded scattered fields from different surfaces are used for developing statistical models. Figure 11.10 shows a sample of $80\text{ m} \times 80\text{ m}$ ocean surface created at a wind speed of 10 m/s and a fetch of 140 km. The generated surface also includes capillary waves. To find the scattered fields, the surface is divided into 64 samples of $10\text{ m} \times 10\text{ m}$ surfaces and each is simulated with the MLFMM solver of Altair FEKO. As illustrated, the plane wave is incident from $(\theta, \phi) = (45^\circ, 180^\circ)$. Figure 11.11 shows the phase and amplitude of the tangential electric field components on a plane at a height of 1.2 m above the ocean surface for an h -polarized incident wave. The fields are shown for one of the $10\text{ m} \times 10\text{ m}$ surfaces. As it can be observed, the amplitude and phase variation of the tangential fields are random. However, tangential fields have a small deterministic

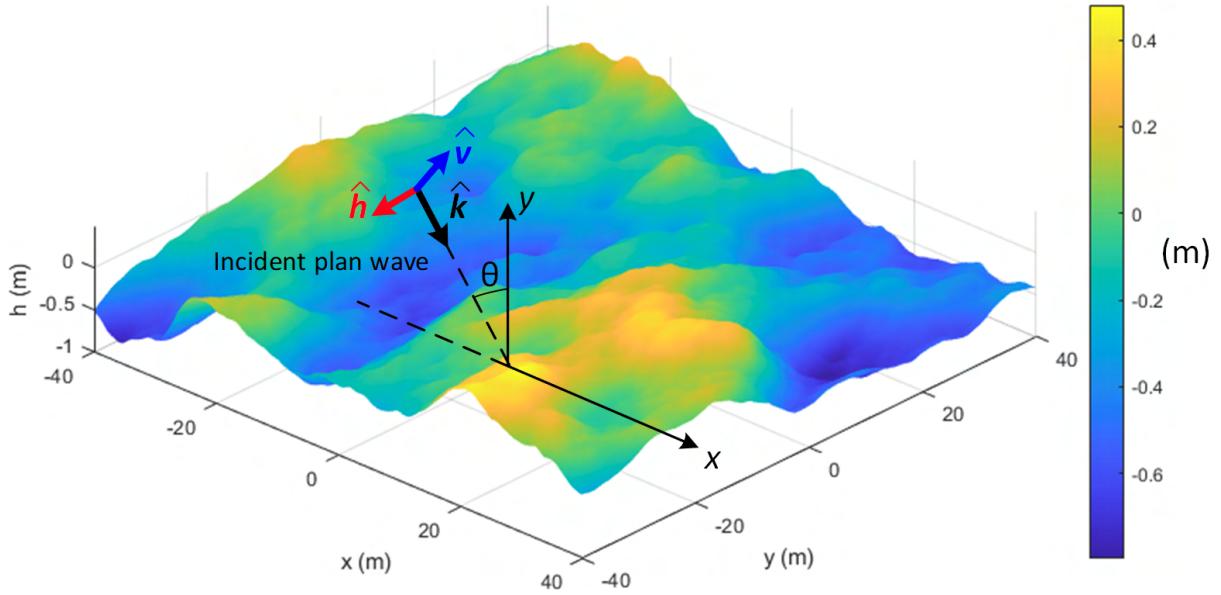


Figure 11.10: Generated ocean surface with area of $80 \text{ m} \times 80 \text{ m}$ at wind speed of $U_{19.5} = 10 \text{ m/s}$ and fetch of 140 km. The incident plane is also shown.

component, known as the mean field, which must be extracted before statistical modeling. The mean field is equal to the mean value of scattered fields among all 64 surfaces. The components of the mean field are shown in Figure 11.12. As it can be observed, its y -component (cross-polarized) is almost zero and its x -component shows almost uniform amplitude and progressive phase similar to scattering from a flat surface, which can be analytically presented. Subtracting the mean field from the total fields provides the fluctuating part of the field that must be statistically modeled.

11.4 Statistical Modeling of Scattering from a Random Ocean Surface

Statistical properties of the scattering from random surfaces

Considering plane wave illumination, if the incident (\bar{E}_i) and scattered (\bar{E}_s) fields are decomposed into vertical (v) and horizontal (h) components, then the scattered field can be related to the incident field by the scattering matrix (S)

$$\begin{bmatrix} E_{s,v} \\ E_{s,h} \end{bmatrix} = \frac{e^{jkr}}{r} \begin{bmatrix} s_{vv} & s_{vh} \\ s_{hv} & s_{hh} \end{bmatrix} \begin{bmatrix} E_{i,v} \\ E_{i,h} \end{bmatrix}. \quad (11.12)$$

The elements of the scattering matrix, s_{pq} , are complex values in general. It is shown in (Sarabandi, 1992) that the real and imaginary parts of each s_{pq} are independent & identically distributed zero-mean Gaussian random variables. It can be shown that the amplitude and phase of s_{pq} follow Rayleigh and uniform distributions, respectively. Based on the observations of scattering data from rough soil surfaces and vegetation, $s_{vh} (= s_{hv})$ are statistically independent from s_{vv} and s_{hh} and so the probability distribution function (PDF) of the real and imaginary parts ($s_{vh} = X_5 + jX_6$) or their joint PDF (= multiplication of the PDFs of the real and imaginary parts) can be described by a single parameter σ_c^2 (the variance)

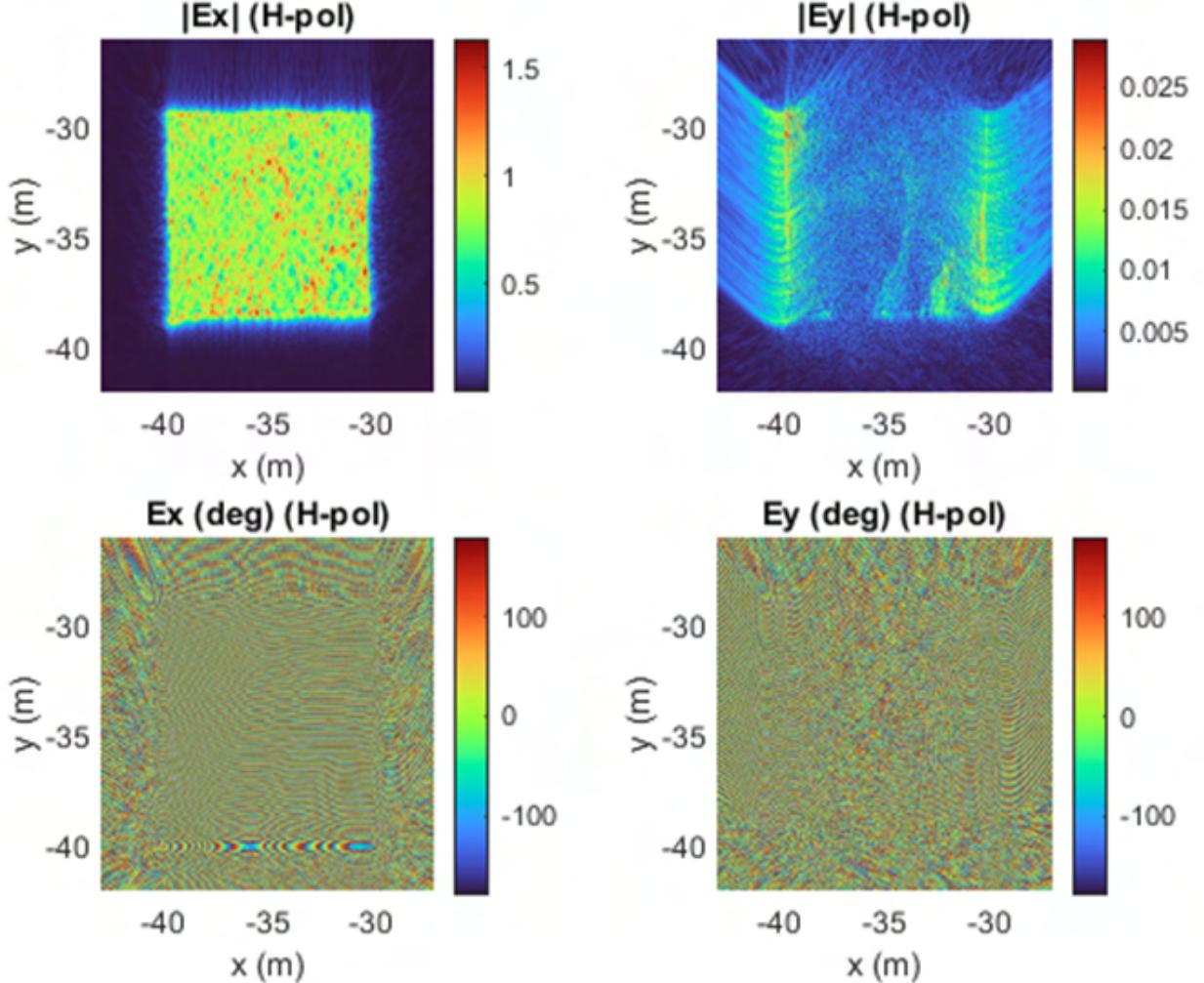


Figure 11.11: Phase and amplitude of the total tangential electric field components on a plane at $z = 0.6$ m above the surface due to h -polarized incident wave. The field is shown for 1/64 of the surface in Figure 11.10.

$$f_{X_5, X_6}(x_5, x_6) = \frac{1}{2\pi^2\sigma_c^2} \exp\left(-\frac{x_5^2 + x_6^2}{2\sigma_c^2}\right). \quad (11.13)$$

Because of the independence from s_{vv} (or s_{hh}), the phase difference between s_{vv} (or s_{hh}) and s_{vh} is uniformly distributed in the range $[-\pi, +\pi]$. The co-polarized elements of the scattering matrix are dependent and jointly Gaussian random variables, denoted by X_{1-4}

$$s_{vv} = X_1 + jX_2, \quad (11.14a)$$

$$s_{hh} = X_3 + jX_4. \quad (11.14b)$$

The joint PDF for X_{1-4} is

$$f_X(x_1, x_2, x_3, x_4) = \frac{1}{4\pi^2\sqrt{|\Lambda|}} \exp\left(X^t \Lambda^{-1} X\right), \quad (11.15)$$

where $X = [X_1, X_2, X_3, X_4]^t$ and t stands for transpose. The matrix Λ is the covariance matrix with elements defined as $\lambda_{ij} = \langle X_i X_j \rangle$. It can be proved that λ_{ij} s have the following properties

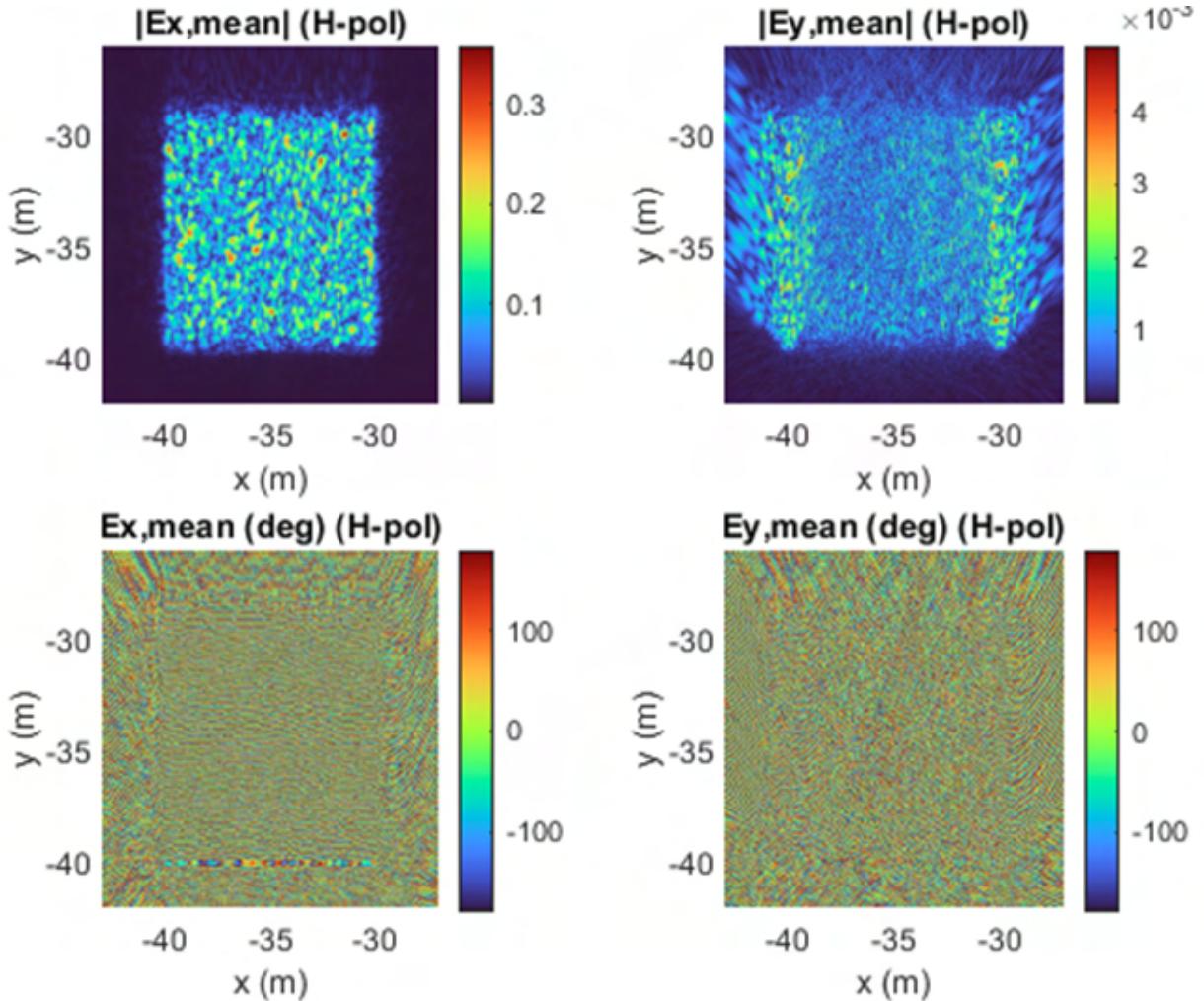


Figure 11.12: Phase and amplitude of the mean tangential electric field components due to h -polarized incident wave. The phase of the x -component is progressive, similar to reflection from a flat surface.

$$\lambda_{11} = \lambda_{22}, \quad \lambda_{33} = \lambda_{44}, \quad \lambda_{12} = 0, \quad \lambda_{34} = 0, \quad \lambda_{13} = \lambda_{24}, \quad \lambda_{14} = -\lambda_{23}. \quad (11.16)$$

Power spectral density

Knowing the power spectral density (PSD) of a field component, random complex samples of the field can be generated using the same method outlined for capillary waves. The PSD of a field component (E) is defined as

$$PSD = \left| \iint E e^{-jk_x x} e^{-jk_y y} dx dy \right|^2. \quad (11.17)$$

Figure 11.13 shows a sample of the PSD for a complex field component (E_{xh}) and for an incident wave direction of $(\theta, \phi) = (45^\circ, 180^\circ)$. The surface is shown in Figure 11.13 (bottom) and the incident wave is h -polarized. The peak in Figure 11.13 (top) indicates the direction of specular reflection in azimuth, i.e. $\phi = 0^\circ$.

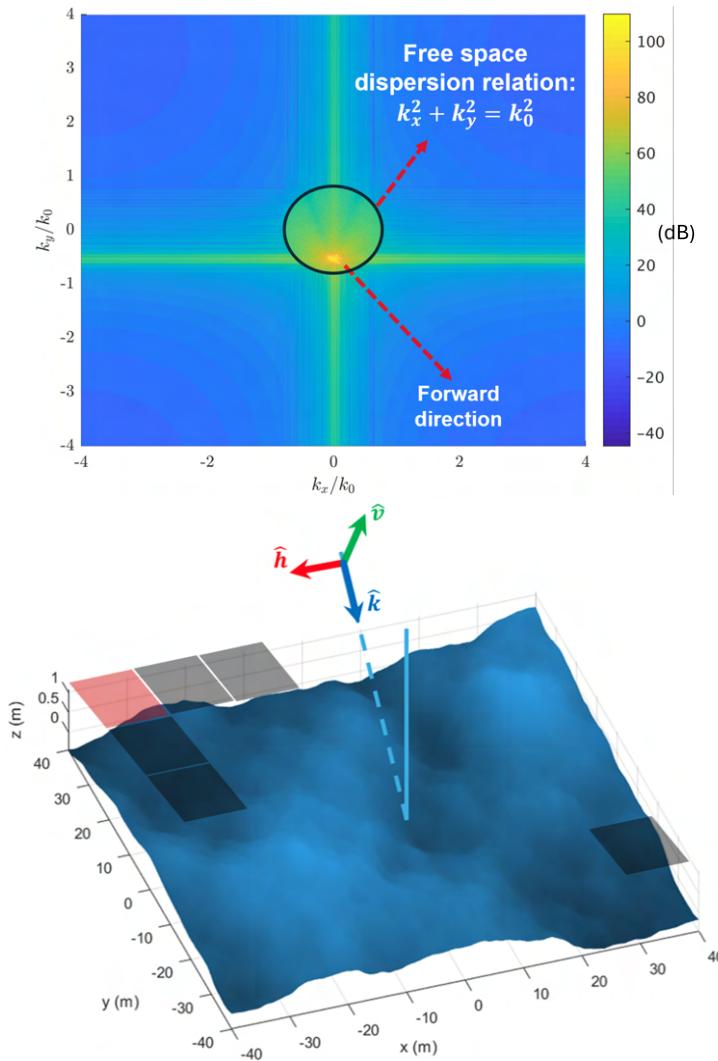


Figure 11.13: *Top*: PSD for E_{xh} scattered from the surface shown in bottom due to an h -polarized wave with angle of incidence of $(\theta, \phi) = (45^\circ, 180^\circ)$.

Procedure to develop a statistical model for scattering from a random ocean surface

The surface profile, and therefore the scattering from the ocean surface, is a function of the wind speed. The scattering also depends on the azimuthal angular difference between the incident wave and the wind direction, as well as the polarization and frequency of the incident wave. As a result, the statistical model must be developed for different wind speeds, frequency bands (e.g., C-, X-, and Ku-band), angle of incidence and wind direction differences, and incident wave polarizations. To develop a statistical model for scattering from random ocean surfaces, the following steps must be followed (wind direction is considered fixed and only the incident wave direction is varied)

1. For a fixed wind speed, incident wave angle, and frequency:
 - a) Generate at least 15 random ocean surfaces.
 - b) For each surface, calculate the co- and cross-polarized components of the scattered fields just above the surface for both h and v polarizations of the incident wave.

- c) Find the mean value of the fields at each sample point among all generated surfaces. The mean field must show a phase progression similar to scattering from a flat surface, but with a relatively small amplitude for the co-polarized wave and almost zero amplitude for the cross-polarized component.
- d) For each surface and for the fluctuating fields, find the non-zero parameters λ_{ij} describing the joint PDFs for co- and cross-polarized components (two PDFs).
- e) For each surface and for one of the co/cross-polarization components of the normalized fluctuating fields, find the PSD.

2. Repeat (1) for different wind speeds, incident wave angles, and frequencies.

Procedure to use the developed statistical model

For a given wind speed, incident wave angle, and frequency

1. Using the PSD of one of the co-polarized components of the fluctuating fields, generate the random complex samples of the field with the same filtering technique described for the generation of capillary waves.
2. Using λ_{ijs} , form the joint PDF and calculate the conditional PDFs for the phase and amplitude of the second co-polarized component of the fluctuating fields. The sample of the second co-polarized component at each point can be generated by knowing the conditional PDFs and the amplitude and phase of the first co-polarized component at the same point (from step 1).
3. Find the total co-polarized fields by adding the generated fluctuating and mean fields.
4. Repeat 1-3 for cross-polarized components.

11.5 Conclusion

An accurate and high-fidelity statistical model representing the statistical properties of electromagnetic waves scattered from a random ocean surface has been presented in this report. With the developed model, scattered electromagnetic fields can be generated above the ocean surface under specific conditions of wind speed, wind direction, and electromagnetic wave frequency and polarization. This serves as an invaluable tool for assessing the performance of communication and radar systems in the presence of the ocean, which acts as a complex random environment. The procedure to develop the model from simulations of very large ocean surfaces has been discussed. The simulation is performed by accurately modeling both large and small features of the ocean surface under various wind conditions, and solving the scattering problem using an efficient and accurate numerical method. The co/cross-polarized scattered fields are represented by joint PDFs with a few parameters and PSDs. For a complete model, the joint PDF parameters and PSDs must be found under different conditions, such as wind speed speeed and direction and incident wave frequency.

Developing a method to reconstruct the fields from the models and finding the model parameters under different conditions are future works in this project.

References

D. J. T. Carter. Prediction of wave height and period for a constant wind velocity using the JONSWAP results. *Ocean Engineering*, 9:17–33, 1982.

B. D. Flury. Acceptance-rejection sampling made easy. *SIAM Review*, 32:474–476, 1990.

K. Hasselmann, T. P. Barnett, E. Bouws, H. Carlson, D. E. Cartwright, K. Enke, J. Ewing, A. Gienapp, D. Hasselmann, P. Kruseman, et al. Measurements of wind-wave growth and swell decay during the joint north sea wave project (JONSWAP). *Ergaenzungsheft zur Deutschen Hydrographischen Zeitschrift, Reihe A*, 1973.

E. Jefferys. Directional seas should be ergodic. *Applied Ocean Research*, 9(4):186–191, 1987.

B. Kinsman. *Wind waves: their generation and propagation on the ocean surface*. Courier Corporation, 1984.

Z. Lin, T. A. A. Adcock, and M. M. L. Estimating ocean wave directional spreading using wave following buoys: a comparison of experimental buoy and gauge data. *Journal of Ocean Engineering and Marine Energy*, 8:83–97, 2022.

D. Long and F. Ulaby. *Microwave radar and radiometric remote sensing*. Artech, 2015.

H. Mitsuyasu, F. Tasai, T. Suhara, S. Mizuno, M. Ohkusu, T. Honda, and K. Rikiishi. Observation of the power spectrum of ocean waves using a cloverleaf buoy. *Journal of Physical Oceanography*, 10(2):286–296, 1980.

J. C. Novarini and J. W. Caruthers. Numerical modeling of acoustic-wave scattering from randomly rough surfaces: an image model. *Journal of Acoustic Society of America*, 53:876–884, 1973.

M. K. Ochi and E. N. Hubble. Six-parameter wave spectra. *Coastal Engineering*, pages 301–328, 1976.

Y. Oh. Microwave polarimetric backscattering from natural rough surfaces. *Ph.D. dissertation, The University of Michigan, Ann Arbor, MI*, 1993.

Y. Oh and K. Sarabandi. An improved numerical simulation of electromagnetic scattering from perfectly conducting random surfaces. *Proceedings of IEEE Antennas and Propagation Society International Symposium and URSI National Radio Science Meeting*, 1994.

K. Sarabandi. Derivation of phase statistics from the mueller matrix. *Radio Science*, 27:553–560, 1992.

K. Sarabandi. *Foundations of Applied Electromagnetics*. Michigan Publishing Services, 2022.

Machine Learning-Aided Flow Field Reconstruction from Sparse and Noisy Measurements

XIANZHANG XU, DARIA SKALITZKY & KRISHNAN MAHESH

We applied a physics-informed deep learning method to reconstruct the flow field from sparse and noisy particle tracks, focusing on accurately assimilating the minimum pressure on a fine-grid domain from very sparse particle trajectories with velocity information. We tested the performance of the deep learning model using a series of analytical vortex models in 2D and 3D. Synthetic PTV data are generated using the analytical solutions of the velocities; the deep learning model trains these sparse PTV data with velocity information, and then the model reconstructs the high-resolution flow field within a fixed domain. The results show that the physics-informed deep learning model can successfully reconstruct the fine flow field of velocity and pressure with an observation rate of less than 1%. This work also investigated the influence of particle density, distributions, and noise levels on the model's performance in pressure field reconstruction. The same approach was applied to a twin-vortex flow field based on Large-Eddy Simulation (LES), and the model successfully reconstructed the pressure field from sparse velocity measurements.

Keywords: data assimilation, physics-informed machine learning, particle tracking velocimetry, vortex flow

12.1 Introduction

Obtaining the pressure field in flows is important to a variety of naval applications. For example, cavitation inception, defined as the initial rapid growth of bubbles occurs when the minimum pressure in the fluid drops below vapor pressure (Rood, 1991). It is important to know the operating conditions at which these incepting pressures are attained. These low pressures typically occur in highly localized small scale strained vortices that are both spatially and temporally intermittent. Particle image velocimetry (PIV) and particle tracking velocimetry (PTV) are widely used in experiments to measure the flow velocities (Westerweel, 1995). However, the pressure fields are not directly measured and must be inferred from the particle measurements. Furthermore, the minimum pressure can be difficult to accurately measure due to the sparse distribution of particles around vortex cores as observed in Knister (2024). These challenges motivate the current work. Figure 12.1 shows a schematic overview of our problem. Consider a complex turbulent flow, such as the flow in the tip gap of a ducted propulsor from the simulations of (Leasca et al., 2025). Note the dominance of small scale co- and counter-rotating vortices whose interaction significantly influences the minimum pressures. We imagine a volume of interrogation illustrated by the box where Lagrangian measurements from particles are obtained. The measurements are assumed sparse, due to particles being flung out of the cores of small vortices, and noisy due to inherent experimental uncertainties. We attempt to obtain the minimum pressure by reconstructing the velocity and pressure fields on a fine Eulerian grid. We first emulate this scenario by generating

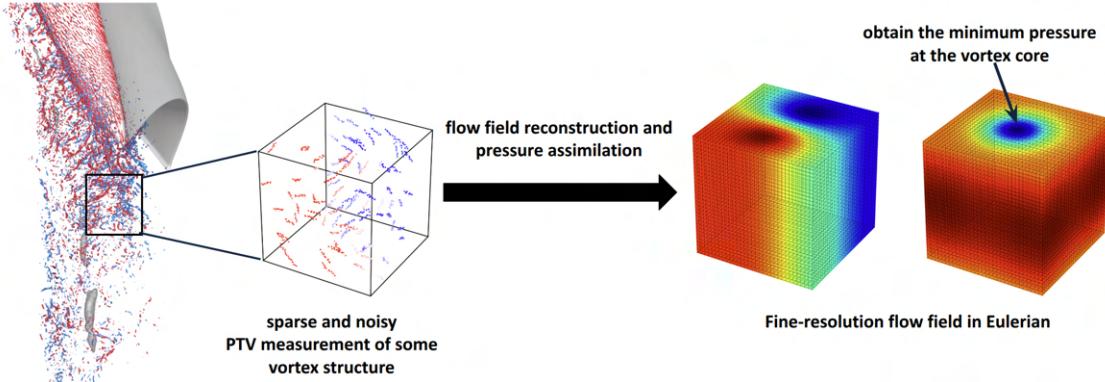


Figure 12.1: Schematic of the flow field reconstruction problem for vortex flow. The figure on the far left is from the simulations of Leasca et al. (2025).

sparse and noisy measurements of velocities for certain canonical vortex flows, and then apply the methodology to a complex turbulent flow.

Flow field reconstruction, especially the pressure field, is challenging. One of the most common approaches is isolating the pressure gradient terms from the incompressible Navier-Stokes equations and solving the Poisson equation for pressure (Van Oudheusden, 2013). Similarly, some works directly integrate the pressure gradient from the Navier-Stokes equation along a specific path in the flow field (Murai et al., 2007). The work in Liu and Katz (2006) introduced a virtual boundary integration scheme to obtain the pressure distribution by integrating the material acceleration. Another approach considers the flow field reconstruction as a data assimilation problem. For example, Zhang et al. (2020) used a generalized least squares method to incorporate the uncertainty of the pressure gradient into the optimization process to reconstruct the pressure field. Some works (Wang et al., 2019; Buchta et al., 2022) use the ensemble-based and adjoint-based variational methods. Recently, machine learning methods have attracted much attention in solving fluid dynamics problems. In Özbay et al. (2021), a convolutional neural network (CNN) is used to infer the solution of the 2D Poisson equation on a Cartesian grid. Physics-informed Neural Networks (PINNs) (Raissi et al., 2019), which is a type of deep learning model that incorporates the physical governing equations into the training loss function of the neural network, have been used in related problems. Zhou et al. (2023) and Hasanuzzaman et al. (2023) have implemented PINNs to improve the accuracy of PIV measurements and PTV reconstructions, and also to assimilate the pressure field of turbulent flow, as shown in Leoni et al. (2023). These works have shown that PINNs have the dual ability to both interpolate the velocity field and assimilate the pressure field, and that it is less sensitive to the noisy data in comparison to the traditional method. Therefore, PINNs could be useful for reconstructing the pressure field and obtaining the minimum pressure from the sparse PTV measurements. However, these previous works focus on the application to different flows and comparison to the traditional methods.

We seek a more comprehensive study on the performance of PINNs for the flow field reconstruction and the acquisition of the global minimum pressure from the sparse and noisy PTV measurements. We apply a framework based on PINNs to reconstruct the fine-resolution flow field in Eulerian from the sparse and noisy velocity measurements in either Lagrangian or Eulerian, assimilate the pressure field, and obtain the global minimum pressure. We also investigate the sensitivities of the PINN model to several uncertainty issues, including the particle densities, measurement noise, and calibration errors. A series of canonical vortex models in 2D and 3D are used as the benchmarks to study the performance of the PINN model towards the goal. Synthetic

PTV velocity data are generated using the analytical solutions of the vortex models, which are also used to generate the fine flow fields of both the velocity and the pressure for validation and testing. Further investigations are performed to test the effects of different particle distributions and the noise levels. Section 12.2 introduces the fundamental background of PINNs and presents the canonical vortex models. Then the entire framework is summarized. Section 12.3 shows the flow field reconstruction results from the synthetic PTV data. Section 12.4 discusses how particle distribution and different noise levels can affect the accuracy and the uncertainty of the flow field reconstruction. Section 12.6 concludes the report.

12.2 Methodology

This section reviews the technical details of the deep learning model used in this work, i.e., Physics-informed Neural Networks (PINNs). It also introduces the analytical vortex flow models in both 2D and 3D, with analytical expressions that are used to generate the training and verification data. Finally, the framework designed to emulate the sparse particle tracks from PTV data and then reconstruct the entire flow field of interest is introduced.

Physics-informed neural networks

Physics-informed Neural Networks (PINNs), was originally introduced to train the deep-learning model to learn the complex physical systems governed by the partial differential equations (PDEs), which are usually difficult to learn by regular neural networks (Raissi et al., 2019). PINNs are based on a multilayer perceptron (MLP) neural network architecture, which contains an input layer, multiple hidden layers, and an output layer. The input layer is the data that has been labeled with different features for the model to evaluate the corresponding output variables. The hidden layers and output layer are represented by the neurons or nodes, where each neuron calculates a linear weighted sum of its inputs plus the bias, and a nonlinear activation function to obtain the output and passes it to the next layer until the output layer is reached. The output layer is also represented by neurons, whose total number is determined by the problem of interest. The number of neurons and hidden layers, however, is usually determined by the model's performance and the problem's complexity. MLP neural networks can be considered as each neuron seeking an optimal fitting that contains the process,

$$z = \mathbf{w}\mathbf{x} + b, \quad a = \sigma(z), \quad (12.1)$$

where \mathbf{w} is the weights of each input neuron \mathbf{x} , b is the bias, and σ is the activation function, which is non-linear. Some of the most commonly used activation functions include hyperbolic tangent, ReLU, sigmoid, gelu, etc.. Therefore, the MLP neural networks are expected to fit both linear and nonlinear models.

The regular MLP neural networks are trained by the loss function, which evaluates the deviations from the outputs of the model and the actual reference data. Therefore, regular MLP neural networks usually take a large amount of data to train the model, obtain the main features of the model by optimizing the weights and bias parameters of each neuron that give the best results, i.e., the minimum residuals of the loss function. After that, another input and output dataset is used to validate the model. The optimization process is performed using the gradient descent method. It works by computing the gradient of the loss function with respect to each weight and then updating the weights in the opposite direction of the gradient to gradually reduce the error. To efficiently compute these gradients for all layers, the backpropagation algorithm is employed. Backpropagation applies the chain rule of calculus to propagate the error backward through the

network, layer by layer, calculating how much each weight contributed to the final error. This combination of gradient descent for optimization and backpropagation for efficient gradient calculation allows neural networks to learn complex patterns from data.

A main feature that distinguishes PINNs from regular MLP neural networks is the loss function. Since many complex physical systems are governed by the PDEs that also require the computation of derivatives of key features with respect to the spatial and temporal variables, the PINN model incorporates these PDEs into the loss function that is used to train the neural networks. Besides the PDEs, it is also important to know the boundary conditions and initial conditions to get the correct solutions that satisfy the PDEs; therefore, the loss function of PINNs also includes the terms that compute the boundary conditions and initial conditions. The total loss function can be expressed as

$$L = \lambda_d L_d + \lambda_p L_p + \lambda_b L_b. \quad (12.2)$$

Here $\lambda_{d,p,b}$ are the weight coefficients corresponding to the loss of data L_d , loss of physical equations (PDEs) L_p , and loss of the boundary and initial conditions L_b . Each term in equation 12.2 can be independent of the others, which means that the sample points used to calculate the data loss are not necessarily the same as the points used to calculate the PDE loss or the boundary loss. This makes PINNs flexible to the data required to train the neural networks, where the regular neural networks often require sufficient good-quality data. When no training data are available, the loss function of PINNs corresponds to solving a PDE with known boundary conditions. When some training data are available, PINNs can be trained with those sample data and the PDE loss to assimilate the unknown variables in the output layer. More details of the data assimilation will be discussed in the later sections. Since the governing PDEs generally compute the derivatives of the physical parameters, such as velocity, temperature, and forces with respect to the spatial and temporal coordinates, the architecture of PINNs is designed to take the spatial and temporal coordinates as the input of the model, and the physical parameters of interest as the outputs. For the fluid problem, it is natural to take the Navier-Stokes equations as the physical loss equation, and the velocities and the pressure as the outputs when the external forces are negligible. Figure 12.2 shows an example of the architecture of PINNs when the 3D incompressible Navier-Stokes equations are used as the physics information. Here, the inputs are the three spatial coordinates x, y, z and the temporal coordinates t , and the outputs are the velocities u, v, w and pressure p corresponding to the inputs. The outputs are post-processed by computing their derivatives and plugging them into the momentum equations in 3 directions and the continuity equation to satisfy the incompressibility condition. The parameters of the neural networks are updated by the backpropagation to make sure that the gradient of the total loss function keeps decreasing throughout iterations.

Analytical vortex models

We use three analytical vortex models in both 2D and 3D as benchmarks to investigate the PINN model in flow field reconstruction. These vortex models with analytical solutions can be used to generate both the synthetic PTV measurements in Lagrangian for training and also the fine-resolution flow field data in Eulerian for validation. Here, we briefly introduce all the vortex models.

a) Taylor-Green vortex

The Taylor-Green vortex is an exact, time-dependent analytical solution to the 2D incompressible Navier-Stokes equations. The velocity field in Cartesian coordinates of the Taylor-Green Vortex

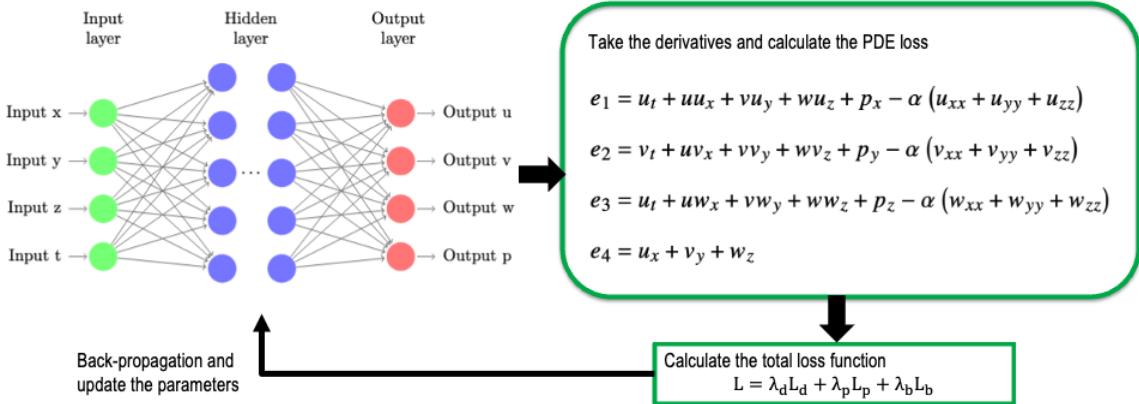


Figure 12.2: The work flow of PINNs for 3D incompressible Navier-Stokes equations.

is given as

$$\begin{aligned} u(x, y, t) &= -U \cos(x) \sin(y) e^{-2\nu t}, \\ v(x, y, t) &= U \sin(x) \cos(y) e^{-2\nu t}, \end{aligned} \quad (12.3)$$

where U is a characteristic velocity and ν is the kinematic viscosity. The pressure field can also be derived from the velocity field expression in equation 12.3,

$$p(x, y, t) = p_0 - \frac{\rho U^2}{4} (\cos(2x) + \cos(2y)) e^{-4\nu t}, \quad (12.4)$$

where ρ is the density of the fluid and p_0 is a reference pressure.

b) Lamb-Oseen vortex

The second analytical vortex model used is the Lamb-Oseen vortex, a solution to the Navier-Stokes equations describing the diffusion of an initially concentrated axisymmetric vortex in a viscous fluid. The velocity field of the Lamb-Oseen vortex is given in cylindrical coordinates:

$$\begin{aligned} v_r &= 0, \\ v_\theta(r, t) &= \frac{\Gamma}{2\pi r} \left(1 - e^{-r^2/(4\nu t)} \right), \\ v_z &= 0, \end{aligned} \quad (12.5)$$

where ν is the kinematic viscosity and Γ is the initial circulation. Since the velocity in the z direction is zero, it can be treated as a 2D flow. To obtain the pressure field data for validation, the velocity expressions in equation 12.5 are plugged into the Navier-Stokes equations in cylindrical coordinates, and then the pressure field can be obtained by integrating the expression

$$\frac{\partial p}{\partial r} = \rho \frac{v^2}{r}, \quad (12.6)$$

where v is the velocity and ρ is the density of the fluid. Lamb-Oseen vortex flow models the structure of real vortices that have a small core surrounded by a nearly inviscid region, and it includes the unsteadiness through viscous diffusion. Therefore, it is a good basic flow to study the physics of vortices. The Lamb-Oseen vortex model is also selected as the canonical vortex flow to help

understand how the distribution and density of the particles affect the flow reconstruction results, where more details will be discussed in Section 12.4.

c) Burgers' vortex

Burgers' vortex is also expressed in cylindrical coordinates (r, θ, z) . It is an exact solution to the Navier-Stokes equations in 3D. Burgers' vortex model describes a steady axisymmetric vortex in a viscous fluid and is subjected to a constant axial stretching rate $a > 0$. Since the stretching balances viscous diffusion, the vortex is in a steady state. The velocity field of the Burgers' vortex is given as

$$\begin{aligned} v_r &= -ar, \\ v_\theta(r) &= \frac{\Gamma}{2\pi r} \left(1 - e^{-ar^2/(2\nu)} \right), \\ v_z &= 2az, \end{aligned} \tag{12.7}$$

where ν is the kinematic viscosity and Γ is the circulation. Similar to the Lamb-Oseen vortex, the differential equation of the pressure can be expressed by plugging the velocity solution into the Navier-Stokes equation, and then the pressure field can be obtained by integrating equation 12.8 in the r and z directions,

$$\frac{\partial p}{\partial r} = -a^2 r + \frac{v_\theta^2}{r}. \tag{12.8}$$

Burgers' vortex flow is selected as a basic vortex flow to study the vortex flow in 3D, which also contains the axial stretching effects.

Flow field reconstruction framework

The objective of this work is to reconstruct the flow field from sparse PTV data, where both the velocity fields and the pressure field are expected in fine resolution. Therefore, PINN is used to solve this problem. The input variables of the PINN model are the spatial coordinates and the time steps, and the output variables are the velocities and the pressure. In the training procedure of the 3D problem, for example, the sparse PTV data which includes (x, y, z, t) and (u, v, w) is given to the model to train. As addressed in the previous sections, here the pressure information is totally unknown and set to zero. Therefore, the model only uses the sparse velocity data to learn the maps between the spatial-temporal coordinates and the flow field quantities with the given physical information, i.e., the incompressible Navier-Stokes equations. In addition to the PTV data, the training samples also collect random points inside the domain of interest where no output information is available. These points are used to compute the physical loss only. In some PINNs works, the sample points also contain the points at boundaries and initial time, but they are not necessary for this work since the PTV data has already provided the data loss contribution. The training procedure ends when the training loss converges or reaches the predefined iterations. Then the trained model is evaluated at a fine resolution of the spatial coordinates; in this work, only the spatial grid sizes are different between the training and testing procedures. The time step sizes are the same, but they could be different as desired by the problem of interest. The velocities and pressure on the fine coordinates predicted by the PINN model will then be compared to the pre-generated validation data, and the \mathcal{L}_2 relative errors for each variable are calculated.

Figure 12.3 summarizes the framework using the Burgers' vortex as an example. The same framework is applied to all the study cases with synthetic PTV data in this work. More details of the data generation process are introduced in Section 12.3.

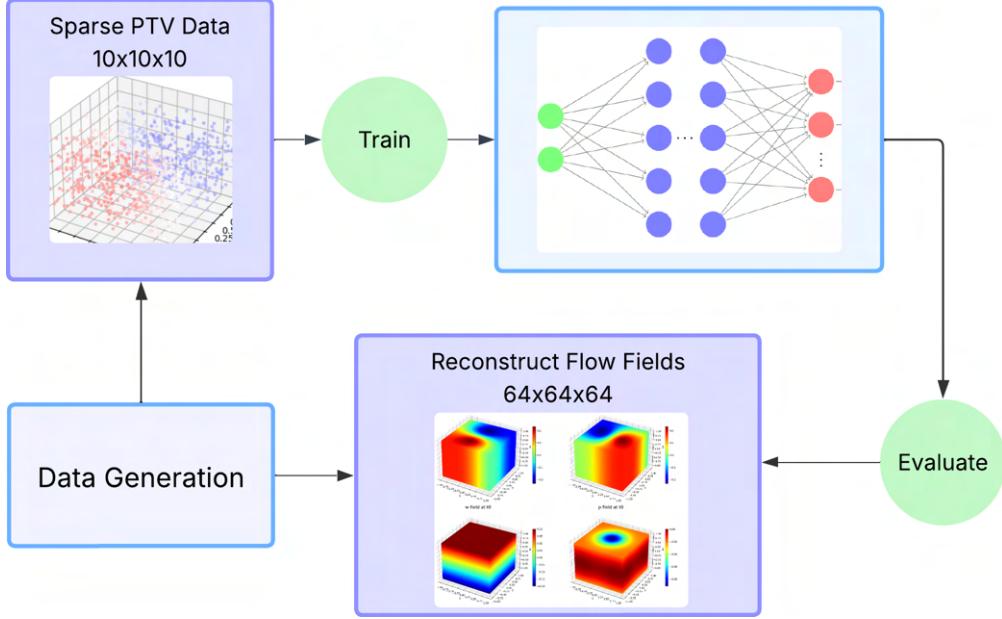


Figure 12.3: Flow field reconstruction framework.

12.3 Synthetic PTV data generation and flow field reconstruction results

In this section, the data generation process for training and validation is introduced. The flow field reconstruction results of the three vortex models are also presented and discussed.

Data generation

In this work, we aim to emulate the situation when the particles are sparse and may be away from the low-pressure core. Therefore, the analytical solution of the vortex models can be used to generate synthetic PTV tracks for training. PTV data collects the velocity information using the Lagrangian method, where the motion of individual tracer particles is tracked through successive image frames to reconstruct their trajectories. Therefore, the synthetic PTV data is initially generated by a low density, such as 10×10 , where each dimension has the corresponding number of random coordinates and results in a total of 100 points. After that, at each time step, every single particle is moving one time step forward based on the velocity, which gives the particle position for the forward time step as

$$\begin{aligned} x(t_{k+1}) &= x(t_k) + \int u(t_k) d\tau \\ y(t_{k+1}) &= y(t_k) + \int v(t_k) d\tau. \end{aligned} \quad (12.9)$$

Here, the forward Euler method is used since the main focus of this work is not on modeling the PTV tracks, and a total of 21 time steps are sampled to generate the training data.

Figure 12.4 shows an example of the PTV particles generated for Taylor-Green vortex with a density of 10×10 in a domain of $[-\pi, \pi]$, with $U = 1$, and $\nu = 0.01$. In Figure 12.4a, the initial distribution of the particles and the corresponding velocities are shown, and Figure 12.4b shows the entire trajectories with the color map indicating the v velocity. The time window is from 0.1 to 2, with a total of 21 time steps. It is clear that the particle trajectories align well with the

analytical solution of the Taylor-Green vortex. Note that the particles are not set to be periodic, so that some particles could move outside the domain of interest. The figure only shows the particles inside the domain. The pressure data are set to be unknown for training, since the PTV data from the experiments only output the velocities, and the objective of this work is to assimilate the pressure field from the sparse PTV tracks.

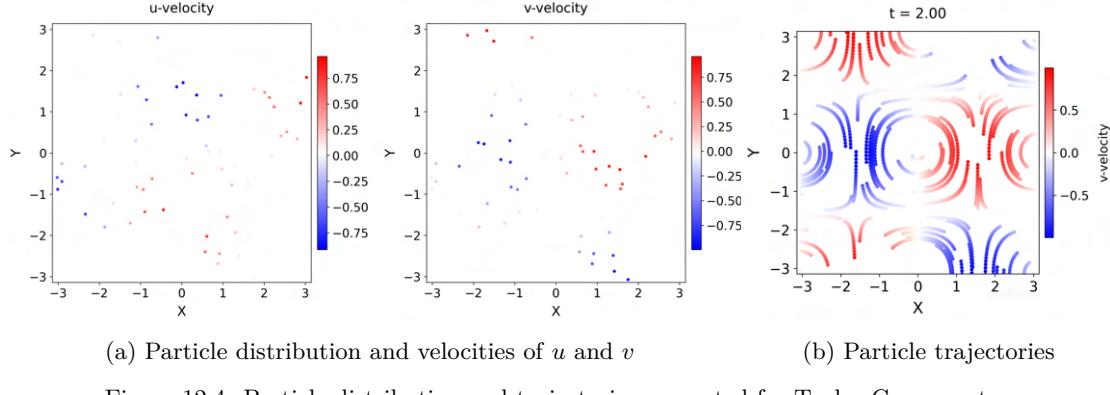


Figure 12.4: Particle distribution and trajectories generated for Taylor-Green vortex.

Similar to the Taylor-Green vortex flow, the training data for the Lamb-Oseen vortex flow are generated by initially sampling 10×10 random points and tracking their positions via the analytical solution of the velocities. Figure 12.5 shows the particle distribution of the Lamb-Oseen vortex flow at one instant time-step. The color map indicates the velocities of u and v of the particles. The velocities are translated from cylindrical coordinates to Cartesian coordinates to be consistent with other vortex models and also easy to follow. These 100 points are randomly generated, and it can be seen that they are sparsely distributed in the domain. The domain size is $[-1, 1]$ in both x and y dimensions, the kinematic viscosity is set to 0.01, and the initial circulation strength Γ is set to 1. The time window is also $[0.1, 2]$ with a total of 21 steps. Similarly, here only the particles inside the domain of interest are shown, and those outside the domain are omitted. For the Taylor-Green vortex and Lamb-Oseen vortex, the validation data is generated inside the same domain but with fixed coordinates in Eulerian, with a fine resolution of 64×64 .

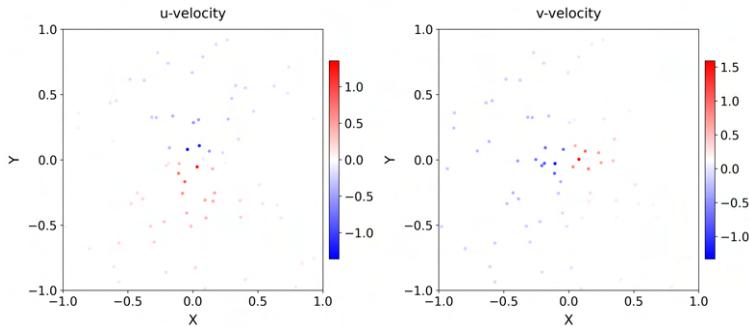


Figure 12.5: Particle distribution for the Lamb-Oseen vortex at the initial time steps.

Figure 12.6 shows the particle distribution at an instant time step. The entire time window is also $[0.1, 2]$ with a total of 21 steps. The flow domain size is $[-1, 1]$ in each dimension. The stretching rate is set as $a = 0.1$, the circulation $\Gamma = 1$, and the kinematic viscosity is $\nu = 0.01$. The density of particles is $10 \times 10 \times 10$, where 10 random coordinates are randomly generated on each dimension, respectively, then they are combined together, resulting in a total of 1000 sample points

at each time step. For the validation data, the flow field is the same size but with a finer resolution

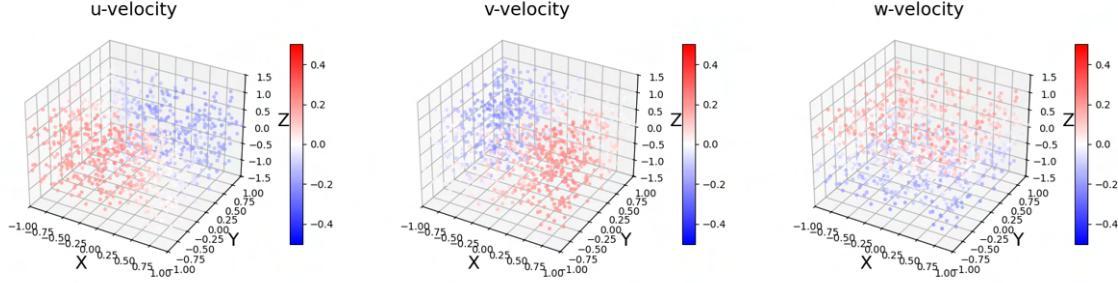


Figure 12.6: Particles generated for Burgers' vortex at one time step.

Overall, three analytical vortex models are selected to study the flow-field reconstruction from sparse particle tracks using PINNs. These three vortex models include the canonical benchmark with an exact solution of the pressure field, the unsteady 2D vortex model with diffusion effects, and the 3D vortex model with stretching effects. As mentioned before, the training data is generated as sparse particles in the Lagrangian method, and the validation data is generated with finer resolution in the Eulerian method.

Results of Taylor-Green vortex

The first case studied is the Taylor-Green vortex model, where the domain is set to be $[-\pi, \pi]$, with $U = 1$, $\rho = 1$, and $\nu = 0.01$. The density of training particles is 10×10 and the testing resolution is 64×64 . Figure 12.7 shows the velocity fields reconstruction results and the comparison to the reference validation data. The outputs from PINNs are almost identical to the ground truth.

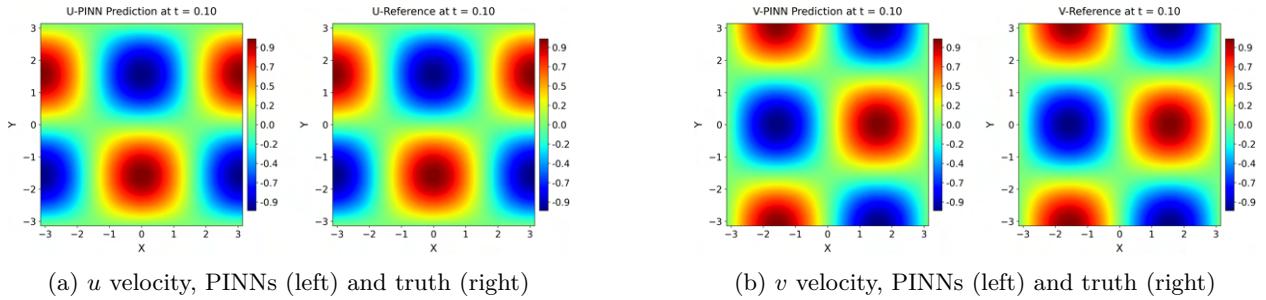


Figure 12.7: Velocity field reconstruction of Taylor-Green vortex.

For the pressure field shown in Figure 12.8a, the prediction result from PINNs also matches well with the reference ground truth. Figure 12.8b shows the pressure distribution along r , where the PINNs result (left) is nearly the same as the ground truth (right), although some tiny deviations can be observed, which make the result slightly noisier than the ground truth.

Overall, the flow field reconstruction results of PINNs for the Taylor-Green vortex show a high accuracy performance. Table 12.1 summarizes the key quantities of interest in this work for flow field reconstruction. The minimum pressure predicted by PINNs is very close to the reference value, and the overall \mathcal{L}_2 errors in all three fields are very small.

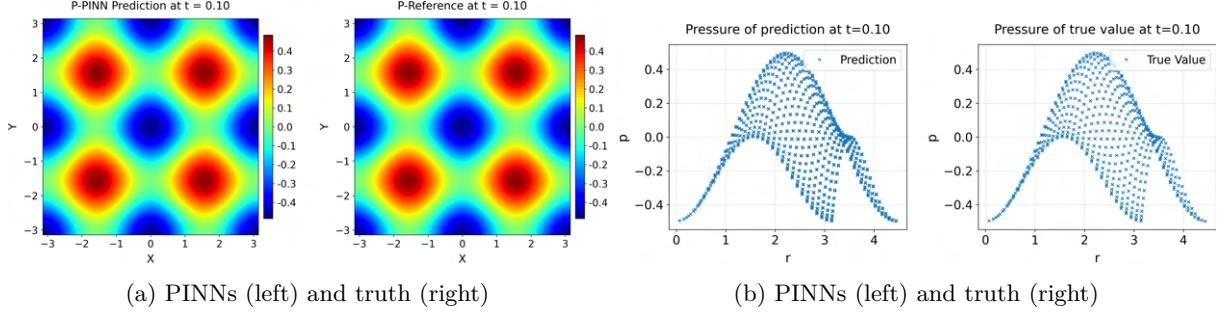


Figure 12.8: (a) Pressure field reconstruction and (b) pressure distribution of the Taylor-Green vortex.

| $\min p$ (ref) | $\min p$ (PINNs) | errors in u | errors in v | errors in p |
|----------------|------------------|---------------|---------------|---------------|
| -0.498 | -0.497 | 0.17% | 0.17% | 0.49% |

Table 12.1: Summary of the flow field reconstruction results of the Taylor-Green vortex.

Results of Lamb-Oseen vortex

The Lamb-Oseen vortex can also be treated as a 2D problem since the velocity in z is zero. With the parameters set in Section 12.2, and also the same density of sampled particles (10×10), the flow reconstruction results of the velocity fields are shown in Figure 12.9. The results show a good agreement between the PINN model and the reference ground truth.

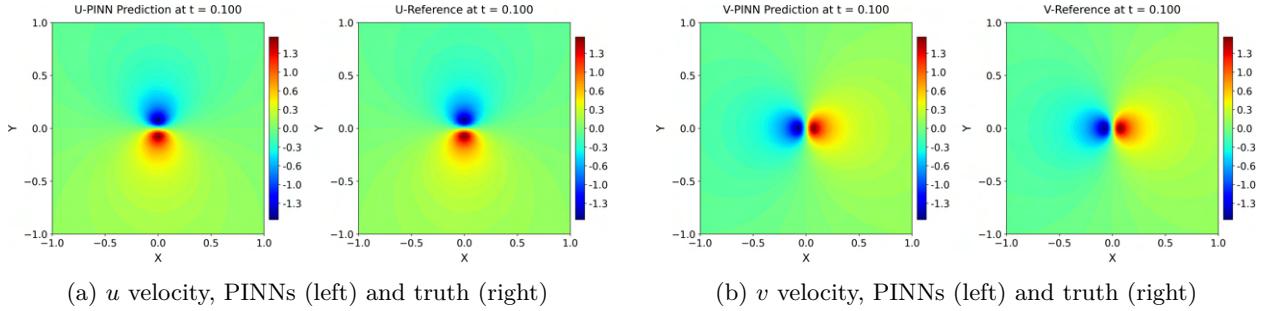


Figure 12.9: Velocity field reconstruction of Lamb-Oseen vortex.

The pressure field results shown in Figure 12.10 indicate that the PINN model successfully captures the patterns of the pressure field, where the minimum pressure point is located at the center of the domain and the pressure values symmetrically diffuse along the r direction. Here, only the pressure field at the initial time ($t = 0.1$) is shown since the pressure values increase over time, and the minimum pressure is obtained at the initial time. It is also observed that the prediction results for later steps show better performance than at the initial time since the pressure value is greater. Therefore, the minimum pressure value at the initial time is considered as a reference to the pressure field reconstruction performance. As shown in Figure 12.10b, the general trend of pressure distribution is well captured by PINNs, although the minimum pressure at the center is larger than the ground truth.

Table 12.2 summarizes the flow reconstruction results for the Lamb-Oseen vortex. The relative errors in u and v are still very small, while the errors in p are larger but acceptable. This is also revealed by the fact that the minimum pressure predicted by the PINN model is greater than the ground truth value.

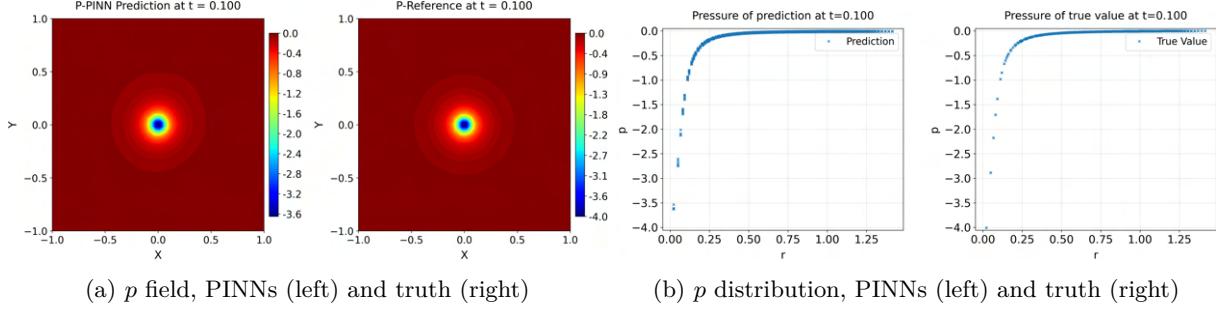


Figure 12.10: Pressure field reconstruction and distribution of Lamb-Oseen vortex.

| $\min p$ (ref) | $\min p$ (PINNs) | errors in u | errors in v | errors in p |
|----------------|------------------|---------------|---------------|---------------|
| -4.0 | -3.63 | 1.01% | 1.32% | 7.62% |

Table 12.2: Summary of the flow field reconstruction results of the Lamb-Oseen vortex.

Results of Burgers' vortex

The Burgers' vortex is used to study the flow field reconstruction performance for the 3D flow field reconstruction problems. The density of particles is $10 \times 10 \times 10$ and the reconstruction field has a resolution of $64 \times 64 \times 64$, which makes the observation rate as low as 0.38%. Since Burgers' vortex is steady, the choice of time step does not affect the results. Here in Figure 12.11, the flow field reconstruction results are presented for the initial time step. The results show that, given the very low observation rate, the PINN model is able to reconstruct the whole flow field in 3D, where the prediction results are nearly identical to the reference ground truth.

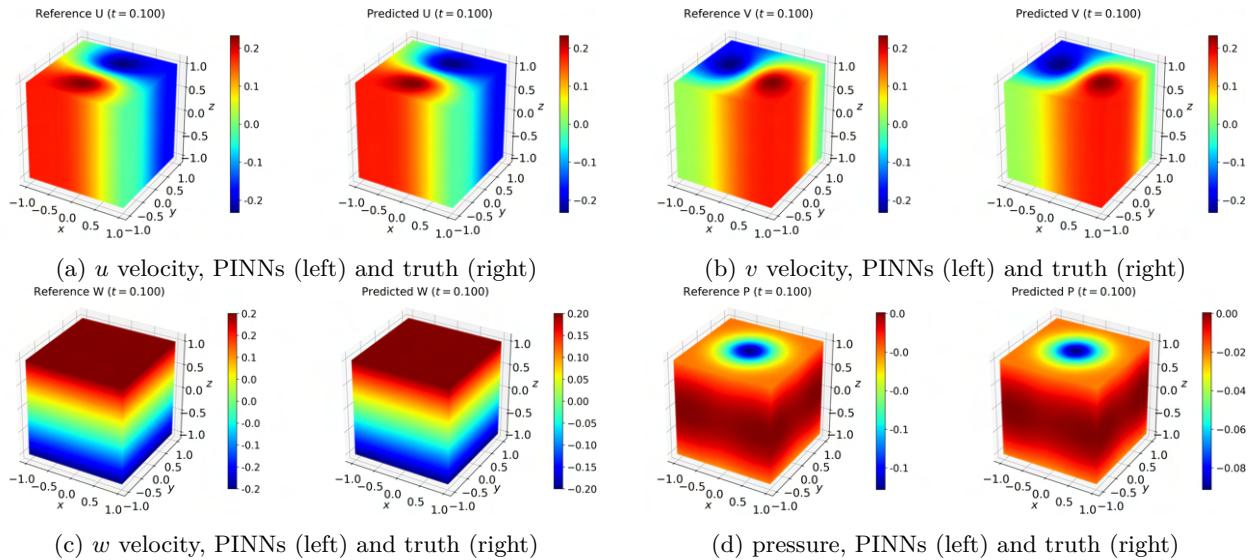


Figure 12.11: Flow field reconstruction of Burgers' vortex.

To further examine the pressure field reconstruction results, Figure 12.12 shows the pressure field and distribution along the r direction on the plane at $z = 0.02$. The prediction results match exactly with the reference ground truth.

Table 12.3 gives the quantified results in terms of the flow field velocities and the pressure. The minimum pressure of the prediction is nearly the same as the reference truth. The \mathcal{L}_2 relative errors

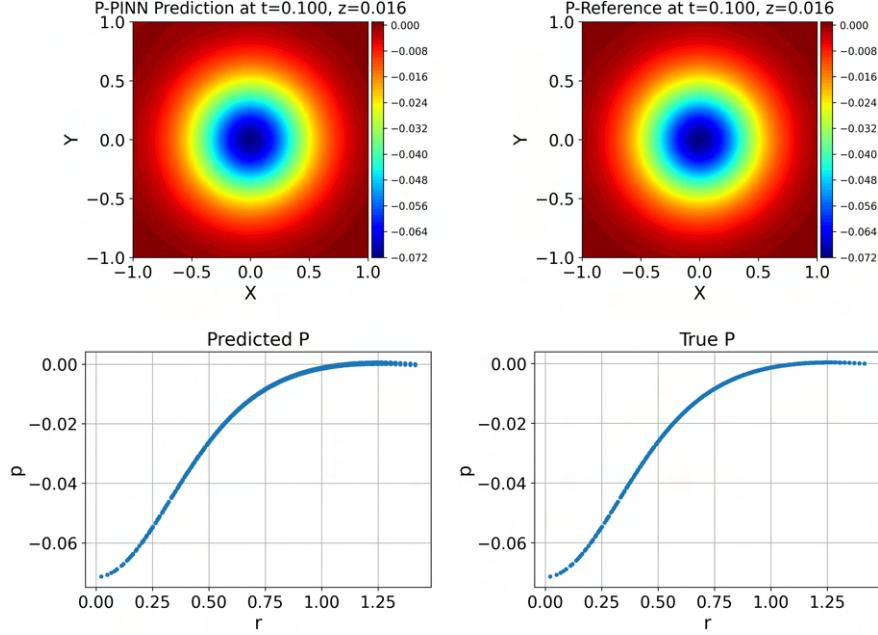


Figure 12.12: Pressure Distribution on a 2D plane of the reconstructed Burgers' vortex flow field.

| $\min p$ (ref) | $\min p$ (PINNs) | errors in u | errors in v | errors in w | errors in p |
|----------------|------------------|---------------|---------------|---------------|---------------|
| -0.0912 | -0.0913 | 0.78% | 0.11% | 0.12% | 0.32% |

Table 12.3: Summary of the flow field reconstruction results of the Burgers' vortex.

are all small and almost negligible. These results indicate that even though the Burgers' vortex is a 3D problem, which is expected to be more complicated than the 2D problem, the PINN model can still successfully reconstruct the flow field from the sparse PTV data. A possible reason could be that the Burgers' vortex is steady compared to the unsteady Lamb-Oseen vortex, which makes the problem easier to solve since the flow does not diffuse over time. Note that, the Taylor-Green vortex is also unsteady, but the flow does not diffuse in the space over time. Future work plans to investigate and quantify the diffusion effects to the flow field reconstruction.

12.4 Model performance on different sample distribution and noise level

In Section 12.3, the flow field reconstruction results of the three analytical vortex models are presented. The proposed method works well with the Taylor-Green vortex and Burgers' vortex, even when the density of the sampled particles is low, as the observation is less than 1%. However, for the Lamb-Oseen vortex, the errors in pressure are slightly higher, and the value of the minimum pressure is also higher than the reference. In order to better understand and investigate the potential effects on the performance of the PINN model in flow field reconstruction, a series of numerical experiments is conducted and presented in this section. Section 12.4 studies the effects of sampled particles' distribution and density on the model prediction accuracy based on the Lamb-Oseen vortex only. This section also discusses the effects of the sampling noise for all three vortex models.

Sample distribution and density

The flow field reconstruction results in Section 12.3 are obtained from the randomly sampled particles in a density of 10×10 . Therefore, it is natural to explore how density affects the accuracy of the minimum pressure. Besides, the particles distribution may also have an influence. In this work, two different methods are used to generate the initial locations of the particles. In equation 12.5, the velocity field of the Lamb-Oseen vortex is a function of r ; therefore, it is worth comparing the randomly and uniformly distributed radius r . As shown in Figure 12.13, the uniform distribution means that the particles are generated with equal distance from the core radius $r_c = \sqrt{4\nu t}$ to the margin of the domain $\sqrt{2}L$, where L is half of the domain length. For example, when the sample density is 10×10 , it generates 10 coordinates equally in the range of $[4\nu t, \sqrt{2}L]$, and then 10 coordinates for angle θ randomly in the range of $[0, 2\pi]$, which gives 100 sample points through the combination of r and θ . The random distribution, however, has 10 random coordinates of both r and θ in the same range.

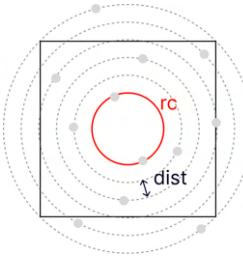


Figure 12.13: Illustration of the particle distribution of the Lamb-Oseen vortex.

The experiments cover 4 cases, with random and uniform radius distribution for the density of 10×10 and 20×20 . Other properties of the Lamb-Oseen vortex are the same as the settings in Section 12.3. For each case, the model is run 3 times and the average results are shown in Table 12.4. The results on velocities are almost the same for all cases since they are nearly the same as the ground truth as well. For the pressure, the uniform radius cases have better results than the random radius cases for both particle densities. Higher density also shows more accurate results for both the minimum pressure and the total relative errors.

| Density 10×10 | min p | errors in p | min/max u | min/max v |
|------------------------|---------|---------------|--------------|--------------|
| random r | -3.36 | 12.6% | -1.55 / 1.53 | -1.54 / 1.50 |
| uniform r | -3.62 | 10.75% | -1.53 / 1.52 | -1.53 / 1.54 |
| Density 20×20 | min p | errors in p | min/max u | min/max v |
| random r | -3.64 | 9.18% | -1.56 / 1.55 | -1.55 / 1.55 |
| uniform r | -3.72 | 9.15% | -1.54 / 1.54 | -1.54 / 1.53 |

Table 12.4: Summary and comparison of the random and uniform radius distributions.

Sampling noise in velocity

Previous results show that the PINN model has good performance based on the synthetic PTV data. In the real-world experiments, however, the PTV data could be noisy due to multiple reasons. This section presents the flow field reconstruction results from noisy data, where two different noise

sources are studied. The first one is the noise on the velocities, which can be caused by measurement errors and uncertainties. The second one is the noise on the spatial coordinates, which can be some errors in the calibration of the particle locations. For the first problem, the particle locations are assumed to be accurate, while Gaussian white noise is added to the velocities.

$$\bar{u} = u + \epsilon_u, \quad \bar{v} = v + \epsilon_v, \quad \bar{w} = w + \epsilon_w, \quad \text{where } \epsilon \sim \mathcal{N}(0, \sigma). \quad (12.10)$$

Three different noise levels are tested, i.e., $\sigma = 0.01$, $\sigma = 0.05$, and $\sigma = 0.1$, which correspond to small, medium, and heavy noise, respectively. The same approach is applied to the noise of spatial coordinates, so that the equations are omitted here. The experiment tests are run for all three vortex models, and for each noise level and vortex model, three runs are taken to check the uncertainties due to the noise, and also to better examine the performance of the PINN model. All the parameters are set the same as Section 12.3, and the low density of particles is used. Figure 12.14 shows the comparison of the velocity histories of a single particle at different noise levels. When the noise level σ increases, significant changes and disturbances in the velocity can be seen. For heavy noise, the time history trajectory looks very different from the clean data without noise.

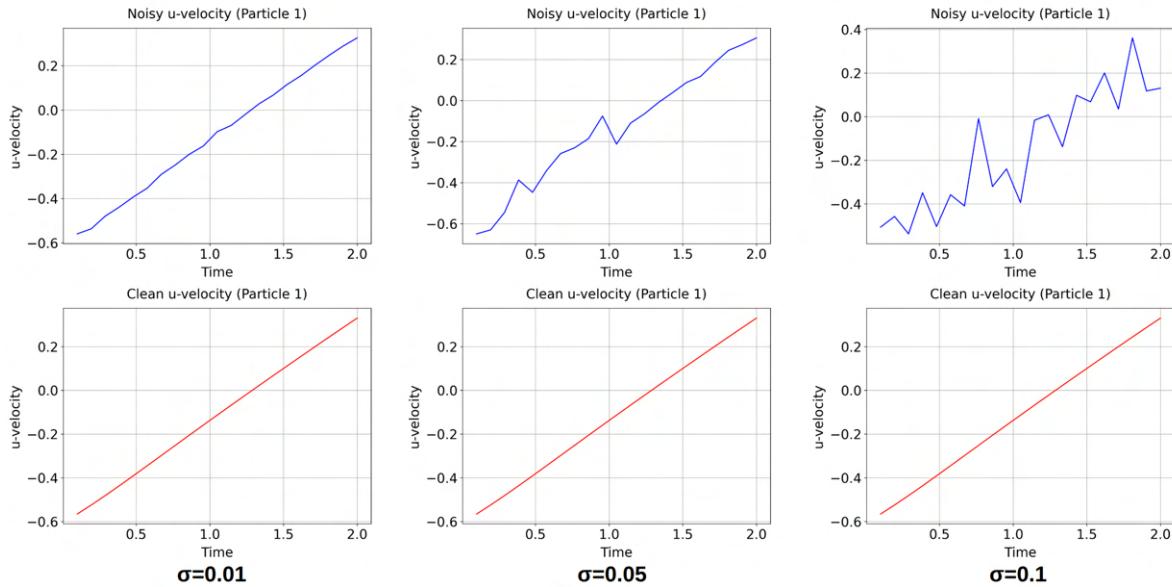


Figure 12.14: Velocity histories of a single particle of Taylor-Green vortex.

The pressure field reconstruction results for the Taylor-Green vortex at different noise levels are shown in Figure 12.15. Since the velocity fields always have good results, they are also omitted here. The results indicate that even for heavy noise, the PINN model can still reconstruct the pressure field with acceptable accuracy. For the specific value of the minimum pressure and the relative errors in velocities, Figure 12.15f shows the results of all three runs, where the left subfigure indicates the uncertainties of the minimum pressure for three noise levels, and the right subfigure shows the average errors in velocities. For the minimum pressure, the larger noise level has a larger influence on the variance of the prediction results; however, the values in general do not deviate a lot from the ground truth. The errors for velocities are less than 5% even for the heavy noise level.

The results for the Lamb-Oseen vortex are shown in Figure 12.16, and the results for the Burgers' vortex are shown in Figure 12.17. A no-noise case is also included as the control group. The same trend in the sensitivity of reconstructed minimum pressure to the noise levels can be

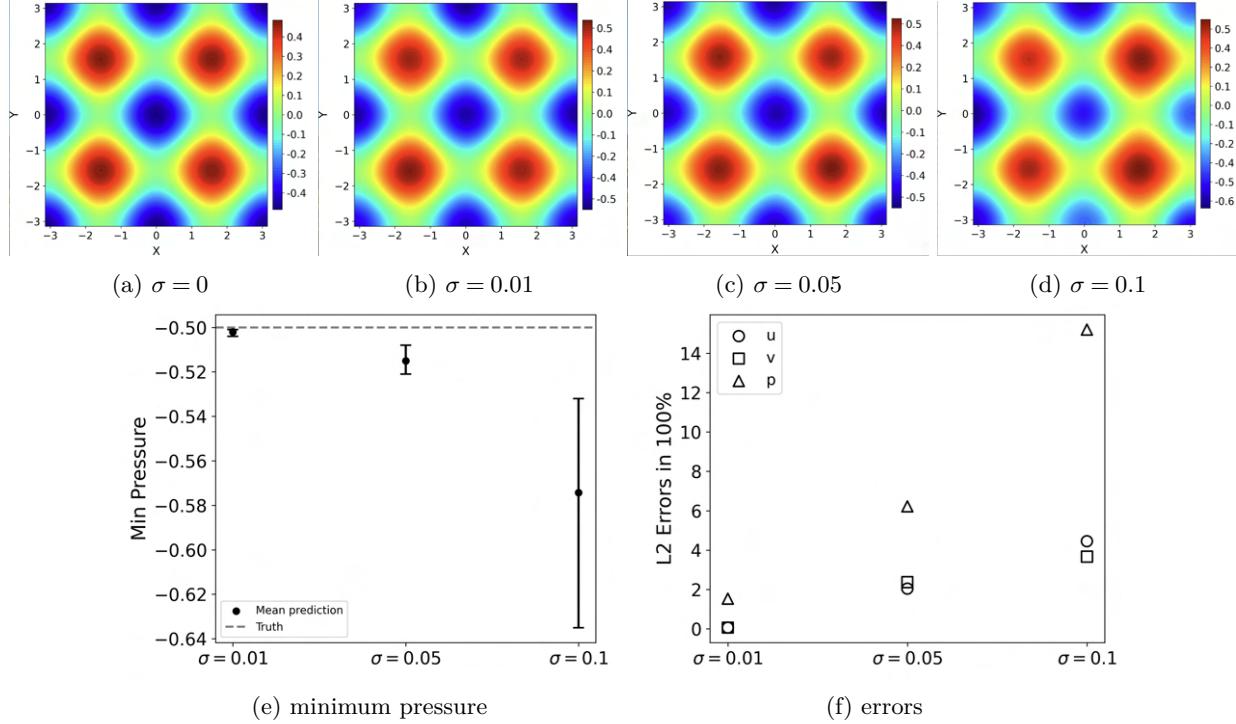


Figure 12.15: Pressure field reconstruction results for different noise levels, Taylor-Green vortex.

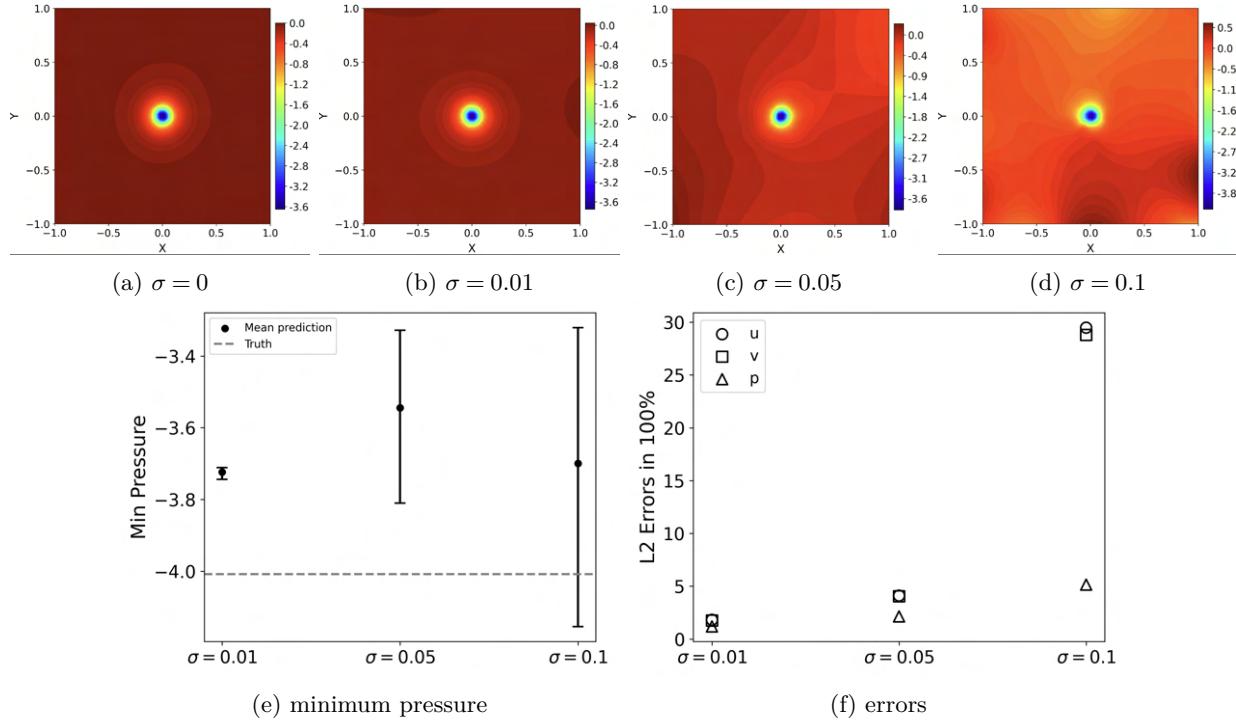


Figure 12.16: Pressure field reconstruction results for different noise levels, Lamb-Oseen vortex.

seen for both of them. When the noise level is higher, the minimum pressure has a larger variance, which could even result in a minimum pressure lower than the reference value when the noise is heavy. The errors of velocities for the Lamb-Oseen vortex, however, could be very large in heavy

noise data, although the minimum pressure is good. The pressure field for the Lamb-Oseen vortex is also well captured by the PINN model, especially the location of the minimum pressure and the general patterns. The heavy noise case also shows the noisiest pressure field in comparison to other noise levels. For the Burgers' vortex, the results are generally very good for all pressure fields, the minimum pressure, and the errors in velocities.

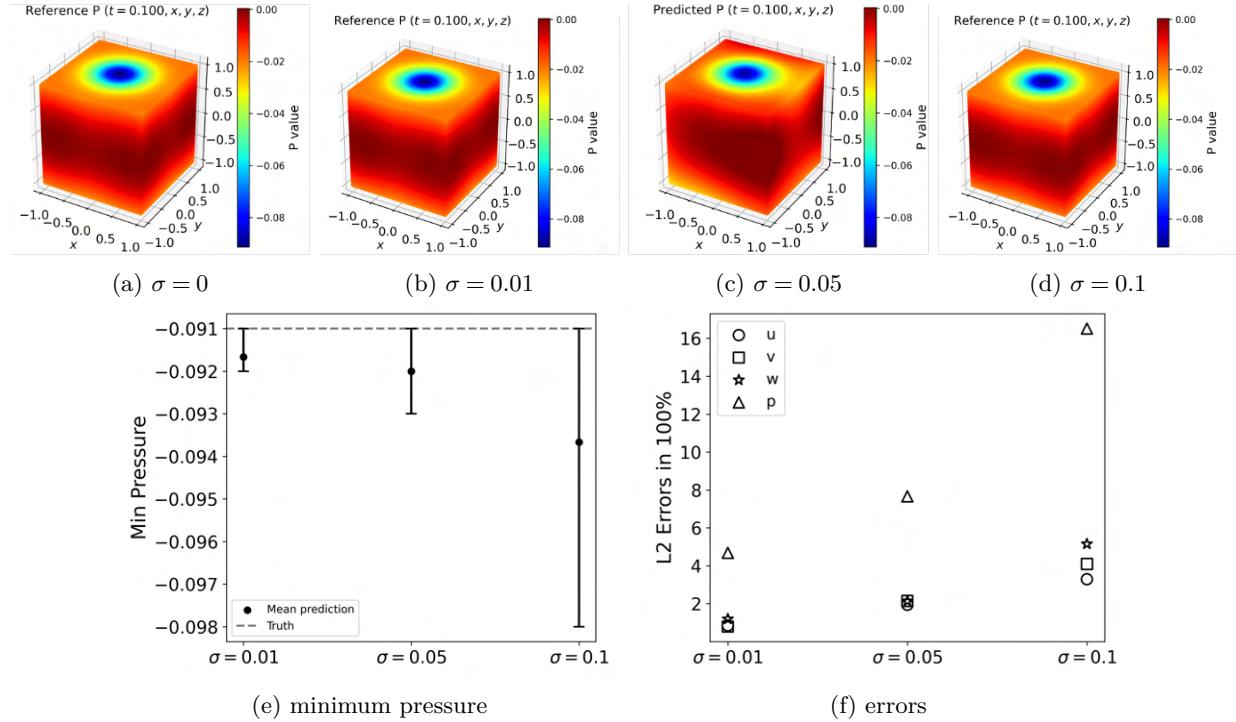


Figure 12.17: Pressure field reconstruction results for different noise levels, Burgers' vortex.

Sampling noise in spatial coordinates

For the noise effects due to the spatial coordinates noise, all three vortex models are examined, and each case was tested with 3 runs. Figure 12.18 shows two examples of the spatial noise and the particle distributions for different noise levels. When the noise level is small, the general trajectories are close to the clean control group, but the particles show large deviation when the noise is heavy.

The flow field reconstruction results are summarized in Figure 12.19. For each noise level, a total of 3 runs were taken to test the uncertainty and also avoid the randomness. The minimum pressure values at different noise levels for the three vortex models are shown in Figures 12.19a, 12.19b, and 12.19c. In general, the values of minimum pressure deviate more when the noise levels are higher, while the variations in each case do not have a clear increasing trend with the noise level like that observed in the velocity noise. For the Taylor-Green vortex, higher noise levels produce smaller minimum pressure, but the Lamb-Oseen vortex and Burgers' vortex show the opposite results: the higher noise levels give larger minimum pressure. The predictions of the Taylor-Green vortex and the Burgers' vortex are more robust among different noise levels in spatial measurements, especially for the Burgers' vortex, where the effects of the spatial noise at small and medium levels are nearly negligible. Similar to the velocity noise, higher noise levels give larger deviations of the minimum pressure from the reference truth. However, the Lamb-Oseen vortex shows poor performance even

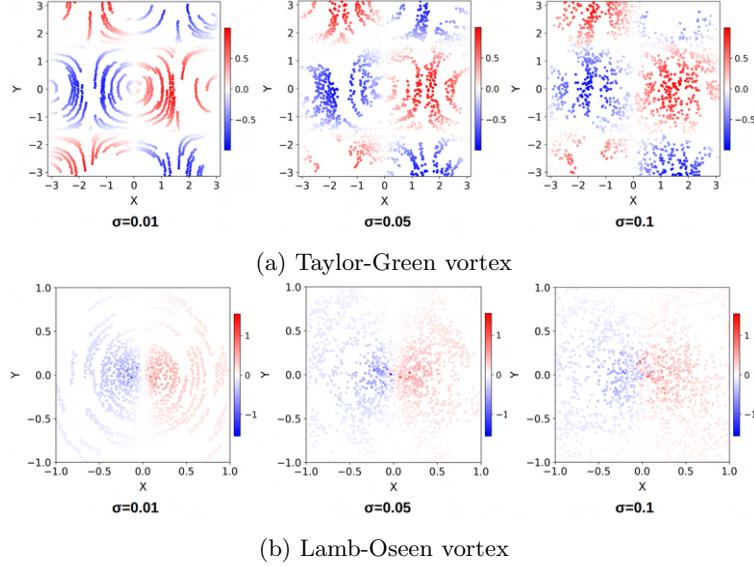


Figure 12.18: Visualization of the spatial coordinates noise at different levels.

in the medium spatial noise. The average errors (in 3 runs) of the reconstructed velocity fields and pressure fields are shown in Figures 12.19d, 12.19e, and 12.19f. In general, the errors of pressure are always higher than the errors of velocity, and the errors increase with the noise level. Lamb-Oseen vortex showed the least robustness to the spatial noise, where the errors of pressure go up to 70% for heavy noise, and 50% for medium noise. Taylor-Green vortex and Burgers' vortex are generally more robust even in the heavy noise, the errors of pressure rise to 14% for the Taylor-Green vortex and only 8% for the Burgers' vortex.

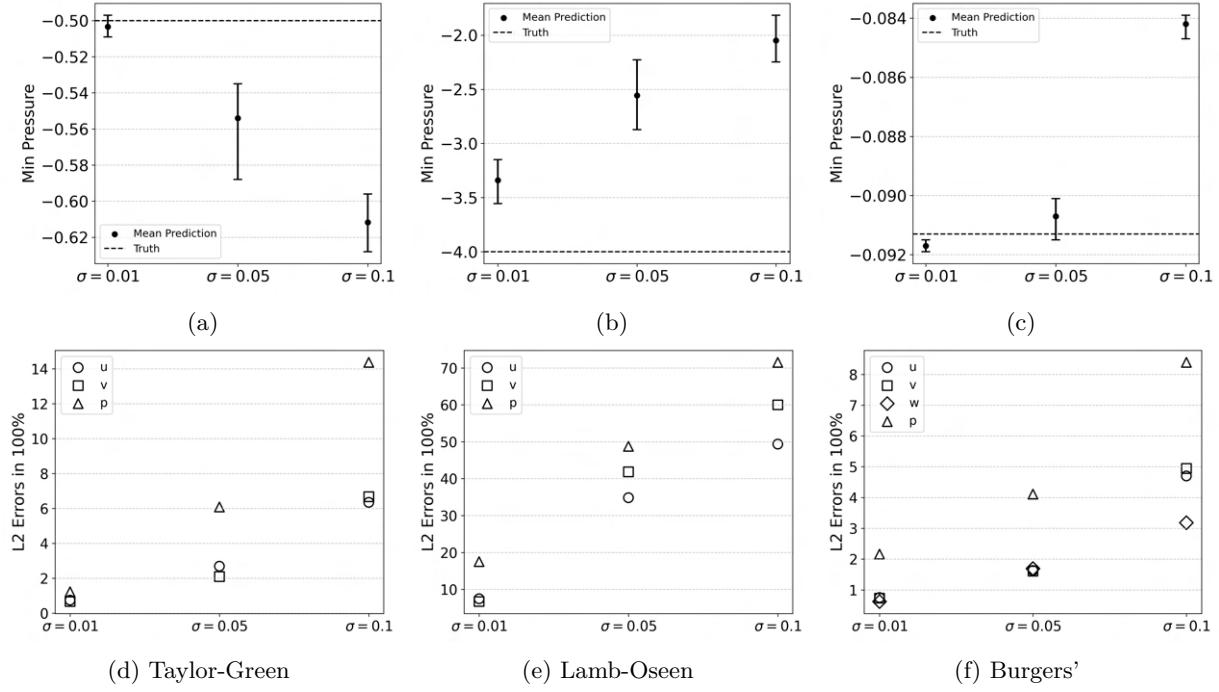


Figure 12.19: Pressure field reconstruction results for different noise levels in spatial coordinates.

12.5 Test on turbulent flow

The proposed framework is also tested using a turbulent flow database (Bappy, M.H, 2025) with the large-eddy simulation (LES) method. The problem aims to simulate a twin-vortex problem where the two vortices are interacting with each other. The data is organized as multiple 2D slices in the $y - z$ plane, from 1 to 20 at an axial distance of 0.01 in the x direction. The time step size is 0.005 time units. The kinematic viscosity is $1e-6$, and the reference velocity is $10m/s$, length of $0.167m$, density of $1000kg/m^3$, which results in a Reynolds number of 1.67×10^6 . The training data samples the velocities with a down-sample size of $1/20$ while the full-size data are used as the validation data. We aim to reconstruct the full-size flow field, assimilating the pressure field, from the sparse training data.

The results are shown in Figure 12.20 as an overview of the entire flow field reconstruction results in comparison to the ground truth, the LES simulation data. It can be seen that the overall results appear close to the LES simulation results. The entire flow fields reconstructed by the PINN model mainly capture the overall patterns of the turbulent flow and also the locations of the twin-vortex as indicated in the pressure field. To further compare the detailed results, the comparisons of the 2D slice of the flow field in the $y - z$ plane are presented in Figure 12.21. The left side of each subfigure shows the ground truth from the LES simulation, the middle column is the result predicted by the PINN model, and the right side is the sampled data for training at the corresponding slice. The pressure values in training data are set to zero to avoid any data leakage.

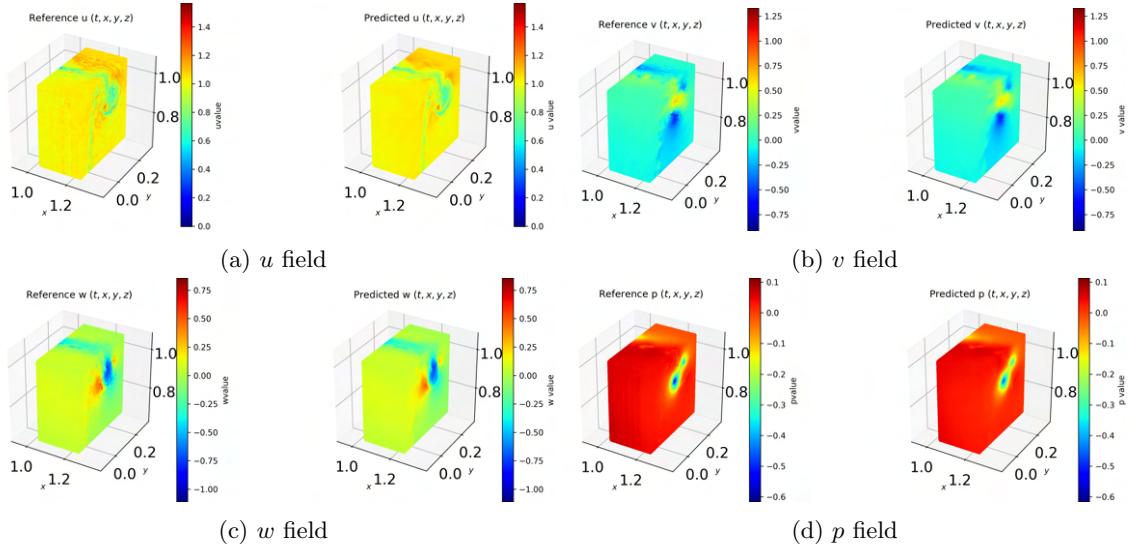


Figure 12.20: Flow field reconstruction results for twin-vortex problem.

| Error in u | Error in v | Error in w | Error in p | min p prediction/truth |
|--------------|--------------|--------------|--------------|--------------------------|
| 23.43% | 4.83% | 5.08% | 1.65% | -0.525/0.612 |

Table 12.5: Summary of the flow field reconstruction for twin-vortex simulation.

The flow fields reconstructed by the PINN model are generally smoother than the LES results, especially for the velocity field. The \mathcal{L}_2 error of the entire flow field over the entire simulation time is computed and summarized in Table 12.5. The minimum pressure obtained in the prediction is

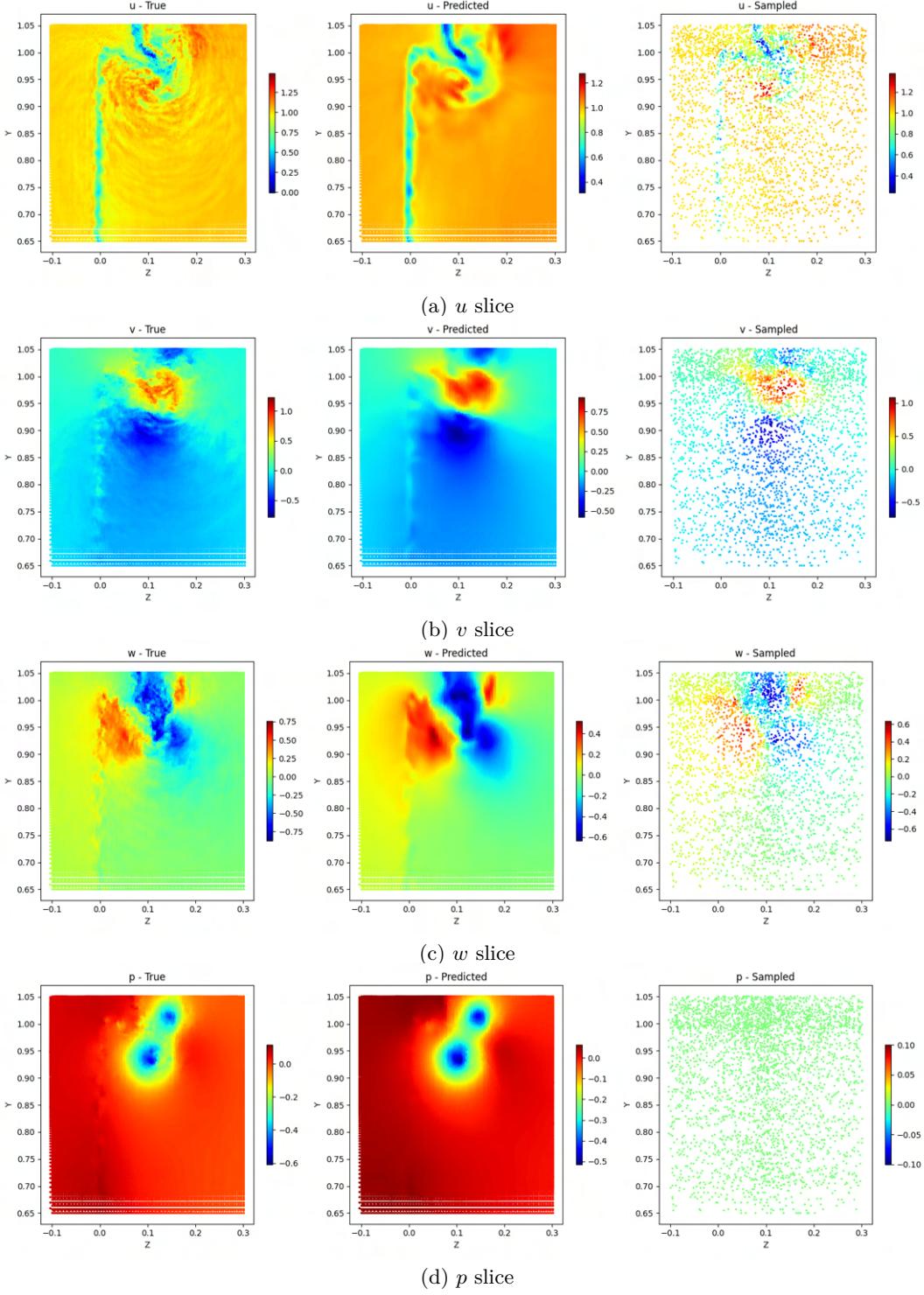


Figure 12.21: Comparison of the results in 2D slice, LES (left), PINNs (middle), sample (right).

0.525, which is slightly higher than the LES results, 0.612. Although the error in the u velocity field has a larger error, the other three fields have very small errors, especially for the pressure field with a relative error as low as 1.65%. A possible reason that the u velocity has a higher error could be that the density of the particles in the x direction is also much lower than in the other

two directions.

12.6 Conclusion and future work

This work demonstrates the capability of machine learning methods—in particular, Physics-Informed Neural Networks (PINNs)—to reconstruct fine-resolution flow fields from sparse particle measurements. Using several analytical vortex models in both 2D and 3D as test cases, the study systematically evaluated the performance of the proposed method in reconstructing both velocity and pressure fields and obtaining the location and value of the minimum pressure from limited and noisy PTV data.

The results show that the PINN model can accurately reconstruct flow fields, even with an observation rate of less than 1%. For the Taylor-Green and Burgers' vortex models, the reconstructed velocity and pressure fields closely match the reference ground truth, including accurate identification of the minimum pressure—a key parameter for understanding cavitation inception in marine applications. Although the Lamb-Oseen vortex case exhibited slightly higher errors, particularly for pressure prediction, these errors remain within an acceptable range, demonstrating the robustness of the approach across diverse flow scenarios. The investigation into sample distribution and noise effects further highlights the strengths and limitations of PINN-based flow reconstruction. Higher particle density and a uniform spatial distribution generally improved reconstruction accuracy, particularly in the pressure field. The model also exhibits notable resilience to measurement noise: the velocity fields remain accurately predicted even under significant noise, and the pressure field, especially the minimum pressure value, is predicted with acceptable uncertainty.

Overall, this study provides valuable insights and a solid foundation for further development in data-driven flow field reconstruction and pressure estimation with physics-informed machine learning tools. Future work will focus on extending the methodology to more complex and unsteady vortex dynamics, refining noise quantification and uncertainty analysis, and ultimately, application to real-world experimental PTV measurement data. This will further strengthen the bridge between sparse experimental observations and high-fidelity flow field information, with important implications for the study of cavitation inception and broader fluid dynamics applications.

References

D. A. Buchta, S. J. Laurence, and T. A. Zaki. Assimilation of wall-pressure measurements in high-speed flow over a cone. *Journal of Fluid Mechanics*, 947:R2, 2022.

G. Hasanuzzaman, H. Eivazi, S. Merbold, C. Egbers, and R. Vinuesa. Enhancement of piv measurements via physics-informed neural networks. *Measurement Science and Technology*, 34(4):044002, 2023.

D. Knister. *Experimental Study of Cavitation Inception During Vortex-Vortex Interaction and Investigation of Constrained Volumetric Velocimetry Methods*. PhD thesis, University of Michigan, Ann Arbor, 2024.

T. J. Leasca, T. B. Kroll, and K. Mahesh. Large-eddy simulation of the tip vortex flow in a ducted propulsor. *Journal of Fluid Mechanics*, 1010:A51, 2025.

P. C. D. Leoni, K. Agarwal, T. A. Zaki, C. Meneveau, and J. Katz. Reconstructing turbulent velocity and pressure fields from under-resolved noisy particle tracks using physics-informed neural networks. *Experiments in Fluids*, 64, May 2023.

X. Liu and J. Katz. Instantaneous pressure and material acceleration measurements using a four-exposure PIV system. *Experiments in Fluids*, 41(2):227–240, 2006.

References

Y. Murai, T. Nakada, T. Suzuki, and F. Yamamoto. Particle tracking velocimetry applied to estimate the pressure field around a savonius turbine. *Measurement science and technology*, 18(8):2491, 2007.

A. G. Özbay, A. Hamzehloo, S. Laizet, P. Tzirakis, G. Rizos, and B. Schuller. Poisson cnn: Convolutional neural networks for the solution of the poisson equation on a cartesian mesh. *Data-Centric Engineering*, 2:e6, 2021.

M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

E. P. Rood. Review—mechanisms of cavitation inception. *Journal of Fluids Engineering*, 113(2):163–175, 06 1991.

B. Van Oudheusden. Piv-based pressure measurement. *Measurement Science and Technology*, 24(3):032001, 2013.

M. Wang, Q. Wang, and T. A. Zaki. Discrete adjoint of fractional-step incompressible Navier-Stokes solver in curvilinear coordinates and application to data assimilation. *Journal of Computational Physics*, 396:427–450, 2019.

J. Westerweel. *Digital particle image velocimetry: Theory and application*. Delft University Press, 1995.

J. Zhang, S. Bhattacharya, and P. P. Vlachos. Using uncertainty to improve pressure field reconstruction from piv/ptv flow measurements. *Experiments in Fluids*, 61(6):131, 2020.

K. Zhou, J. Li, J. Hong, and S. J. Grauer. Stochastic particle advection velocimetry (spav): theory, simulations, and proof-of-concept experiments. *Measurement Science and Technology*, 34(6):065302, 2023.

Optimizing Hydrofoil using Gradient-Based and Deep Reinforcement Learning Approach

MOHAMED A. S. ABDUL-KAIYOOM & JOAQUIM R. R. A. MARTINS

We consider gradient-based and machine learning-based optimizations for optimizing hydrofoils. In gradient-based cases, we explore different objective functions for single-point optimizations and their effect on the optimal design. We also perform multipoint optimization to optimize the hydrofoil, which encounters different inflow angle of attack values. To perform gradient-based hydrofoil optimization, we require an efficient CFD solver with analytical derivatives. However, many multiphase solvers do not have these features. Therefore, machine-learning based optimization seems to be a promising alternative to gradient-based optimization. Therefore, we develop a deep reinforcement learning framework to perform hydrofoil optimization.

Keywords: hydrofoil optimization, deep reinforcement learning

13.1 Introduction

Hydrofoil design optimization represents a critical component in modern fluid dynamics applications, where efficiency, performance, and reliability are paramount. As industrial and research applications increasingly demand better performance in fluid-interacting designs, the need for comprehensive analysis and optimization of hydrofoil designs becomes important. Hydrofoil shape optimization is crucial in the early design stage of the design to find an optimal design that satisfies several lift, cavitation, and geometric constraints. Therefore, we perform several gradient-based hydrofoil optimizations to understand the design trends and develop a machine-learning based framework for effective design optimization.

Hydrofoil design optimization requires a computational fluid dynamics (CFD) model with appropriate fidelity that captures viscous effects, such as the Reynolds-averaged Navier-Stokes (RANS) equations. RANS-based design optimization has been applied extensively to the design of aircraft (He et al., 2019; Mangano and Martins, 2021; Abdul-Kaiyoom et al., 2023a,b), wind turbines (Dhert et al., 2017; Madsen et al., 2019), turbomachinery (Briones et al., 2020; Pátý and Lavagnoli, 2018) and watercraft (Garg et al., 2017; Ploe, 2018). Numerical optimization techniques in hydrofoil design have traditionally focused on minimizing drag while satisfying a given lift constraint and cavitation constraints. Cavitation is a major concern in the hydrodynamic designs because it can degrade the performance and cause structural failure. Therefore, eliminating cavitation is critical in the optimal design. This report presents a detailed examination of hydrofoil optimization, where we focus on optimizing the hydrofoil for different lift constraints with cavitation constraints and performance metrics that directly impact design effectiveness.

Although the above cases utilize gradient-based optimizations to efficiently find an optimal design, machine-learning has been widely used in many airfoil shape optimization (Du et al., 2021; Bouhlel et al., 2020; Dussauge et al., 2023). Machine-learning based optimizations can explore a wide range of design spaces and are more flexible in using a wide range of CFD solvers. Thus, we also develop a machine-learning based framework for effective hydrofoil design optimization

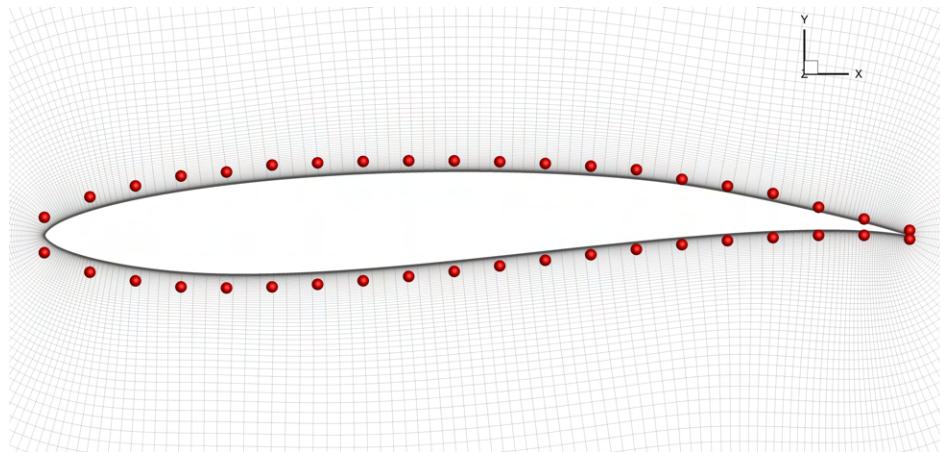


Figure 13.1: FFD parameterization for Eppler 817 hydrofoil. The red points are the FFD control points.

to incorporate CFD solvers that are not equipped with gradients. This framework will facilitate effective design optimization with multiphase solvers in the future. In this framework, we use a deep reinforcement learning (DRL) based approach to optimize hydrofoil designs. In DRL, the agent identifies the best course of action to take with the environment and maximizes the cumulative reward function. Using this approach, Dussauge et al. (2023) demonstrated that the DRL method reveals increased efficiency in exploring the design space and successfully converges on airfoils that exhibit improved performance compared to the traditional approach. Thus, DRL based hydrofoil optimization is a potential approach for future design optimization.

The remainder of the report is structured as follows. Section 13.2 explains the methodology used for addressing the hydrodynamic design optimization. Section 13.3 details the optimization problems solved in this study. In Sec. 13.4, we present and discuss the results of the traditional optimization problem and the features of the ongoing DRL framework. Finally, the conclusions are presented in Sec. 13.5.

13.2 Methodology

For traditional and DRL based optimizations, we use the OpenMDAO framework (Gray et al., 2019), which enables a flexible component-based architecture. In these optimizations, the primary CFD solver is ADflow¹ (Mader et al., 2020), which solves the RANS equations on structured multiblock and overset grids. We use the Spalart–Allmaras (SA) turbulence model (Spalart and Allmaras, 1992) in all the results. ADflow also uses a robust approximate Newton–Krylov solver algorithm (Yildirim et al., 2019) and an efficient adjoint solver to compute the derivatives required for gradient-based optimization (Kenway et al., 2019). Besides ADflow, pyHyp² (Secco et al., 2021), a grid generation package, is used for generating volume meshes. pyGeo³ (Kenway et al., 2010) updates the geometry model and IDWarp⁴ (Secco et al., 2021) warps the volume mesh. pyGeo offers several parameterization techniques. In this work, the airfoil shape is parameterized using a free-form deformation (FFD) approach (Sederberg and Parry, 1986), as shown in Figure 13.1.

¹<https://github.com/mdlolab/adflow>

²<https://github.com/mdlolab/pyhyp>

³<https://github.com/mdlolab/pygeo>

⁴<https://github.com/mdlolab/idwarp>

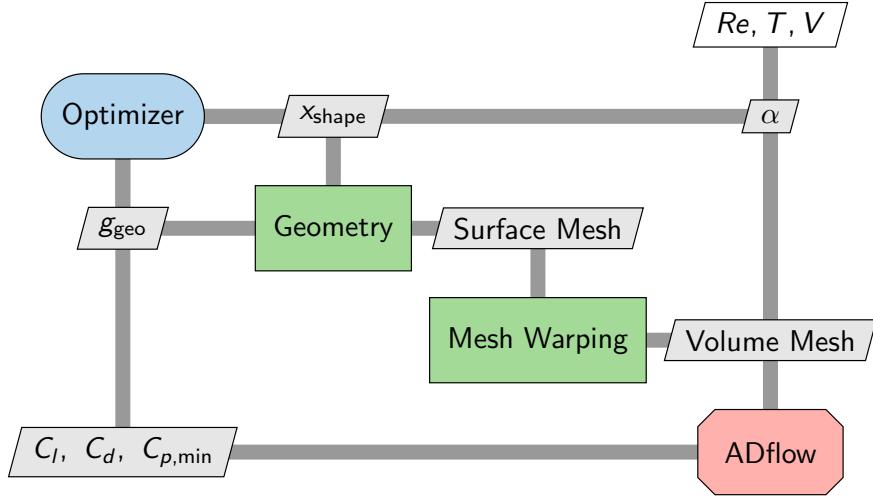


Figure 13.2: XDSM diagram for the hydrofoil optimization.

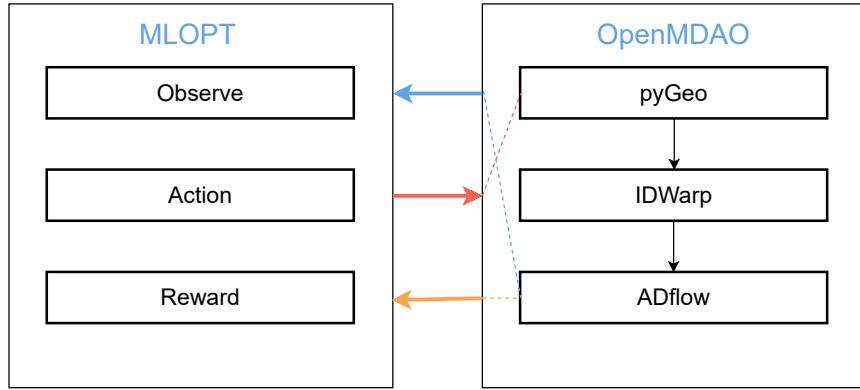


Figure 13.3: DRL framework that is coupled with OpenMDAO components.

For gradient-based optimization, within pyOptSparse, we use the sparse nonlinear optimizer (SNOPT) (Gill et al., 2002), a sequential quadratic programming (SQP) algorithm. SNOPT can solve large-scale nonlinear optimization problems. The finite difference method is the most common method for calculating derivatives for gradient-based calculations. However, finite differences scale poorly with the number of design variables and are subject to significant numerical errors (Martins and Ning, 2022, Sec. 6.4). The adjoint method is the most cost-effective approach in this case because its computational cost does not depend on the number of design variables (Martins and Ning, 2022, Sec. 6.7).

Figure 13.2 shows the eXtended Design Structure Matrix (XDSM) diagram of the hydrofoil optimization. In the gradient-based optimization, the optimizer feeds pyGeo the new design variables at each iteration, updating the geometry model and related CFD surface mesh. IDWarp then updates the volume mesh through the volume mesh warping module. Then, the ADflow is converged using the robust approximate Newton–Krylov solver algorithm (Yildirim et al., 2019). ADflow also computes the adjoint derivatives of the modified mesh. The function and derivative evaluations are passed into the optimizer for the next update.

For DRL based optimization, we develop a DRL framework, which uses the *Stable Baselines*

package in Python and is coupled with the OpenMDAO framework to observe, modify, and calculate the reward functions. Figure 13.3 shows this DRL framework, which has been developed recently. The DRL agent observes the current FFD control point positions and takes an action to modify the shape by moving the FFD control points. Then, we pass these new FFD control point positions to pyGeo to update the geometry and update the volume mesh using IDWarp. Following that, we run aerodynamic analysis using the ADflow on the updated volume mesh and collect the performance metrics such as lift coefficient (C_l), drag coefficient (C_d), and minimum pressure coefficient (C_p, \min). These performance metrics are used to compute the reward function, which gives an agent numerical feedback and guides it towards learning beneficial behaviors by gradually increasing the total reward value. In deep reinforcement learning, the reward function is a mathematical function that gives an agent scalar feedback regarding the quality of its actions. It is the main learning feedback that directs the agent to perform desired actions. In this research, we provide a positive value when the constraints are met and a negative value when the constraints are violated. This iterative process continues to maximize the cumulative reward function.

13.3 Problem formulation

In general, we minimize C_d with respect to the angle of attack and shape design variables while satisfying a set of constraints. The geometric constraint functions involve the hydrofoil's thickness (t), curvature (k), and LE radius (R).

| | Function/Variable | Description | Quantity |
|------------|-----------------------------|------------------------|----------|
| Minimize | C_d | | |
| Or | | | |
| Maximize | C_l/C_d | | |
| By varying | x_{shape} | Shape design variables | 20 |
| | α | Angle of attack | 1 |
| Subject to | $C_l = 0.2$ | Lift coefficient | 1 |
| | $0.8t_0 \leq t \leq 1.5t_0$ | Thickness constraint | 200 |
| | $0.8k_0 \leq k \leq 4.0k_0$ | Curvature constraint | 2 |
| | $0.8R_0 \leq R \leq 2.5R_0$ | LE radius constraint | 1 |
| | $C_p, \min \geq -0.5$ | cavitation constraint | 1 |

Table 13.1: The single-point hydrofoil optimization problem formulation.

We use three different baseline designs. The first baseline is the Eppler 817 hydrofoil because several studies have investigated its cavitation development (Astolfi et al., 1999; Amromin, 2001). Therefore, we perform hydrodynamic shape optimization to optimize the baseline shape by removing the cavitation. For removing the cavitation, we use a cavitation model in ADflow, where we find the minimum C_p over the hydrofoil surface and constrain it. This constraint is computed by applying KS aggregation (Kreisselmeier and Steinhauser, 1979) to the C_p values computed at each location, which gives a smooth and differentiable approximation of the minimum C_p ,

$$C_{p,\min} = \text{KS}(-C_p) = \max(-C_p) + \frac{1}{\rho_{\text{KS}}} \ln \left(\sum_{n=1}^{N-1} e^{\rho_{\text{KS}}(-C_p - \max(-C_p))} \right), \quad (13.1)$$

where $C_{p,i}$ are the coefficient of pressure values over the surface, N is the number of C_p over the surface, and ρ is KS weight parameter. This cavitation model has been used successfully to alleviate cavitation in several cases (Liao et al., 2021; Ng et al., 2025).

Table 13.1 shows the optimization problem formulation for the single-point optimization cases. In the first case, we minimize the C_d subject to the constraints. In the second case, to achieve a maximum hydrodynamic efficient foil, we also perform an optimization to maximize L/D with the same set of constraints. We compare these two cases in Section 13.4.

The second baseline is NACA 66 for propeller designs because it has been used in the KP 505 benchmark propeller. We are interested in studying the multipoint optimization for this baseline because the propeller encounters different inflow angles of attack. Table 13.2 shows the problem formulation for this case. Therefore, we formulate a multipoint optimization to minimize the drag at different angles of attack values. We consider three points for this problem. The center point has an angle of attack (α) design variable, and the other two points have an angle of attack of $\alpha - 1$ degrees and $\alpha + 1$ degrees, respectively. In this formulation, the optimizer explores various inflow angles for the optimal design and achieves a better performance. We also enforce cavitation constraint on all three points using Eq. 13.1.

| | Function/Variable | Description | Quantity |
|------------|---|--|----------|
| Minimize | $C_{d, \alpha-1} + C_d + C_{d, \alpha+1}$ | | |
| By varying | x_{shape} | Shape design variables | 20 |
| | α | Angle of attack | 1 |
| Subject to | $C_l = 0.2$ | Center point Lift coefficient | 1 |
| | $C_{p, \min} \geq -0.5$ | Cavitation constraint for all three points | 1 |
| | $0.8t_0 \leq t \leq 1.5t_0$ | Thickness constraint | 200 |
| | $0.8k_0 \leq k \leq 4.0k_0$ | Curvature constraint | 2 |
| | $0.8R_0 \leq R \leq 2.5R_0$ | LE radius constraint | 1 |

Table 13.2: The multipoint hydrofoil optimization problem formulation.

The third baseline is the RAE 2822 airfoil. We use this baseline because it has been used in several airfoil optimizations (Du et al., 2021; He et al., 2019). We are interested in testing our DRL framework with the well-known baseline design. We will first perform transonic airfoil shape optimization using our DRL framework to validate the results with the gradient-based optimization. This will guarantee that the DLR-based optimization is comparable with the gradient-based optimization. Table 13.3 shows the problem formulation for this case.

The Reynolds number for the single-point and multipoint hydrofoil cases is 2×10^{-6} , and the Mach number is 0.18. In the airfoil optimization, the Mach number is 0.8 and the altitude is 10,000 meters.

| | Function/Variable | Description | Quantity |
|------------|------------------------------|------------------------|----------|
| Minimize | C_d | | |
| By varying | x_{shape} | Shape design variables | 20 |
| | α | Angle of attack | 1 |
| Subject to | $C_l = 0.2$ | Lift coefficient | 1 |
| | $0.8t_0 \leq t \leq 1.5t_0$ | Thickness constraint | 200 |
| | $0.25k_0 \leq k \leq 3.5k_0$ | Curvature constraint | 2 |

Table 13.3: The single-point airfoil optimization problem formulation.

13.4 Results

Gradient-based hydrofoil optimization

We first perform several single-point optimizations in this section to optimize the hydrofoil for the given problem formulation in Table 13.1. For the drag-minimized problem, we chose five different target lift values. For this case, the maximum target lift value we were able to optimize was 0.6. For the C_l/C_d maximized problem, we remove the lift constraint. The rest of the constraints in this case are identical to the drag-minimized problem.

Figure 13.4 shows the optimal hydrofoil shapes for the problem formulation in Table 13.1. When the target lift value increases in the drag-minimized problem, the optimal design achieves higher camber. On the upper surface of the hydrofoil, the pressure also decreases. Thanks to the cavitation constraint, the optimizer limits the optimal design to fall below the constrained C_p value. We notice a flattening pressure value over the upper surface due to the cavitation constraint. As the camber increases due to the higher target lift value, more of the optimal design upper surface has a constant pressure value. Similarly, the C_l/C_d maximized case also achieves a high cambered hydrofoil design. This design also has a constant pressure value over the upper surface because of the cavitation constraint.

Figure 13.5 shows the C_l vs C_d plot for all the cases in Table 13.1. When the target lift value increases, the optimal drag value also increases. The C_l/C_d maximized have the largest C_l and C_d value. Figure 13.6 shows the polar study for the drag-minimized and C_l/C_d maximized cases. Each optimal point has lower drag in its polar plot. These results show the important trends of optimizing the hydrofoil using two different ways: minimizing drag and maximizing C_l/C_d . Both of these cases satisfy the constraint and achieve different optimal designs.

Following that, we perform a multipoint optimization for the problem detailed in Table 13.2. The ship propellers encounter different inflow angles of attack because of the wake from the ship hull. Therefore, we consider these variations in the optimization problem to optimize the baseline design. We also perform a single-point optimization to show the importance of this multipoint optimization.

Figure 13.7 shows the multipoint optimal design and C_p plots. We also show the single-point optimal design for comparison. Both of these cases achieve different optimal designs. The optimal α for single-point and multipoint optimizations are 0.42 and 0.63, respectively. The optimizer changes the multipoint optimal design significantly to reduce the drag across all three points and satisfy the constraints. The single-point optimal design violates the cavitation constraint at the $\alpha + 1$ point because this point was not considered in the problem formulation. This performance degradation

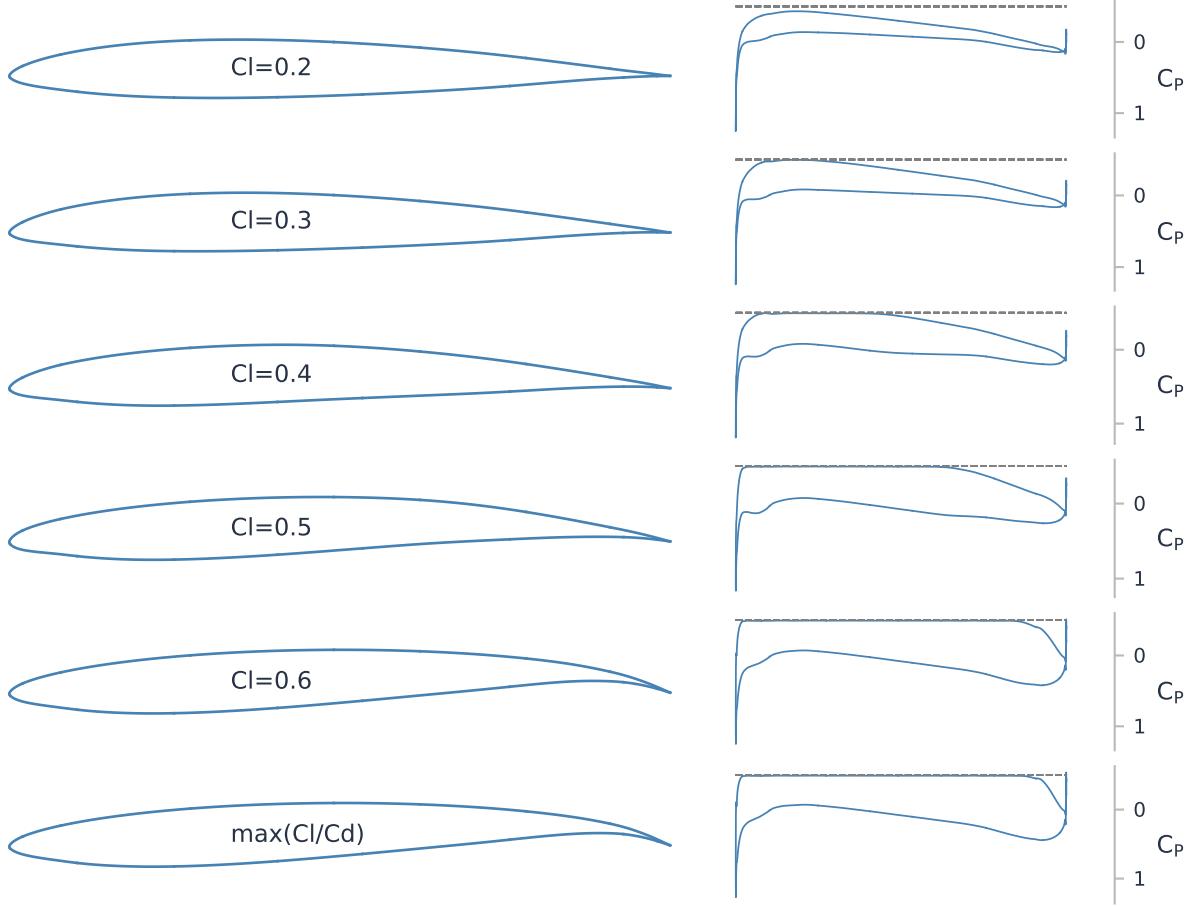


Figure 13.4: Optimal shapes for the drag minimized and C_l/C_d maximized problem.

is known in single-point optimizations. However, the multipoint optimal design satisfies all the constraints, including the cavitation constraint across all the points.

Figure 13.8 shows the α vs C_l and α vs C_d plots. For a given α , the multipoint design achieves a larger C_l compared to the single-point design because it has a higher camber. However, it achieves a larger drag compared to the single-point optimal design. Nevertheless, the multipoint optimal design satisfies the cavitation by penalizing the drag. Therefore, multipoint optimizations are crucial for designs that experience different inflow angles of attack values.

DRL-based hydrofoil optimization

The policy network in the DRL algorithm needs to observe the environment and take actions to optimize the design. As explained in Sec. 13.2, the environment in the DRL framework is the OpenMDAO group, which contains pyGeo, IDWarp, and ADflow components. Reinforcement learning is distinct from other machine learning techniques. Instead of using a fixed dataset as in supervised learning, the agent in reinforcement learning itself gathers the data for training through interactions with the environment. We want to optimize the airfoil problem detailed in Table 13.3. Therefore, the DRL agent is first trained using the environment to map from states, which are design variables, lift coefficient, drag coefficient, and geometric constraints, to actions, where the agent updates the design variables. This mapping process is done internally in the DRL agent

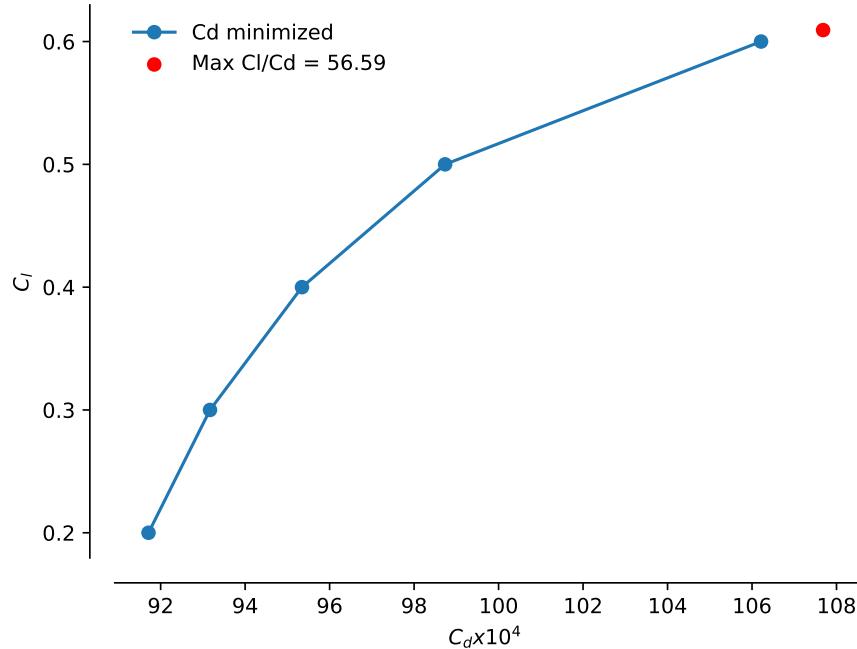


Figure 13.5: C_l vs C_d plot for the problem in Table 13.1.

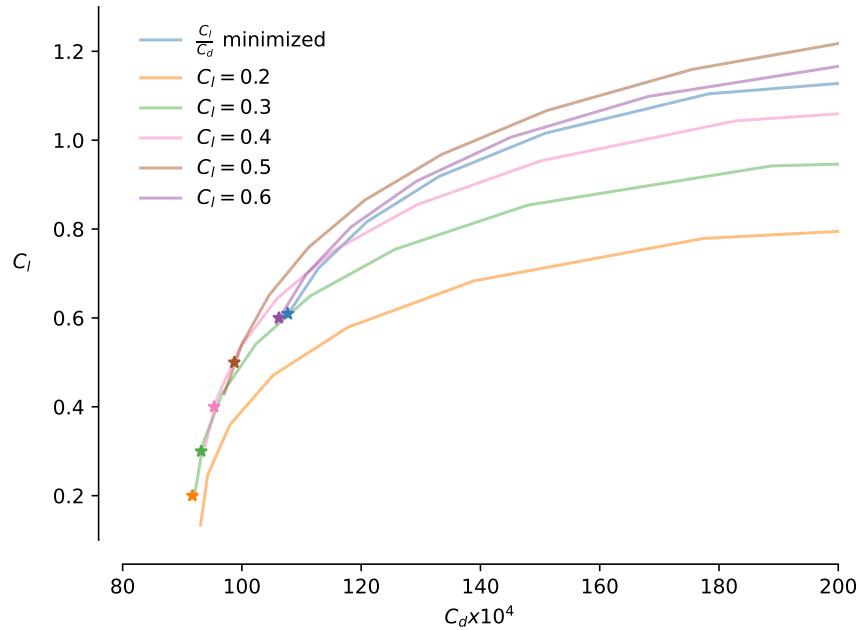


Figure 13.6: Drag polar for the problem in Table 13.1.

to take actions. In this way, the DRL agent interacts with the environment and learn about the quality of its actions. The DRL reward function reflects the quality of action taken by the agent. The reward function contains the objective function (drag coefficient) and the constraints (lift and geometric constraints). When the drag decreases, the reward function value increases. Similarly, when the constraints are met, the reward function value increases as described in Sec. 13.2. In this way, the DRL agent finds the optimal design. During the training process, the agent observes

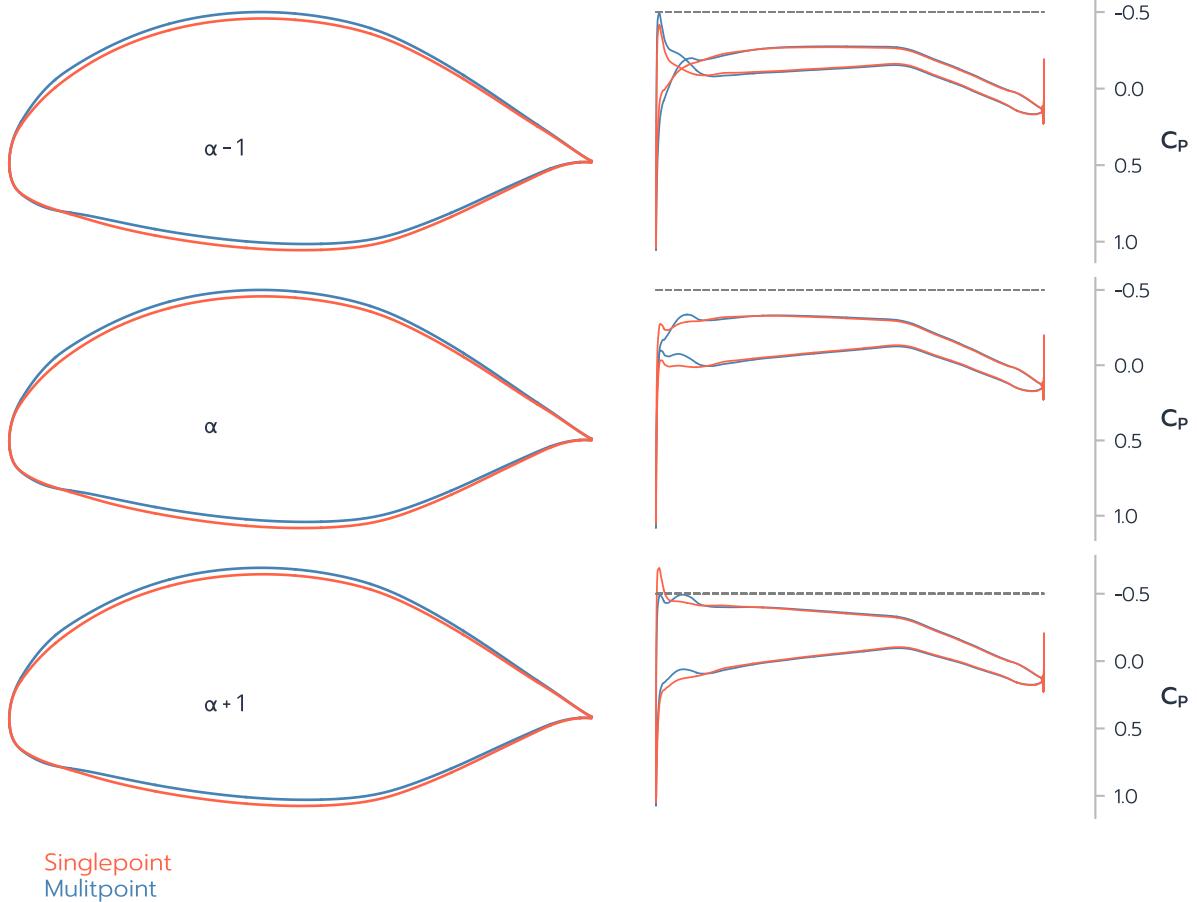


Figure 13.7: The single-point and multipoint optimal designs.

and takes action to learn the design space by updating the angle of attack and FFD control point positions. The updated hydrofoil shape for the given action is required to be a valid shape to run the CFD simulation to get the performance metrics. The CFD simulation frequently fails because these shapes have negative volume cells and are therefore invalid. To overcome these difficulties, we enforce thickness and curvature constraints at each iteration. At the beginning of training, we first compute the initial thickness (t_0) using pyGeo for enforcing the relative thickness at each iteration (left figure in Figure 13.9). The thickness corrections are made in the following steps:

1. At each iteration, we compute the thickness t_i (right figure in Figure 13.9) to check if the relative thickness t_i/t_0 violates its bounds.
2. If the relative thickness violates the bounds, we move the upper FFD (FFD_{upper}) and lower FFD (FFD_{lower}) control point positions by the required amount to satisfy the constraints,

if $t_i/t_0 < t_{lower}$,
 $FFD_{upper}+ = |FFD_{upper}| \times (t_{lower} - t_i/t_0), \quad FFD_{lower}- = |FFD_{lower}| \times (t_{lower} - t_i/t_0);$
 if $t_i/t_0 > t_{upper}$,
 $FFD_{upper}+ = |FFD_{upper}| \times (t_{upper} - t_i/t_0), \quad FFD_{lower}- = |FFD_{lower}| \times (t_{upper} - t_i/t_0),$

where t_{lower} and t_{upper} are lower and upper bounds for the relative thickness, respectively.

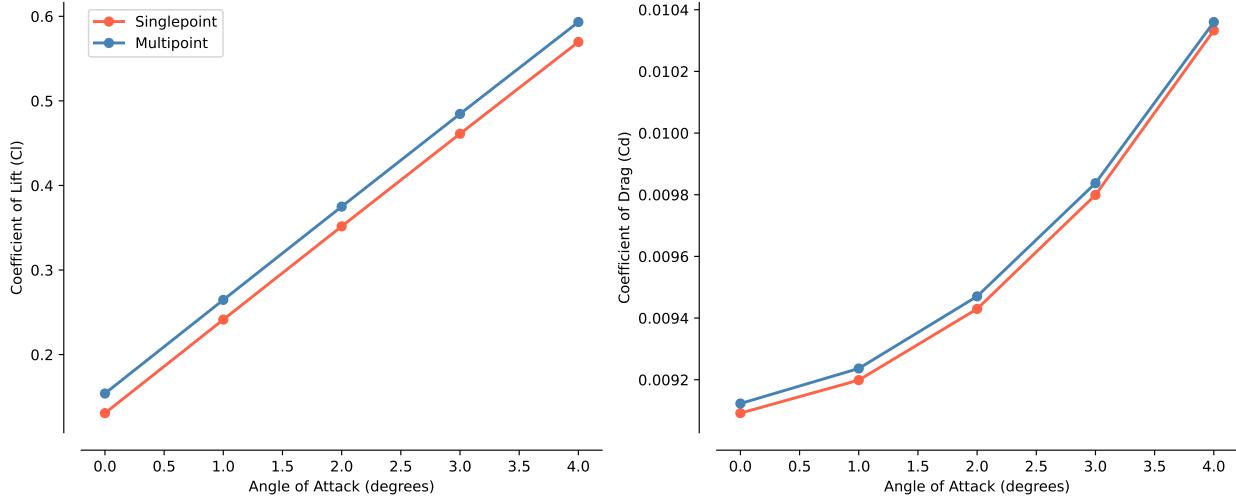


Figure 13.8: Polar study for the optimal designs.

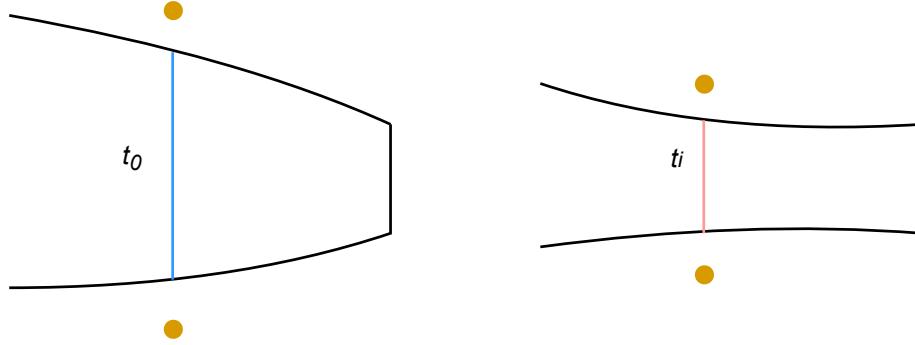


Figure 13.9: Enforcing thickness and curvature constraints at each iteration. The orange points are the FFD control points. The airfoil shape change is shown in this figure. The blue line shows the initial thickness and the red line shows the thickness at the i^{th} iteration.

3. Following that, we check the curvature of the FFD control point positions. High curvature FFD control point variation leads to relatively less smooth shapes that are difficult to converge. Figure 13.10 shows how the high curvature point is corrected to get a relatively smooth curve. We remove the high curvature FFD control point and interpolate with the rest of the FFD control point positions to find the new FFD control point position for the high curvature FFD control point. We use the *PCHIP* interpolate function in python, monotonic cubic splines, to find the value of new points to eliminate high curvature points.
4. We then repeat step 3 to enforce the thickness constraint again to satisfy them. Figure 13.11 shows how the bad shape is turned into a valid shape by enforcing thickness and curvature constraints.
5. Finally, we update the volume mesh and run the CFD simulation.

We calculate how much the FFD control point positions need to be changed based on the required thickness and curvature changes to satisfy their bounds. After we update the new FFD control point positions, IDWarp warps the volume mesh. Following that, ADflow runs the CFD

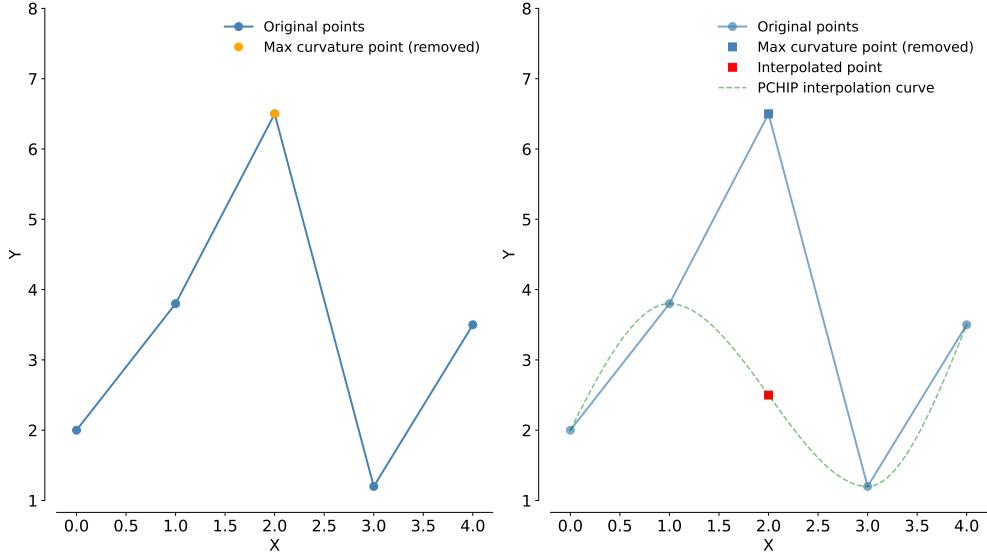


Figure 13.10: Removing high curvature point and interpolating to find new points.

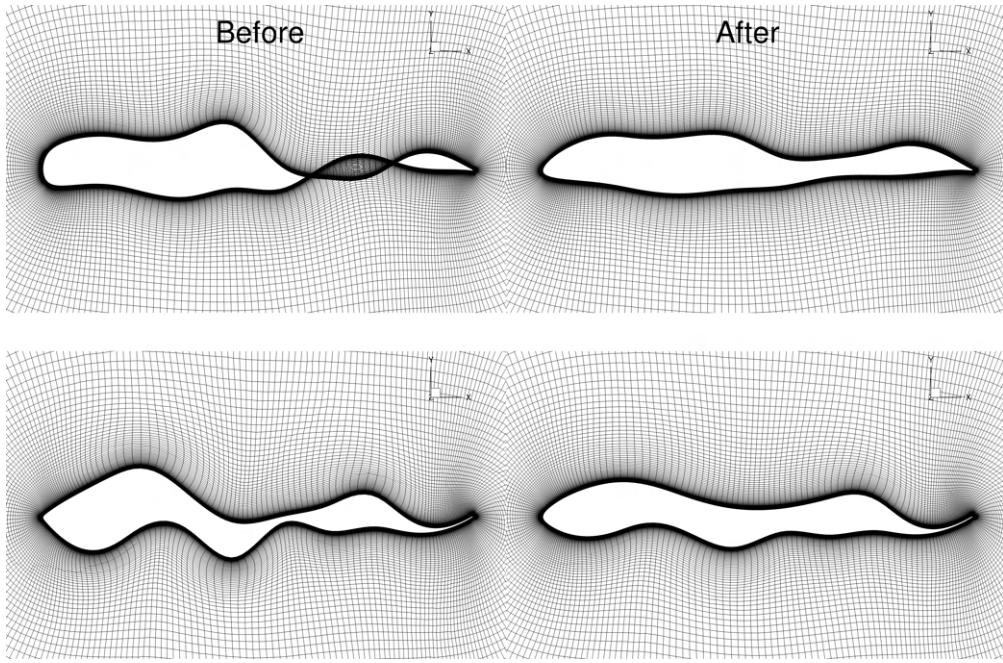


Figure 13.11: Before and after enforcing thickness and curvature constraints at each iteration.

analysis and provides the performance metrics. This training process is carried out for a number of iterations. The policy network in DRL learns from the trained dataset. After training, the DRL agent will optimize the baseline. However, at this stage, we are exploring more robust options to train the environment and get practical shapes. After training, we will explore the optimization part. In the next process, we will explore the options to tune hyperparameters for optimal training and performance.

13.5 Conclusions

Hydrofoil optimizations are crucial for naval designs. As a result, we perform several gradient-based hydrofoil optimizations to understand the design trends and develop a machine-learning based framework for effective design optimization.

In the gradient-based optimizations, we show the optimal design differences between minimizing the drag and maximizing the lift over drag ratio. The latter case achieves higher drag and lift values to maximize the lift-to-drag ratio. Both of these cases satisfy the cavitation constraints, where we noticed a flattening pressure value over the upper surface. Furthermore, we also perform multipoint optimization to effectively optimize the different inflow angles of the attack case. We also perform a single-point optimization to compare it to the multipoint optimization. The multipoint case successfully satisfies the cavitation constraint across all the points, while the single-point case violates the cavitation constraint at the high inflow angle of attack point. These results show the importance of the multipoint optimization in hydrofoil optimization.

Finally, we develop a deep reinforcement learning (DRL) based framework for hydrofoil optimization. Currently, we are developing geometric constraint functions to obtain valid shapes for training. We are also exploring tuning the hyperparameter for optimal performance.

References

M. A. S. Abdul-Kaiyoom, A. Yildirim, and J. R. R. A. Martins. Coupled aeropropulsive design optimization of an over-wing nacelle configuration. In *AIAA SciTech Forum*, 2023a. AIAA 2023-0327.

M. A. S. Abdul-Kaiyoom, A. Yildirim, and J. R. R. A. Martins. RANS-based multipoint aeropropulsive design optimization of an over-wing nacelle configuration. In *AIAA Aviation Forum*, San Diego, CA, 2023b. AIAA 2023-3588.

E. L. Amromin. Scale effect of cavitation inception on a 2d Eppler hydrofoil. *Journal of Fluids Engineering*, 124(1):186–193, 2001.

J.-A. Astolfi, P. Dorange, J.-Y. Billard, and I. C. Tomas. An experimental investigation of cavitation inception and development on a two-dimensional Eppler hydrofoil. *Journal of Fluids Engineering*, 122(1):164–173, 1999.

M. A. Bouhlel, S. He, and J. R. R. A. Martins. Scalable gradient-enhanced artificial neural networks for airfoil shape design in the subsonic and transonic regimes. *Structural and Multidisciplinary Optimization*, 61:1363–1376, 2020.

A. M. Briones, T. J. Erdmann, and B. A. Rankin. LES-verified RANS-based deterministic and continuous multiple-objective combustor design optimization. In *Proceedings of the ASME Turbo Expo*, volume Volume 4A: Combustion, Fuels, and Emissions of *Turbo Expo: Power for Land, Sea, and Air*, 2020.

T. Dhert, T. Ashuri, and J. R. R. A. Martins. Aerodynamic shape optimization of wind turbine blades using a Reynolds-averaged Navier–Stokes model and an adjoint method. *Wind Energy*, 20(5):909–926, 2017.

X. Du, P. He, and J. R. R. A. Martins. Rapid airfoil design optimization via neural networks-based parameterization and surrogate modeling. *Aerospace Science and Technology*, 113:106701, 2021.

T. P. Dussauge, W. J. Sung, O. J. P. Fischer, and D. N. Mavris. A reinforcement learning approach to airfoil shape optimization. *Scientific Reports*, 2023.

N. Garg, G. K. W. Kenway, J. R. R. A. Martins, and Y. L. Young. High-fidelity multipoint hydrostructural optimization of a 3-D hydrofoil. *Journal of Fluids and Structures*, 71:15–39, 2017.

P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM Journal of Optimization*, 12(4):979–1006, 2002.

References

J. S. Gray, J. T. Hwang, J. R. R. A. Martins, K. T. Moore, and B. A. Naylor. OpenMDAO: An open-source framework for multidisciplinary design, analysis, and optimization. *Structural and Multidisciplinary Optimization*, 59(4):1075–1104, 2019.

X. He, J. Li, C. A. Mader, A. Yildirim, and J. R. R. A. Martins. Robust aerodynamic shape optimization—from a circle to an airfoil. *Aerospace Science and Technology*, 87:48–61, 2019.

G. K. W. Kenway, G. J. Kennedy, and J. R. R. A. Martins. A CAD-free approach to high-fidelity aerostructural optimization. In *Proceedings of the 13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference*, Fort Worth, TX, 2010. AIAA 2010-9231.

G. K. W. Kenway, C. A. Mader, P. He, and J. R. R. A. Martins. Effective adjoint approaches for computational fluid dynamics. *Progress in Aerospace Sciences*, 110:100542, 2019.

G. Kreisselmeier and R. Steinhauser. Systematic control design by optimizing a vector performance index. In *International Federation of Active Controls Symposium on Computer-Aided Design of Control Systems, Zurich, Switzerland*, volume 12, pages 113–117, 1979.

Y. Liao, J. R. R. A. Martins, and Y. L. Young. 3-D high-fidelity hydrostructural optimization of cavitation-free composite lifting surfaces. *Composite Structures*, 268:113937, 2021.

C. A. Mader, G. K. W. Kenway, A. Yildirim, and J. R. R. A. Martins. ADflow: An open-source computational fluid dynamics solver for aerodynamic and multidisciplinary optimization. *Journal of Aerospace Information Systems*, 17(9):508–527, 2020.

M. H. A. Madsen, F. Zahle, N. N. Sørensen, and J. R. R. A. Martins. Multipoint high-fidelity CFD-based aerodynamic shape optimization of a 10 MW wind turbine. *Wind Energy Sciences*, 4:163–192, 2019.

M. Mangano and J. R. R. A. Martins. Multipoint aerodynamic shape optimization for subsonic and supersonic regimes. *Journal of Aircraft*, 58(3):650–662, 2021.

J. R. R. A. Martins and A. Ning. *Engineering Design Optimization*. Cambridge University Press, Cambridge, UK, 2022.

G. W. Ng, Y. Liao, A. Yildirim, and J. R. R. A. Martins. Hydrostructural optimization of subcavitating cambered and symmetric composite foils. *Composite Structures*, 351:118545, 2025.

M. Pátý and S. Lavagnoli. Accuracy of RANS CFD methods for design optimization of turbine blade tip geometries. In *2018 AIAA Joint Propulsion Conference*, 2018. AIAA 2018-4435.

P. Ploe. *Surrogate-based optimization of hydrofoil shapes using RANS simulations*. Theses, École centrale de Nantes, 2018.

N. Secco, G. K. W. Kenway, P. He, C. A. Mader, and J. R. R. A. Martins. Efficient mesh generation and deformation for aerodynamic shape optimization. *AIAA Journal*, 59(4):1151–1168, 2021.

T. W. Sederberg and S. R. Parry. Free-form deformation of solid geometric models. *SIGGRAPH Computer Graphics*, 20(4):151–160, 1986.

P. Spalart and S. Allmaras. A one-equation turbulence model for aerodynamic flows. In *30th Aerospace Sciences Meeting and Exhibit*, 1992. AIAA 1992-439.

A. Yildirim, G. K. W. Kenway, C. A. Mader, and J. R. R. A. Martins. A Jacobian-free approximate Newton–Krylov startup strategy for RANS simulations. *Journal of Computational Physics*, 397:108741, 2019.

Adjoint-Based Assimilation of External Data with High-Fidelity Simulations of Complex Flows

GAO JUN WU, SREEVATSA ANANTHARAMU¹ & KRISHNAN MAHESH

This work introduces an adjoint-based data assimilation algorithm that integrates high-fidelity flow simulations with external velocity data. The external data sources include theoretical solutions, numerical simulations from other flow solvers and experimental measurements. The algorithm finds a numerical solution to the incompressible Navier-Stokes equations that best matches the external velocity data. It distinguishes itself from other adjoint-based assimilation methods in two ways. Firstly, the modeled solution is obtained using a high-order, non-dissipative Hybridizable Discontinuous Galerkin (HDG) scheme on unstructured meshes (Anantharamu and Mahesh, 2024). Secondly, beyond optimizing the unknown initial condition of the incompressible Navier-Stokes (N-S) equations, the assimilation also optimizes unknown time-varying inflow/outflow boundary conditions. The objective function is defined as the sum of squared differences between the numerical solutions and external data, with the discretized N-S equations imposed as equality constraints through Lagrange multipliers. Using the adjoint of the discrete system, the optimization problem is solved iteratively with the quasi-Newton method. The algorithm has been implemented both in MATLAB for small test cases and in parallel for large flow problems on high-performance computing (HPC) systems. Various canonical test cases are presented to highlight the key features of the algorithm, followed by a discussion of the computational cost.

Keywords: data assimilation, HDG, adjoint

14.1 Introduction

Experiments, analytical and numerical solutions of fluid dynamic systems are typically viewed as distinct data sources that need to be combined for cross-validation or complementary insights. By directly injecting existing data into a high-fidelity complex-geometry Navier-Stokes (N-S) solver, we develop a framework that finds the best-matching simulation to observations or theory. This approach has the potential to enhance both the accuracy and efficiency of simulations, as well as improve measurement techniques for quantities that are currently challenging to obtain.

Specifically, this work introduces an adjoint-based data assimilation algorithm that assimilates external time-varying velocity data with a numerical solution to the incompressible N-S equations. The external samples of velocity can be obtained in an Eulerian or Lagrangian frame of reference, allowing the algorithm to be applied to data from various measurement sources such as hot wires, particle image velocimetry (PIV), and particle tracking velocimetry (PTV), etc. As illustrated by Figure 14.1, given a set of coarse velocity samples inside a control volume (marked as gray dots), the algorithm solves a minimization problem for the L^2 norm of velocity difference between numerical simulations and external data. The independent parameters to be optimized are the unknown

¹University of Minnesota, Minneapolis, MN

initial and boundary conditions of velocity. The mass and momentum conservation laws governed by the incompressible N-S equations are included as equality constraints of the optimization. By optimizing the initial and boundary conditions given to the N-S solver, the algorithm finds a simulated velocity field that best matches the given external data and reconstructs non-measured quantities such as pressure and vorticity.

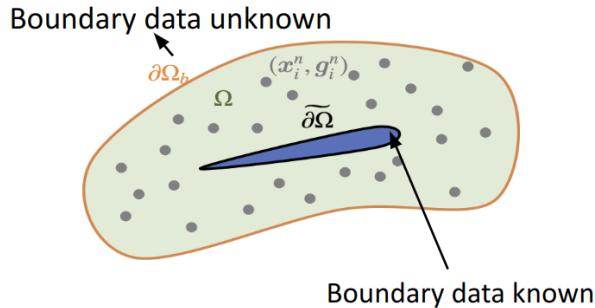


Figure 14.1: Schematic of the data assimilation problem.

Firmly grounded in numerical analysis and the first principles of non-linear flow physics, the current data assimilation algorithm distinguishes itself from other adjoint-based methods in two ways. Firstly, the modeled solution is obtained using a high-order, non-dissipative numerical scheme on unstructured meshes. This enables flexible and adaptive mesh designs around complex geometries tailored to the flow configuration and sample distribution. Secondly, beyond optimizing the unknown initial condition of the flow field, the algorithm also optimizes unknown time-varying inflow/outflow boundary conditions. This would allow for the data assimilation domain to follow Lagrangian particles in space and time.

The remainder of the report is organized as follows. Section 14.2 summarizes the assimilation algorithm; section 14.3 presents a few canonical test cases where synthetic PIV or PTV sample data are used to showcase various capabilities of the algorithm; section 14.4 discusses the computational cost of the algorithm, followed by conclusions and future work in section 14.5.

14.2 Methods

Overview of the assimilation algorithm

Figure 14.2 shows the data assimilation algorithm as an iterative process of finding the optimal initial and boundary conditions for the N-S solver. At the start of the process, an initial guess on the initial and boundary conditions for the velocity field is generated based on interpolating the available sample data. This initial guess is then fed into the numerical solver, and the resulting velocity solutions are compared with the external velocity data. The differences between the data become the forcing in the adjoint momentum equation. The adjoint equation is solved backward in time and gives the most sensitive direction in perturbing the initial and boundary conditions that will lead to a decrease in velocity differences. With the adjoint solution, the quasi-Newton method is then used to systematically determine a search direction and a step size to update the initial and boundary conditions. With the newly updated conditions, a new iteration will start. This iterative process proceeds until a local minimum of the objective function is found. As part of the numerical solution, this process also finds quantities such as pressure and vorticity that are consistent with the assimilated velocity field.

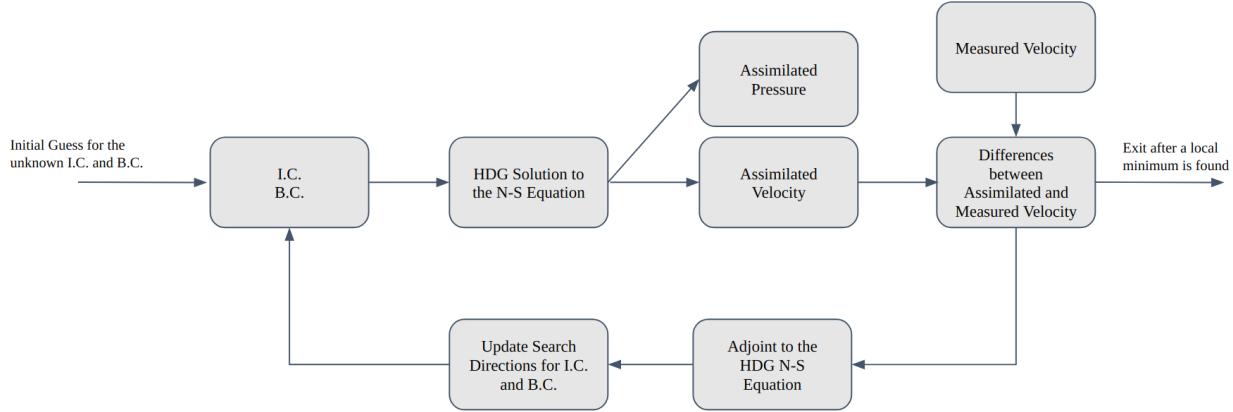


Figure 14.2: Schematic of the assimilation algorithm.

HDG numerical scheme

The Hybridizable Discontinuous Galerkin (HDG) method is used to solve the incompressible Navier-Stokes equations. Full details of the method are presented in Anantharamu and Mahesh (2024). The computational domain is discretized by unstructured tetrahedral cells. The algorithm seeks a numerical solution for the velocity gradient \mathbf{L} and velocity \mathbf{u} in each cell K , as well as pressure p and the hybridized velocity $\hat{\mathbf{u}}$, defined only on the faces of the discretized domain, that satisfy

$$(\mathbf{L}, \mathbf{G})_K + (\mathbf{u}, \nabla \cdot \mathbf{G})_K = \langle \mathbf{u}, \mathbf{G} \mathbf{n} \rangle_{\partial K}, \quad (14.1)$$

$$\left(\frac{\partial \mathbf{u}}{\partial t}, \mathbf{v} \right)_K - (\mathbf{u} \otimes \mathbf{u}, \nabla \mathbf{v})_K + \langle (\hat{\mathbf{u}} \cdot \mathbf{n}) \hat{\mathbf{u}}, \mathbf{v} \rangle_{\partial K} = -(\nabla \cdot \nu \mathbf{L}, \mathbf{v})_K + \langle p, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial K} + (\mathbf{f}, \mathbf{v})_K, \quad (14.2)$$

$$\langle \mathbf{u} \cdot \mathbf{n}, q \rangle_{\partial K} = \langle \hat{\mathbf{u}} \cdot \mathbf{n}, q \rangle_{\partial K}, \quad (14.3)$$

where $(\mathbf{w}_1, \mathbf{w}_2)_K$ and $\langle \mathbf{w}_1, \mathbf{w}_2 \rangle_F$ are short hand notations for volume and surface integrals of integrand $\mathbf{w}_1 \cdot \mathbf{w}_2$ in a volume K and a face F respectively. ∂K refers to the set of bounding faces of the element K . $\mathbf{G}, \mathbf{v}, q$ are test functions in the finite dimensional functional spaces that guarantee well-posedness of the weak formulation (see Anantharamu and Mahesh (2024) for detailed definitions). The solution space to approximate \mathbf{u} and its test function space are spanned by element-wise divergence free polynomial basis functions. This means the incompressibility condition is satisfied by construction inside K . The solutions for all subdomains $K \oplus \partial K$ are then stitched together via the face conditions. For each face F in the discretized computational domain,

$$\begin{cases} \sum_{K \in K^+, K^-} \langle (\nu \mathbf{L} - p \mathbf{I}) \mathbf{n}, \mu \rangle_F = 0, & \text{if } F \in F_I, \\ \langle (\nu \mathbf{L} - p \mathbf{I}) \mathbf{n}, \mu \rangle_F = -\langle p_\infty \mathbf{n}, \mu \rangle_F, & \text{if } F \in F_O, \\ \langle \hat{\mathbf{u}}, \mu \rangle_F = \langle \mathbf{u}_D, \mu \rangle_F, & \text{if } F \in F_D, \end{cases} \quad (14.4)$$

where F_I, F_O, F_D denote the set of interior, farfield outflow and Dirichlet boundary (for velocity) faces. For numerical stability, the momentum flux term on the faces, $(\hat{\mathbf{u}} \cdot \mathbf{n}) \hat{\mathbf{u}}$ in Eqn. 14.2, is

approximated by

$$(\hat{\mathbf{u}} \cdot \mathbf{n}) \hat{\mathbf{u}} \approx \begin{cases} (\hat{\mathbf{u}} \cdot \mathbf{n}) \left(\frac{1}{2} (\mathbf{u}_T^- + \mathbf{u}_T^+) + (\hat{\mathbf{u}} \cdot \mathbf{n}) \mathbf{n} \right), & \text{if } F \in F_I, \\ (\hat{\mathbf{u}} \cdot \mathbf{n}) \mathbf{u}, & \text{if } F \in F_O, \\ (\hat{\mathbf{u}} \cdot \mathbf{n}) (H(\hat{\mathbf{u}} \cdot \mathbf{n}) \mathbf{u} + H(-\hat{\mathbf{u}} \cdot \mathbf{n}) \hat{\mathbf{u}}), & \text{if } F \in F_D, \end{cases} \quad (14.5)$$

where \mathbf{u}_T^\pm is the tangential velocity inside cells that share the interior face $F \in F_I$. H denotes the Heaviside step function. The equations are advanced in time using the 2nd-order Adams-Bashforth scheme for the convective terms and the implicit midpoint scheme for the viscous terms.

Adjoint of the HDG system

The adjoint equations to the discretized N-S system is used to ensure the numerical consistency between the gradient evaluation and the objective function evaluation. In each element, the adjoint velocity gradient \mathbf{L}_h^\dagger , the adjoint velocity \mathbf{u}_h^\dagger and the adjoint pressure $p_{h,0}^\dagger$ are governed by,

$$(\mathbf{L}_h^{\dagger n}, \mathbf{G})_K + \frac{1}{2} \left(\mathbf{u}_h^{\dagger n - \frac{1}{2}} + \mathbf{u}_h^{\dagger n + \frac{1}{2}}, \nabla \cdot \mathbf{G} \right)_K = \langle \hat{\mathbf{u}}^{\dagger n}, \mathbf{G} \mathbf{n} \rangle_{\partial K \setminus F_D}, \quad (14.6)$$

$$\begin{aligned} \frac{1}{\Delta t} \left(\mathbf{u}_h^{\dagger n - \frac{1}{2}} - \mathbf{u}_h^{\dagger n + \frac{1}{2}}, \mathbf{v}_h \right)_K - \left(\nabla \cdot \nu \mathbf{L}_h^{\dagger n}, \mathbf{v}_h \right)_K + \langle p_{h,0}^{\dagger n}, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} + \mathbf{C} \cdot \mathbf{T}^n \\ = \sum_{i=1}^{P_k} \left(\mathbf{u}_h^n(\mathbf{x}_i^{(K)}) - \mathbf{g}_h^n(\mathbf{x}_i^{(K)}) \right) \cdot \mathbf{v}_h(\mathbf{x}_i^{(K)}), \end{aligned} \quad (14.7)$$

where $\mathbf{C} \cdot \mathbf{T}^n$ is the nonlinear adjoint convective terms at time n , expressed in full detail in the appendix. Equation 14.7 shows that the differences between the simulated velocity and the sample velocity, $\mathbf{u}_h^n(\mathbf{x}_i^{(K)}) - \mathbf{g}_h^n(\mathbf{x}_i^{(K)})$, summed over all the sample locations inside each element, becomes an effective impulse forcing term for the adjoint momentum evolution. Further, we have

$$\frac{1}{2} \left\langle (\mathbf{u}_h^{\dagger n - \frac{1}{2}} + \mathbf{u}_h^{\dagger n + \frac{1}{2}}) \cdot \mathbf{n}, q_{h,0} \right\rangle_{\partial K} = \langle \hat{\mathbf{u}}^{\dagger n} \cdot \mathbf{n}, q_{h,0} \rangle_{\partial K \setminus F_D}, \quad (14.8)$$

$$\langle \hat{\mathbf{u}}^{\dagger n} \cdot \mathbf{n}, 1 \rangle_{\partial K \setminus F_D} = 0. \quad (14.9)$$

Over the faces F , we have

$$\sum_{K \in \{K^-, K^+\}} \left\langle \left(\nu \mathbf{L}_h^{\dagger n} - (p_{h,0}^{\dagger n} + \bar{p}_{h,\partial K}^{\dagger n}) \mathbf{I} \right) \mathbf{n}, \boldsymbol{\mu}_h \right\rangle_F = 0, \quad \forall F \in F_I, \quad (14.10)$$

$$\left\langle \left(\nu \mathbf{L}_h^{\dagger n} - (p_{h,0}^{\dagger n} + \bar{p}_{h,\partial K}^{\dagger n}) \mathbf{I} \right) \mathbf{n}, \boldsymbol{\mu}_h \right\rangle_F = 0, \quad \forall F \in F_O, \quad (14.11)$$

$$\begin{aligned} \langle \hat{\mathbf{u}}^{\dagger n}, \boldsymbol{\mu}_h \rangle_F = - \left\langle \left(\nu \mathbf{L}_h^{\dagger n} - (p_{h,0}^{\dagger n} + \bar{p}_{h,\partial K}^{\dagger n}) \mathbf{I} \right) \mathbf{n}, \boldsymbol{\mu}_h \right\rangle_F \\ - \left\langle \left(\frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right) (H(-\hat{\mathbf{u}}_h^n \cdot \mathbf{n}) \mathbf{u}_h^n \cdot \mathbf{n}), \boldsymbol{\mu}_h \right\rangle_F, \quad \forall F \in F_D. \end{aligned} \quad (14.12)$$

This set of equations is solved with the same algorithm as described by the original HDG algorithm. With the obtained adjoint solutions, the gradients of the objective function with respect to the initial and boundary conditions can be computed.

The Fréchet derivative with respect to the initial condition is, for each element K ,

$$\begin{aligned} \left(\frac{\delta \mathcal{L}}{\delta \mathbf{u}_h^0}, \mathbf{v}_h \right)_K &= \left(\mathbf{u}_h^{\dagger \frac{1}{2}}, \mathbf{v}_h \right)_K + \frac{\Delta t}{2} \left(\nu (\nabla \cdot \mathbf{L}_h^{\dagger \frac{1}{2}}), \mathbf{v}_h \right)_K \\ &\quad - \frac{\Delta t}{2} \left\langle p_{h,0}^{\dagger \frac{1}{2}}, \mathbf{v}_h \cdot \mathbf{n} \right\rangle_{\partial K} - \Delta t \mathbf{C} \cdot \mathbf{T}^0, \end{aligned} \quad (14.13)$$

where $\mathbf{C} \cdot \mathbf{T}^0$ is the contribution from the nonlinear convective term at time 0.

The Fréchet derivative with respect to the Dirichlet boundary condition is, for each $F \in \mathcal{F}_D$

$$\left\langle \frac{\delta \mathcal{L}}{\delta \hat{\mathbf{u}}_h^n}, \boldsymbol{\mu}_h \right\rangle_F = -\Delta t \left\langle \left(\nu \mathbf{L}_h^{\dagger n} - (p_{h,0}^{\dagger n} + \bar{p}_{h,\partial K}^{\dagger n}) \mathbf{I} \right) \mathbf{n}, \boldsymbol{\mu}_h \right\rangle_F - \Delta t \mathbf{C} \cdot \mathbf{T} \cdot \mathbf{B} \cdot \mathbf{C}^n, \quad (14.14)$$

where n varies from 1 to N , and $\mathbf{C} \cdot \mathbf{T} \cdot \mathbf{B} \cdot \mathbf{C}^n$ is the contribution from the nonlinear convective term at time n with full details given in the appendix. Lastly, at the time step 0,

$$\left\langle \frac{\delta \mathcal{L}}{\delta \hat{\mathbf{u}}_h^0}, \boldsymbol{\mu}_h \right\rangle_F = -\frac{1}{2} \Delta t \left\langle \left(\nu \mathbf{L}_h^{\dagger \frac{1}{2}} - (p_{h,0}^{\dagger \frac{1}{2}} + \bar{p}_{h,\partial K}^{\dagger \frac{1}{2}}) \mathbf{I} \right) \mathbf{n}, \boldsymbol{\mu}_h \right\rangle_F - \Delta t \mathbf{C} \cdot \mathbf{T} \cdot \mathbf{B} \cdot \mathbf{C}^0. \quad (14.15)$$

These Fréchet derivatives are then used by the quasi-Newton solver to solve the optimization problem for the unknown initial and boundary conditions of the assimilated flow system.

Pressure reconstruction from assimilated velocity field

To obtain the full pressure field, some extension to the original formulation (Anantharamu and Mahesh, 2024) is required. Recall the momentum equation of the HDG method starts with the weak form

$$\left(\frac{\partial \mathbf{u}_h}{\partial t}, \mathbf{v}_h \right)_K - (\nabla \cdot \nu \mathbf{L}_h, \mathbf{v}_h)_K + (\nabla p_h, \mathbf{v}_h)_K = (\mathbf{g}_h, \mathbf{v}_h)_K. \quad (14.16)$$

Applying the divergence theorem,

$$\left(\frac{\partial \mathbf{u}_h}{\partial t}, \mathbf{v}_h \right)_K - (\nabla \cdot \nu \mathbf{L}_h, \mathbf{v}_h)_K + \langle p_h, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} - (p_h, \nabla \cdot \mathbf{v}_h)_K = (\mathbf{g}_h, \mathbf{v}_h)_K, \quad (14.17)$$

and using $\nabla \cdot \mathbf{v}_h = 0$, the equation reduces to

$$\left(\frac{\partial \mathbf{u}_h}{\partial t}, \mathbf{v}_h \right)_K - (\nabla \cdot \nu \mathbf{L}_h, \mathbf{v}_h)_K + \langle p, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} = (\mathbf{g}_h, \mathbf{v}_h)_K. \quad (14.18)$$

Note that here we have absorbed the non-linear convective term into the body force term on the right hand side by invoking the explicit time-stepping scheme and the numerical flux approximation scheme on cell faces. Because the solution of elementary velocity field \mathbf{u}_h to the HDG equations is expanded using divergence-free basis functions, $\nabla \cdot \mathbf{v}_h = 0$ by construction. As a result, p only appears in the surface integral terms of the PDE system, therefore, the volumetric pressure field inside each element becomes undefined. The approximated solution for p is defined only on the faces of the discretized domain,

$$p_h(\partial K) = \bar{p}_{h,\partial K} + p_{h,0}. \quad (14.19)$$

The term $\bar{p}_{h,\partial K}^n$ is the non-zero average of the pressure over all faces of each element K , and the second term is the zero average part

$$\bar{p}_{h,\partial K} = \frac{1}{|\partial K|} \langle p_h(\partial K), 1 \rangle_{\partial K}, \quad (14.20)$$

$$\langle p_{h,0}, 1 \rangle_{\partial K} = 0. \quad (14.21)$$

Only the second term will appear in Eqn. 14.18, since the constant part will integrate to zero:

$$\begin{aligned} \langle p, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} &= \langle \bar{p}_{h,\partial K}, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} + \langle p_{h,0}, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} \\ &= \bar{p}_{h,\partial K} \langle 1, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} + \langle p_{h,0}, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K} \\ &= \langle p_{h,0}, \mathbf{v}_h \cdot \mathbf{n} \rangle_{\partial K}. \end{aligned}$$

To reconstruct the pressure inside each cell volume, denoted as p_h^V , we need to retain the term $\langle p, \nabla \cdot \mathbf{v}_h \rangle_K$ by extending the test functional space $\mathbf{v}_h \in \mathcal{S}_{DF} \oplus \mathcal{S}_{DF}^\perp$. So the test function is composed of a divergence-free and a curl-free portion, $\mathbf{v}_h = \mathbf{v}_h^{DF} + \mathbf{v}_h^\perp$. Substituting into Eqn. 14.17, we have

$$\begin{aligned} \left(\frac{\partial \mathbf{u}_h}{\partial t}, \mathbf{v}_h^{DF} \right)_K + \left(\frac{\partial \mathbf{u}_h}{\partial t}, \mathbf{v}_h^\perp \right)_K - \left(\nabla \cdot \nu \mathbf{L}_h, \mathbf{v}_h^{DF} \right)_K - \left(\nabla \cdot \nu \mathbf{L}_h, \mathbf{v}_h^\perp \right)_K \\ + \langle p_h, \mathbf{v}_h^{DF} \cdot \mathbf{n} \rangle_{\partial K} + \langle p_h, \mathbf{v}_h^\perp \cdot \mathbf{n} \rangle_{\partial K} - \left(p_h^V, \nabla \cdot \mathbf{v}_h^\perp \right)_K \\ = (\mathbf{g}_h, \mathbf{v}_h^{DF})_K + (\mathbf{g}_h, \mathbf{v}_h^\perp)_K. \end{aligned} \quad (14.22)$$

Subtracting this equation from Eqn. 14.18,

$$\left(\frac{\partial \mathbf{u}_h}{\partial t}, \mathbf{v}_h^\perp \right)_K - \left(\nabla \cdot \nu \mathbf{L}_h, \mathbf{v}_h^\perp \right)_K + \langle p_h, \mathbf{v}_h^\perp \cdot \mathbf{n} \rangle_{\partial K} - \left(p_h^V, \nabla \cdot \mathbf{v}_h^\perp \right)_K = (\mathbf{g}_h, \mathbf{v}_h^\perp)_K. \quad (14.23)$$

Because \mathbf{u}_h is also in the functional space \mathcal{S}_{DF} , the first term drops out. Upon rearranging, we have

$$\left(p_h^V, \nabla \cdot \mathbf{v}_h^\perp \right)_K = - \left(\nabla \cdot \nu \mathbf{L}_h, \mathbf{v}_h^\perp \right)_K + \langle p_h, \mathbf{v}_h^\perp \cdot \mathbf{n} \rangle_{\partial K} - (\mathbf{g}_h, \mathbf{v}_h^\perp)_K. \quad (14.24)$$

Note the second term on the right hand side requires the evaluation of p_h along the bounding faces of each cell, which is obtained from the original HDG formulation. In short, the original surface-valued pressure solution becomes the boundary condition for the pressure reconstruction problem in each element.

Parallel implementation of the algorithm

The computational cost of the data assimilation algorithm is determined by the numerical and physical constraints imposed by the dimensionality of the unsteady non-linear flow dynamics. The numerical cost for obtaining an accurate simulation to a flow field increases as the range of spatial and temporal scales of the flow system increase. Additionally, if nonlinearity of the flow dynamics leads to a chaotic system, the time horizon in which the adjoint solution can trace the flow dynamics backward in time will be limited by the Lyapunov timescale. Therefore, a long temporal series of chaotic flow data would need to be assimilated in multiple realizations of shorter time windows.

Given these numerical and physical considerations, an in-house parallel implementation of the assimilation algorithm has been completed. We have also developed a right conditioned generalized minimal residual (GMRES) scheme for the HDG global linear system, coupled with the Trilinos library for its algebraic multigrid iterative solver.

14.3 Results and discussions

We first present three small test cases to demonstrate the capabilities of the assimilation algorithm, followed by a more complex test case involving a forced isotropic turbulent flow field. The first three cases are presented to highlight three key features of the algorithm: the ability to assimilate velocity field with unknown Dirichlet boundary conditions while enforcing mass and momentum conservation, the ability to reconstruct non-sampled quantities such as pressure, and the ability to handle PTV-like datasets. The last example is used to showcase large-scale flow assimilation achieved by the parallel implementation of the algorithm.

2D convecting vortex

Figure 14.3 shows the first assimilation case for a 2D convecting vortex, described in Poinsot and Lele (1992). At time 0 the vortex center is placed at the left inflow boundary of the domain, with an analytical expression:

$$u = U_0 + \frac{\partial \psi}{\partial x_2}, \quad v = -\frac{\partial \psi}{\partial x_1}, \quad \psi = C \exp\left(-\frac{x^2 + y^2}{2R_c^2}\right),$$

where C determines the vortex strength and R_c sets the vortex radius. As time advances, the vortex will move into the domain with a convecting velocity U_0 while its core vorticity decays slowly by viscous dissipation. For the case presented, $C = 1$, $U_0 = 20$, $R_c = \sqrt{2}/4$. Figure 14.3(a-c) shows a temporal snapshot of the flow field considered as ground truth, obtained by time-advancing the N-S equations at $\Delta t = 10^{-4}$ with $\nu = 10^{-2}$ for 30 time steps using the prescribed analytical boundary condition.

The assimilation domain, mesh setup, and sampled points are shown in Figure 14.3(g). The assimilation takes the sampled velocity data as input and seeks an optimal combination of initial and time-varying boundary conditions that will yield a solution that minimizes its differences from the sampled velocity. The assimilated results are plotted in Figure 14.3(b, d, f), in comparison to the true fields shown in Figure 14.3(a, c, e). The RMS of the assimilation errors over each timestep is presented in Figure 14.3(h). Without specifying the exact initial condition and the boundary condition on the left inlet plane, the assimilation algorithm successfully finds the optimal conditions to achieve a low assimilation error. With the assimilated solution, non-sampled quantities such as vorticity can also be obtained accurately.

3D Beltrami flow

The second assimilation case is a 3D Beltrami flow field, given as an exact analytical solution to the incompressible N-S equations with periodic boundary conditions in all three directions,

$$\begin{aligned} u &= (\sin z + \cos y) e^{-\nu t}, & v &= (\sin x + \cos z) e^{-\nu t}, & w &= (\sin y + \cos x) e^{-\nu t}, \\ p &= -(\cos x \sin y + \sin x \cos z + \sin z \cos y) e^{-2\nu t}. \end{aligned}$$

This case highlights the algorithm's ability to reconstruct pressure field based on purely velocity data. For the assimilation, scattered velocity samples are provided to the assimilation using the analytical solution. By optimizing the initial condition for velocity, the algorithm obtains a 3D velocity field that closely matches the analytical solution, as seen in Figure 14.4. The velocity field from the assimilated solution is compared with the ground truth on the plane $z = \pi$ at one specific time instant. Using the assimilated velocity solution, the pressure field can be reconstructed accurately by solving Eqn. (14.24).

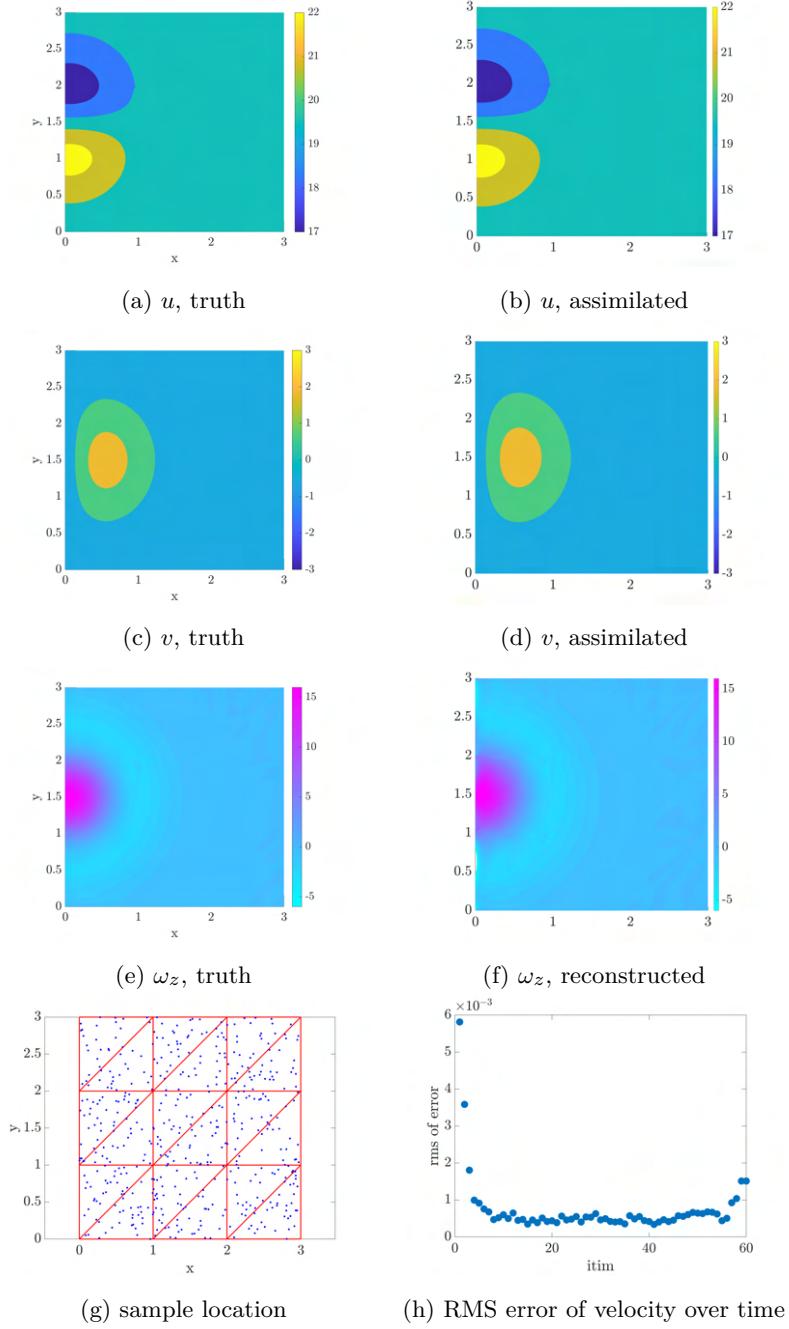


Figure 14.3: Assimilation of a 2D convecting vortex, where the left Dirichlet boundary condition is unknown.

Sub-volume assimilation with 3D synthetic PTV data

The third case illustrates the usage of data from a Lagrangian frame of reference and the algorithm's potential to break down a larger flow domain into smaller subdomains for assimilation. The ground truth is obtained by numerically solving the N-S equations initiated with 16 randomly oriented Fourier modes in a 3D periodic box of length 2π . Scattered velocity data are then sampled in a Lagrangian frame of reference within a subvolume of the periodic box. At time 0, 16^3 points are randomly selected inside the subdomain at $[\pi/2, 3\pi/2] \times [\pi/2, 3\pi/2] \times [\pi/2, 3\pi/2]$. The Lagrangian

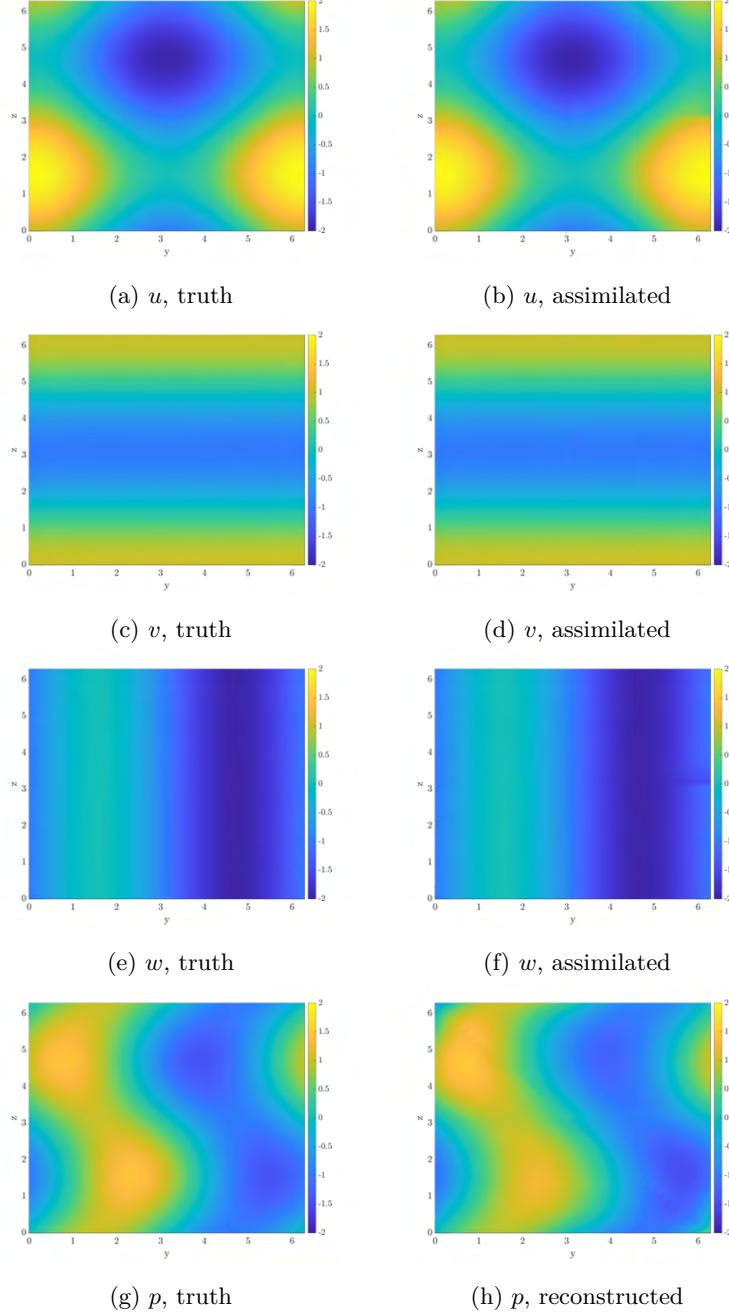


Figure 14.4: Assimilation of a 3D Beltrami flowfield. Note that pressure is reconstructed as a result of the assimilated velocity field.

pathlines are obtained by integrating the velocity over time from these initially selected locations. All sampled pathlines are plotted in Figure 14.5(a), where the line color changes from yellow to red as time increases. Figure 14.5(b) shows a zoom-in view of the pathlines in the region over $[3\pi/4, 5\pi/4] \times [3\pi/4, 5\pi/4] \times [3\pi/4, 5\pi/4]$. Using these PTV like samples, the assimilation algorithm finds the optimal combination of initial condition and time-varying boundary conditions for the smaller subdomain, bounded by the red box in Figure 14.5(c, e, g). The assimilated field shown in Figure 14.5(d, f, h) closely resembles the ground truth data.

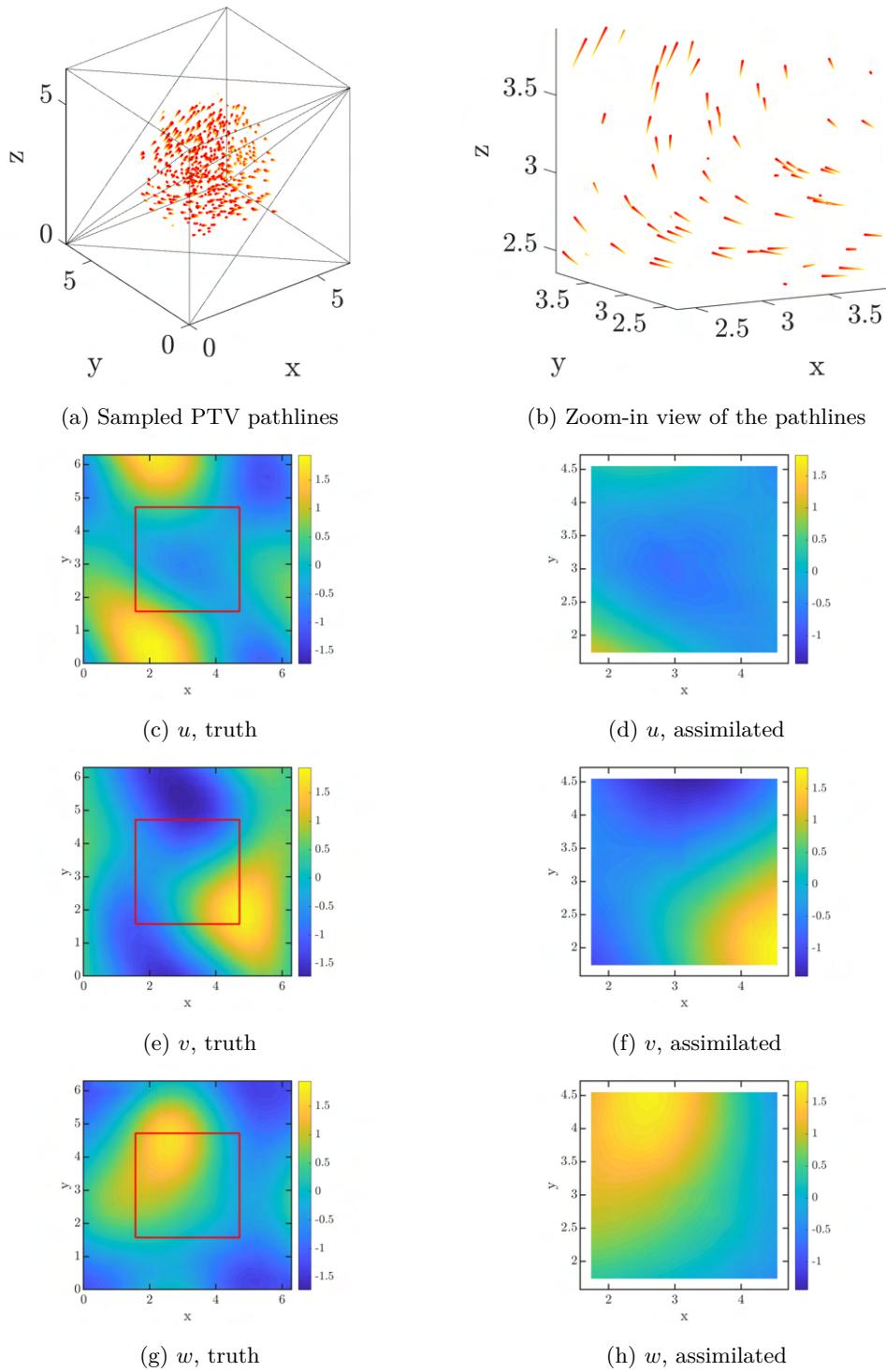


Figure 14.5: Assimilation of a synthetic PTV data set from a 3D flow initiated with randomly oriented Fourier modes. Red boxes show the boundaries of the assimilated domain.

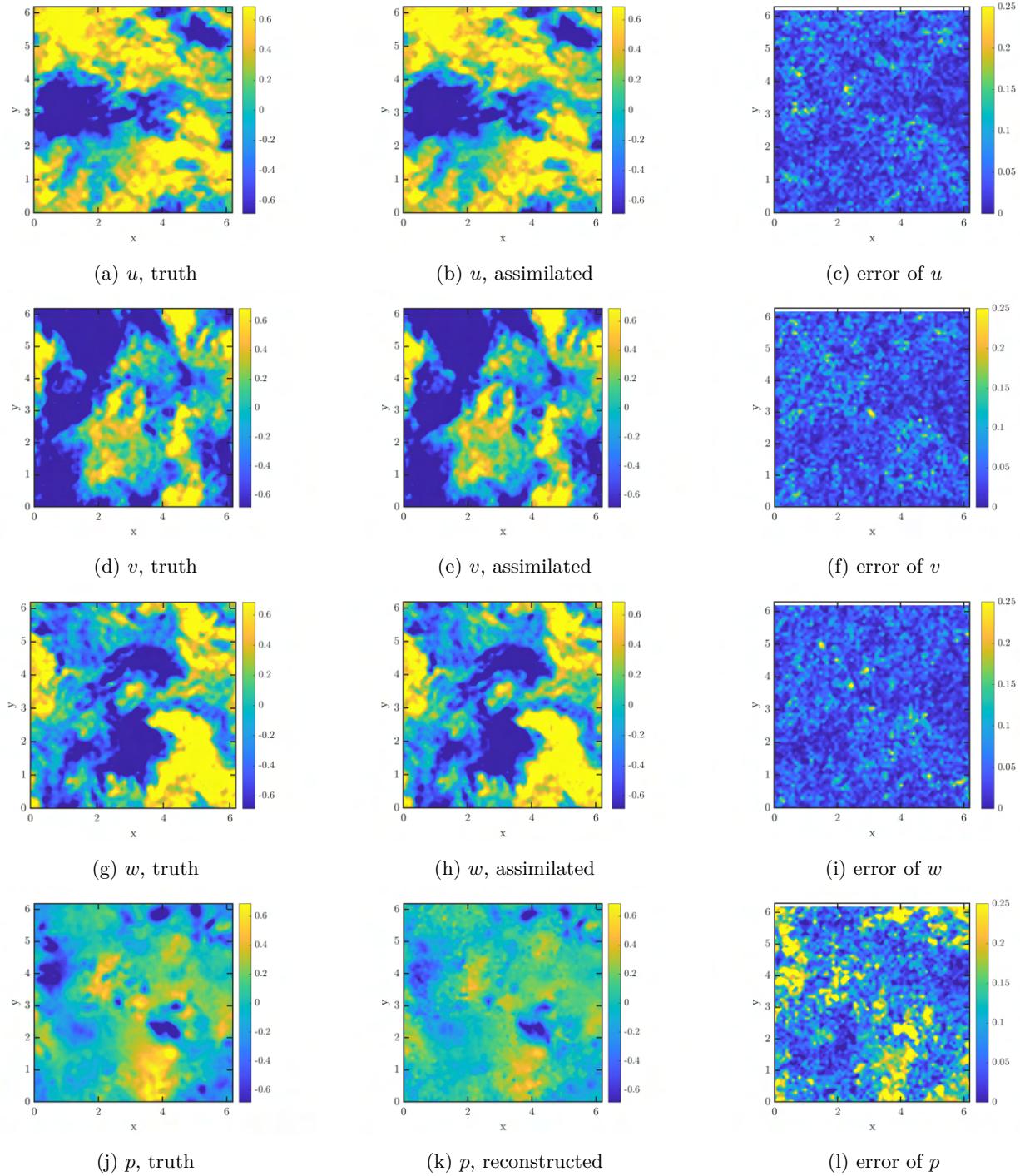


Figure 14.6: Assimilation of a forced isotropic turbulent flow field. Sample obtained from the JHU turbulent database (Li et al., 2008; Perlman et al., 2007).

Forced isotropic turbulence

The last test case involves a forced isotropic turbulent field in a periodic box. The parameters to be optimized are the initial velocity condition. The ground truth data are sampled from a direct numerical simulation (DNS) provided by the Johns Hopkins University Turbulence Database (Li

et al., 2008; Perlman et al., 2007). The original DNS was conducted on a structured grid with 1024 points over a length of 2π in each dimension, while the data sample used for assimilation is collected for 64 points along each dimension. At the same spatial sample locations, four time steps of data are collected at an interval of 0.0001 eddy turnover time. With these samples as ground truth, the algorithm conducted assimilation using an unstructured grid containing roughly 20 elements per dimension and 3rd-order basis functions. Figure 14.6 compares the ground truth data and the assimilated solution at the sampled location. The velocity and pressure fields are visualized on the plane $z = 1.5\pi$ for the last sampled time step. The results indicate the assimilation algorithm successfully obtains the range of turbulent flow scales contained in the sample data. The error plots, generated by computing the absolute value of the relative error between the assimilated field and the truth field, indicate the error distribution in space is non-uniform and containing high wavenumber components. Moreover, the error of pressure is much higher than that of velocity, which is expected since the pressure solution is approximated with lower-order basis functions than velocity. Figure 14.7 shows the change in RMS error of velocity, normalized by u_{rms} , over each quasi-Newton iteration. Before the first quasi-Newton iteration, an initial guess on the element-wise initial condition is obtained via a least-square fit to the sample data at the first sampled time. The error starts off with a sharp dropoff from iteration 1 to 2, followed by a slower decay rate afterwards. The assimilation results are quite satisfactory after 14 iterations, and further iterations will drop the error to even smaller levels.

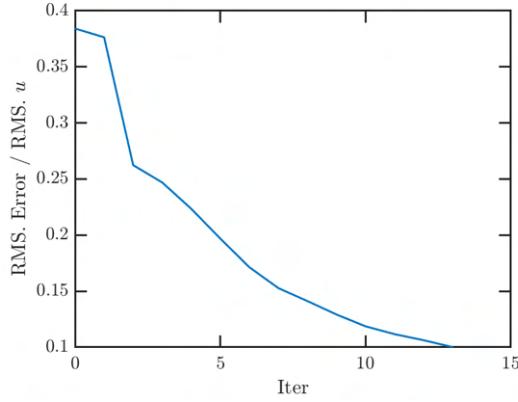


Figure 14.7: RMS error of velocity over each quasi-Newton iteration for the JHU turbulence case.

14.4 Computational cost of adjoint-based assimilation

The adjoint-based assimilation algorithm can be broken down into four main computational tasks: basis generation, local solution, global system assembly and optimization. Basis generation refers to the construction of basis functions used to approximate the discretized velocity and pressure solutions. For each mesh element, the velocity basis is computed using a modified Arnoldi iteration to generate the divergence-free and orthogonal polynomial functions locally. Local solution refers to the stage where the local solution of the HDG equation is obtained; it is posed as a boundary value problem for each single mesh element. Global system assembly is to assemble local solutions from each element into a global left hand side (LHS) matrix, where the trace variables on the element faces are coupled globally to enforce the stress continuity across elements and boundary conditions. The last segment refers to the quasi-Newton iteration to solve the optimization problem for data

assimilation. During each quasi-Newton iteration, a number of forward and adjoint evolutions of the HDG global system of equations are computed.

The first three data assimilation test cases shown in Section 14.3 were computed using a serial implementation in MATLAB running on a single PC, equipped with a 16-core Intel i7-13700 CPU processor and 64GB RAM. Given the simplicity of these flow configurations, the algorithm completed the data assimilation within a few minutes. Figure 14.8 shows the computational time for each of the four main computational tasks. The two most costly segments are local solution and assimilation. The first three segments were only done once at the start of the assimilation, while the last segment required multiple iterations of temporal evolutions.

The computational costs for the basis generation and local solution scale as $\mathcal{O}(N_e p^2)$, where N_e is the number of mesh elements and p being the spatial order. The size of the global LHS matrix also scales as $\mathcal{O}(N_e p^2)$, which sets the memory requirement of the algorithm. The computational cost for the last segment scales as $\mathcal{O}(N_e p^2 N_t N_o)$, where N_t denotes the number of discretized time steps over the temporal domain of the data assimilation problem and N_o denotes the number of iterations required for the quasi-Newton method to reach a satisfactory solution.

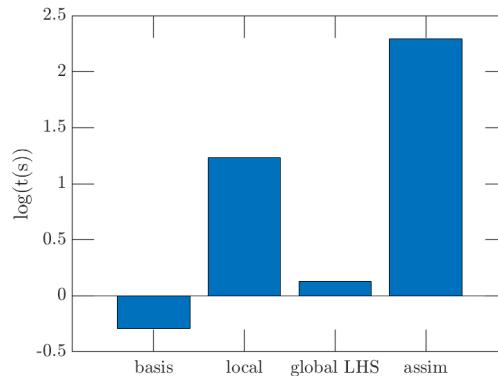


Figure 14.8: Computational cost for the serial version of the assimilation algorithm on a small scale test problem.

The last data assimilation test case shown in Section 14.3 was performed using the MPI-based parallel implementation in Fortran. The number of degrees of freedom (DOFs) for the optimization is about two million, and the computation was conducted with two 96-core AMD EPYC 9654 CPUs over 48 physical hours on the Lighthouse supercomputing cluster. The detailed breakdown of the computational cost, measured by physical time in unit of seconds, is plotted in Figure 14.9. Comparing Figure 14.9(a) with Figure 14.8, it can be seen that the parallel implementation keeps the cost scalable for the first three computational segments: basis generation, local solution and global system assembly. However, for the assimilation phase, the cost increases by three orders of magnitude. Figure 14.9(b) shows the computational time for each iteration step of the quasi-Newton method during assimilation. After the fourth iteration, the computational time history contains various spikes, which is a result of the line search algorithm embedded in the quasi-Newton iteration performing multiple sweeps of forward and backward evolution of the HDG system. This significantly increases the computational cost despite only obtaining a small reduction in the RMS. error as seen in Figure 14.7. To improve the performance, additional tuning to dynamically adjust the initial step size used in the line search is being explored.

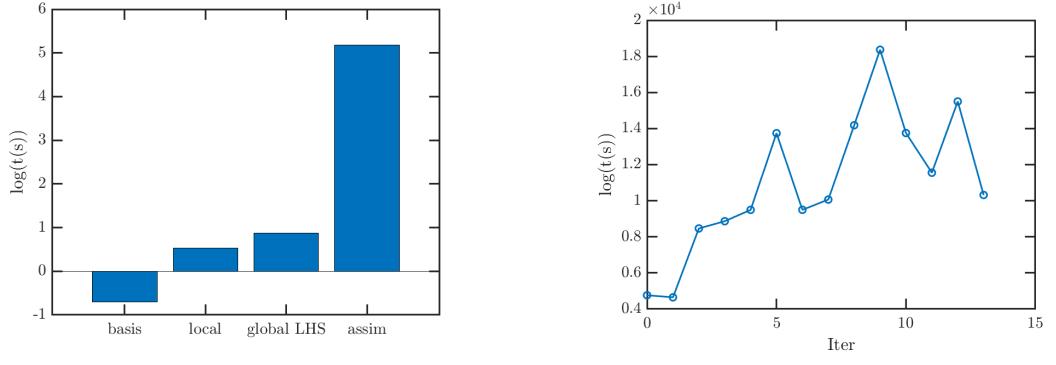


Figure 14.9: Computational cost for the parallel version of the assimilation algorithm for the JHU turbulence test case.

14.5 Conclusions and future work

This work introduces a novel data assimilation algorithm for integrating unsteady high-fidelity simulations with external data generated by theory, numerical simulations, or experimental velocity measurements. The algorithm is capable of minimizing velocity differences between the simulation and the given data source, as well as reconstructing volumetric flow quantities, such as pressure and vorticity, that are not provided in the external data. The demonstrated test cases have assumed sample data to be in the form of PIV or PTV samples, but the method is not limited to these measurement techniques and has the potential to incorporate mean flow data or hot wire measurements.

The algorithm is constructed as a gradient-based optimization, with the sum of squared differences between simulated and external velocity data as the objective function. The free parameters to be optimized are the initial and boundary conditions used as input to the numerical solver. The discretized N-S equations are imposed as equality constraints through Lagrange multipliers. The adjoint of the discretized system is solved to obtain the gradient of objective function with respect to the free parameters. The optimization is solved iteratively with the quasi-Newton method. The flow solver uses an arbitrarily high-order HDG scheme on unstructured mesh with a non-dissipative and discrete kinetic energy conserving flux approximation. To preserve such property on unstructured mesh, the adjoint momentum flux and boundary conditions are derived consistently with the forward discretization scheme. After obtaining the HDG solution for velocity, the pressure field in each element can be reconstructed by extending the test functional space for the momentum equation to include the curl-free basis.

Given the computational cost required for assimilating flow fields involving a wide range of scales, a parallel implementation of the algorithm has been completed. Various canonical flow cases are tested and analyzed, including a fully turbulent flow field. Future work will apply the algorithm for realistic experimental test cases and conduct an in-depth analysis of the computational accuracy and efficiency of the method.

Acknowledgments

This work was partially supported by Grant N00014-21-1-2293. We thank Professor Louis Cattafesta and his team for helpful discussions about PTV data assimilation.

Appendix

We first define two short notations \mathcal{T}_1 and \mathcal{T}_2 . Let \mathbf{a} , \mathbf{b} and \mathbf{c} be three arbitrary vectors, and \mathbf{n} be the face normal vector, then

$$\mathcal{T}_1(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} \otimes \mathbf{b} : \nabla \mathbf{c}, \quad \mathcal{T}_2(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a}(\mathbf{b} \cdot \mathbf{n}) \cdot \mathbf{c}. \quad (14.25)$$

Let $H(\eta)$ denote the unit step function, the nonlinear convective terms are as follows.

$$\begin{aligned} \mathbf{C.T.}^n &= \int_K \mathcal{T}_1 \left(\left(\frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right), \mathbf{u}_h^n, \mathbf{v}_h \right) d\Omega \\ &\quad + \int_K \mathcal{T}_1 \left(\left(\frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right), \mathbf{v}_h, \mathbf{u}_h^n \right) d\Omega \\ &\quad - \sum_{F \in \partial K \cap F_I} \int_F \mathcal{T}_2 \left(\left\{ \left\{ \frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right\} \right\}, \hat{\mathbf{u}}_h^n, \mathbf{v}_h \right) d\Gamma \\ &\quad - \sum_{F \in \partial K \cap F_I} \int_F \mathcal{T}_2 \left(\left\{ \left\{ \frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right\} \right\}, \mu, \mathbf{u}_h^n \right) d\Gamma \\ &\quad - \sum_{F \in \partial K \cap F_D} \int_F \mathcal{T}_2 \left(\left(\frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right), H(-\hat{\mathbf{u}}_h^n \cdot \mathbf{n}) \hat{\mathbf{u}}_h^n, \mathbf{v}_h \right) d\Gamma \\ &\quad + \sum_{F \in \partial K \cap F_D} \int_F \mathcal{T}_2 \left(\left(\frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right), \mu, H(-\hat{\mathbf{u}}_h^n \cdot \mathbf{n}) (\hat{\mathbf{u}}_h^n - \mathbf{u}_h^n) \right) d\Gamma. \end{aligned} \quad (14.26)$$

$$\begin{aligned} \mathbf{C.T.}^0 &= \int_K \mathcal{T}_1 \left(\left(\mathbf{u}_h^{\dagger \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger \frac{3}{2}} \right), \mathbf{u}_h^0, \mathbf{v}_h \right) d\Omega \\ &\quad \int_K \mathcal{T}_1 \left(\left(\mathbf{u}_h^{\dagger \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger \frac{3}{2}} \right), \mathbf{v}_h, \mathbf{u}_h^0 \right) d\Omega \\ &\quad - \sum_{F \in \partial K \cap F_I} \int_F \mathcal{T}_2 \left(\left\{ \left\{ \mathbf{u}_h^{\dagger \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger \frac{3}{2}} \right\} \right\}, \mathbf{u}_h^0, \mathbf{v}_h \right) d\Gamma \\ &\quad - \sum_{F \in \partial K \cap F_I} \int_F \mathcal{T}_2 \left(\left\{ \left\{ \mathbf{u}_h^{\dagger \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger \frac{3}{2}} \right\} \right\}, \mathbf{v}_h, \mathbf{u}_h^0 \right) d\Gamma \\ &\quad - \sum_{F \in \partial K \cap F_D} \int_F \mathcal{T}_2 \left(\left(\mathbf{u}_h^{\dagger \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger \frac{3}{2}} \right), H(-\hat{\mathbf{u}}_h^0 \cdot \mathbf{n}) \mathbf{u}_h^0, \mathbf{v}_h \right) d\Gamma \\ &\quad + \sum_{F \in \partial K \cap F_D} \int_F \mathcal{T}_2 \left(\left(\mathbf{u}_h^{\dagger \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger \frac{3}{2}} \right), \mathbf{v}_h, H(-\hat{\mathbf{u}}_h^0 \cdot \mathbf{n}) (\hat{\mathbf{u}}_h^0 - \mathbf{u}_h^0) \right) d\Gamma. \end{aligned} \quad (14.27)$$

$$\begin{aligned} \mathbf{C.T.B.C.}^n &= \sum_{F \in \partial K \cap F_D} \int_F \mathcal{T}_2 \left(\left(\frac{3}{2} \mathbf{u}_h^{\dagger n + \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger n + \frac{3}{2}} \right), H(-\hat{\mathbf{u}}_h^n \cdot \mathbf{n}) \mathbf{u}_h^n, \boldsymbol{\mu}_h \right) d\Gamma, \\ \mathbf{C.T.B.C.}^0 &= \sum_{F \in \partial K \cap F_D} \int_F \mathcal{T}_2 \left(\left(\mathbf{u}_h^{\dagger \frac{1}{2}} - \frac{1}{2} \mathbf{u}_h^{\dagger \frac{3}{2}} \right), H(-\hat{\mathbf{u}}_h^0 \cdot \mathbf{n}) \mathbf{u}_h^0, \boldsymbol{\mu}_h \right) d\Gamma. \end{aligned} \quad (14.28)$$

References

S. Anantharamu and K. Mahesh. A non-dissipative, energy-conserving, arbitrary high-order numerical method and its efficient implementation for incompressible flow simulation in complex geometries. *International Journal for Numerical Methods in Fluids*, 2024.

References

Y. Li, E. Perlman, M. Wan, Y. Yang, C. Meneveau, R. Burns, S. Chen, A. Szalay, and G. Eyink. A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence. *Journal of Turbulence*, 9(31), 2008.

E. Perlman, R. Burns, Y. Li, and C. Meneveau. Data exploration of turbulence simulations using a database cluster. In *SC '07: Proceedings of the 2007 ACM/IEEE Conference on Supercomputing*, 2007.

T. J. Poinsot and S. K. Lele. Boundary conditions for direct simulations of compressible viscous flows. *Journal of Computational Physics*, 1992.

Development, Application, and Refinement of a New Data-Driven Turbulence Modeling Methodology

NILOY GUPTA & KARTHIK DURAISAMY

This report describes the work performed by the authors over the past year to develop, apply, and improve a systematic methodology for developing data-driven turbulence models. We propose and outline the steps of the methodology, which uses gradient-based optimization and the Learning and Inference Assisted by Feature-space Engineering framework. This methodology is developed in the context of improving turbulent heat transfer modeling for hypersonic flows. Our work on this problem shows that piecewise multilinear maps and properly designed features can provide better accuracy, robustness, and generalizability compared to neural networks, resulting in a preliminary model that successfully generalizes to a different geometry with a different shock mechanism. We also apply this methodology to improve smooth body separation prediction and investigate optimal feature design. Analyzing validation case performance is necessary to determine an optimal set of features. Furthermore, theoretically better features do not always yield improved predictive performance as they can lead to overfitting during training. The proposed methodology is not complete; improvements to the feature design process and piecewise multilinear map implementation will be worked on during the upcoming reporting period.

Keywords: turbulence modeling, data-driven modeling, separated flows

15.1 Introduction

Reynolds-averaged Navier-Stokes (RANS) simulations remain the dominant choice for the analysis of engineering flows due to their balance between accuracy and computational cost. However, the assumptions and simplifications inherent to RANS models result in systematic errors that prevent them from accurately representing complex turbulent flows (Duraisamy et al., 2019). Recently, data-driven modeling and machine learning have emerged as promising avenues for addressing such deficiencies while still maintaining computational efficiency. However, some of the existing approaches in the literature neglect a key concept–model consistency. For turbulence modeling for RANS simulations, it is imperative that the model is trained with consideration to the low-fidelity model. If an augmentation is only trained using data directly taken from high fidelity sources, mismatches between features during training and prediction, accumulation of errors, and introduced imbalances between model terms will prevent it from accurately estimating model discrepancy (Duraisamy, 2021).

To address this concern, Parish and Duraisamy (2016) and Singh and Duraisamy (2016) developed the field inversion and machine learning (FIML) approach. It is a two-step approach where field inversion is first performed to learn a spatially defined discrepancy field quantifying model-form error before it is approximated in terms of non-dimensional input features to create a predictive model. It has been successfully applied to various fluid flow problems, including turbomachinery

modeling (Ferrero et al., 2020), separation prediction (Singh et al., 2017b), and transition prediction (Yang and Xiao, 2020).

Holland et al. (2019) and Sirignano et al. (2020) improved upon FIML by combining discrepancy field estimation and model parameter inference into a single coupled procedure referred to as integrated inference and machine learning (IIML). The augmentation is embedded within the solver during training, and model parameters are optimized directly across all cases simultaneously. This inherently accounts for the model’s functional form during training and improves upon the consistency between the learning and prediction environments, resulting in a more accurate and generalizable model (Holland et al., 2019; Sirignano et al., 2020; Duraisamy, 2021; Fang et al., 2023). Despite these improvements, simple IIML techniques lack safeguards for ensuring the robustness and generalizability of the model. Srivastava and Duraisamy (2021) address this by proposing the Learning and Inference Assisted by Feature-space Engineering (LIFE) framework, which provides guiding principles for developing robust, generalizable turbulence model augmentations using limited training data.

However, further improvements are still required for data-driven modeling frameworks to be used for engineering design and analysis. Data-driven models, particularly those utilizing neural networks, often lack physical interpretability due to their complex, nonlinear structure, which obscures the relationship between inputs and outputs (Kutz, 2017). As a result, they can behave unpredictably when extrapolated beyond their training regime (Zhu et al., 2023; Barnard and Wessels, 1992). Feature-based models are generally more interpretable, but require users to make numerous modeling decisions: the choice of augmented terms, model structure, hyperparameters, and, most critically, the set of input features. These features must identify regions where the baseline turbulence model fails and distinguish relevant physical phenomena. While the proposing papers and other studies justify their chosen features, they often fail to describe the details of their derivation, lack systematic guidance on feature derivation, or use unrecommended quantities according to Spalart (2015).

In this report, we summarize the work conducted over the reporting period to address existing gaps between proposed frameworks for data-driven turbulence modeling in the literature and their practical implementation. Section 15.2 describes the methodology that has been developed. Section 15.3 describes the modeling effort for turbulent heat transfer in hypersonic flows that was used to develop the methodology, and Section 15.4 describes the ongoing work to apply this methodology for improving smooth body separation prediction. Finally, Section 15.5 summarizes our progress and findings and outlines our plans for the upcoming period.

15.2 Turbulence model augmentation and inference methodology

During the reporting period, we developed an inference methodology for designing and training robust, generalizable turbulence model augmentations for RANS simulations. It was developed during the work outlined in Section 15.3. The steps of the methodology are as follows:

1. Identify a physically meaningful augmentation target by addressing the most critical assumptions or modeled terms causing inaccuracies for the flow of interest.
2. Evaluate the viability of the augmentation and determine its optimal form through field inversion (FI), ensuring that the predicted corrections are physically consistent.

3. Simplify the development process via an idealized learning problem, which clarifies how feature choices, feature-space representations, and hyperparameters affect the accuracy and robustness of the final augmentation.
4. Incrementally refine the model by progressively adding features or complexity to address specific shortcomings in the initial augmentation, ensuring targeted performance gains without compromising interpretability.

For our work, this methodology is enabled by two specific components: gradient based optimization using the discrete adjoint method and the LIFE framework.

RANS augmentation using the discrete adjoint method

We introduce an augmentation model into the RANS equations by implicitly defining it as a function of features. With \mathcal{R}_m representing the RANS residual operator that solves for states $\tilde{\mathbf{u}}_m$, the augmented equations can be expressed as

$$\mathcal{R}_{m,\text{aug}}(\tilde{\mathbf{u}}_m; \beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m), \mathbf{w}), \boldsymbol{\xi}) = 0. \quad (15.1)$$

The augmentation model $\beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m), \mathbf{w})$ is evaluated at every finite volume cell in the mesh and parameterized by \mathbf{w} . The cell's states are used to calculate a feature vector $\boldsymbol{\eta}$, which is the input to the model that calculates β , the discrepancy term for that point in the computational domain. Turbulence model performance is assessed using a cost function \mathcal{J} that is explicitly a function of $\tilde{\mathbf{u}}_m$ and implicitly a function of $\beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m), \mathbf{w})$. Provided that a model structure is specified, obtaining the best turbulence model is accomplished by solving the optimization problem

$$\min_{\mathbf{w}} \mathcal{J}(\tilde{\mathbf{u}}_m; \beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m), \mathbf{w})) \quad \text{s.t.} \quad \mathcal{R}_{m,\text{aug}}(\tilde{\mathbf{u}}_m; \beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m), \mathbf{w}), \boldsymbol{\xi}) = 0. \quad (15.2)$$

Since \mathbf{w} is often high-dimensional, this optimization problem is best solved using gradient-based techniques. For our work, we use the discrete adjoint method. Differentiating the cost function $\mathcal{J}(\tilde{\mathbf{u}}_m; \beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m), \mathbf{w}))$ with respect to \mathbf{w} and substitution of terms yields the desired gradient

$$\frac{d\mathcal{J}}{d\mathbf{w}} = \left(\frac{\partial \mathcal{J}}{\partial \beta} + \boldsymbol{\psi}^T \frac{\partial \mathcal{R}_{m,\text{aug}}}{\partial \beta} \right) \frac{\partial \beta}{\partial \mathbf{w}}, \quad (15.3)$$

where $\boldsymbol{\psi}^T$ is adjoint vector that is the solution to the adjoint system given by Giles and Pierce (2000); Giles et al. (2003),

$$\boldsymbol{\psi}^T \frac{\partial \mathcal{R}_{m,\text{aug}}}{\partial \tilde{\mathbf{u}}_m} = - \frac{\partial \mathcal{J}}{\partial \tilde{\mathbf{u}}_m}. \quad (15.4)$$

Learning and inference assisted by feature-space engineering

The Learning and Inference Assisted by Feature-space Engineering (LIFE) framework, proposed by Srivastava and Duraisamy (2021), is an extension of integrated inference and machine learning (IIML) frameworks. IIML frameworks were developed to improve field inversion and machine learning (FIML) by directly optimizing model parameters when learning optimal augmentation fields to ensure that they are realizable by the model structure. This helps to enforce consistency in the learned model across all training cases (Holland et al., 2019; Sirignano et al., 2020; Fang et al., 2023). LIFE builds upon IIML by providing guiding principles and an implementation of them for developing model augmentations that: 1) generalize to unseen configurations and cases,

minimize spurious predictions, and 3) train using limited data. The framework's key contribution is the concept of localized learning. Localized learning techniques restrict modifications of the baseline model to regions of the feature space with training data, ensuring that the model behaves according to the baseline model elsewhere. Since the baseline model is a physically valid solution, this prevents spurious behavior that can arise from simple nonlinear augmentation functions that are commonly used in machine learning (Srivastava and Duraisamy, 2021).

15.3 Application to turbulent heat transfer in hypersonic flows

This section briefly summarizes the work done that yielded the augmentation and inference approach described in Section 15.2. For brevity of the report, only the major conclusions are provided. We refer the reader to Gupta and Duraisamy (2025) for the full details of this effort.

Modeling problem and augmentation choice

Since the most commonly used turbulence models were designed and calibrated for subsonic incompressible flows without significant pressure gradients or heat transfer, they struggle to accurately predict hypersonic flows. In particular, when applied to hypersonic flows exhibiting shock-boundary layer interactions (SBLIs), they grossly overestimate the predicted heat transfer Raje et al. (2025). One of the sources for this error is the turbulent heat transfer model, typically given by

$$c_p \overline{\rho u_j'' T''} \approx - \frac{c_p \bar{\mu}_t}{\text{Pr}_t} \frac{\partial \hat{T}}{\partial x_j}, \quad (15.5)$$

where c_p is the isobaric specific heat constant, and T is the temperature. It is analogous to the model for the mean laminar heat transfer and makes use of Morkovin's hypothesis to provide a constant value for the turbulent Prandtl number Pr_t close to unity (Smits and Dussauge, 2006; Duan et al., 2011). However, in hypersonic flows and flows with shockwaves, compression and dilatation cannot be ignored, invalidating this assumption and requiring a variable- Pr_t model (Rotta, 1965; Horstman and Owen, 1972; Priebe and Martín, 2021). Consequently, we augment our baseline turbulence model, the Wilcox 2006 $k - \omega$ model (Wilcox, 2008), through the modification

$$\text{Pr}_t = f(\beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m))) \text{Pr}_{t,0}, \quad (15.6)$$

where f is some function of the discrepancy field and $\text{Pr}_{t,0} = 0.9$ is the baseline value.

Full and idealized learning problems

We assess the turbulence model and RANS performance using the objective function

$$\mathcal{J} = \frac{1}{n_x} \sum_{i=1}^{n_x} \left(q_{w,m}(x_i) - q_{w,ref}(x_i) \right)^2, \quad (15.7)$$

which is the mean-squared error of the wall heat transfer predictions by the model with respect to the reference data constructed from the experimental data provided in the studies listed in Table 15.1 at n_x wall faces in the computational domain. The reference data is constructed to isolate the increase in heat transfer due to the SBLI and ignore other external phenomena and discrepancies (Gupta and Duraisamy, 2025). The full learning problem can be stated as Eqn. (15.2) with the objective function \mathcal{J} provided by Eqn. (15.7).

| Case | Study | Data Type | Geometry | M_∞ | Angle (°) |
|------|-----------------------------|--------------|-----------------------------|------------|-----------|
| S6 | Schülein (2006) | Experimental | Oblique incident shock wave | 5.00 | 6 |
| S10 | | | | | 10 |
| S14 | | | | | 14 |
| C15 | Coleman and Stollery (1972) | Experimental | Compression corner | 9.22 | 15 |
| C30 | | | | | 30 |

Table 15.1: External data used for training augmentations.

However, directly proceeding to designing a generalizable augmentation is challenging due to the number of design decisions required. Therefore, we formulate an idealized learning problem to learn an existing variable turbulent Prandtl number model. The cost function for this problem is given by

$$\mathcal{J}_k = \frac{1}{n_x} \sum_{i=1}^{n_x} (q_{w,m}(x_i) - q_{w,KC}(x_i))^2, \quad (15.8)$$

which is the mean-squared error of the wall heat transfer predictions by the model and the variable- Pr_t model of Kays and Crawford (1993), henceforth referred to as the idealized model, which can be stated as

$$\text{Pr}_t = \frac{1}{0.5882 + 0.228(\nu_t/\nu) - 0.0441(\nu_t/\nu)^2 \left[1 - \exp\left(\frac{-5.165}{\nu_t/\nu}\right) \right]}. \quad (15.9)$$

The goal of solving this problem is to develop a single-feature model where a function of ν_t/ν is used to recreate the idealized model. However, it cannot be exactly recreated, as many model structures require closed intervals for inputs or bounded inputs for improved performance. Therefore, it serves as an excellent problem for characterizing the numerics of the full learning problem and developing guidelines for optimal feature design.

Findings from field inversion

Following our proposed methodology, we next ensure the viability of our augmentation by performing field inversion (FI). Furthermore, we use FI to determine the best mathematical form of the augmentation among four options: βPr_t , Pr_t/β , $e^\beta\text{Pr}_t$, and $e^{-\beta}\text{Pr}_t$. These four options provide different benefits in terms of numerical stability, unconstrained versus constrained optimization, and emphasis on different inferred values of Pr_t .

Figure 15.1 shows the results of this effort. We ascertain that the variable- Pr_t augmentation can yield the desired improvement in wall heat transfer predictions and identify the multiplicative exponential form ($\text{Pr}_t \rightarrow e^\beta\text{Pr}_t$) as the optimal augmentation. It performs strongly on the Mach 5 oblique shock cases, comparably to Pr_t/β , and provides the most stable optimization for the Mach 9 compression corner cases. Its unconstrained nature allows for simpler optimization methods to be used. We adopt this augmentation for the remainder of the hypersonics modeling work.

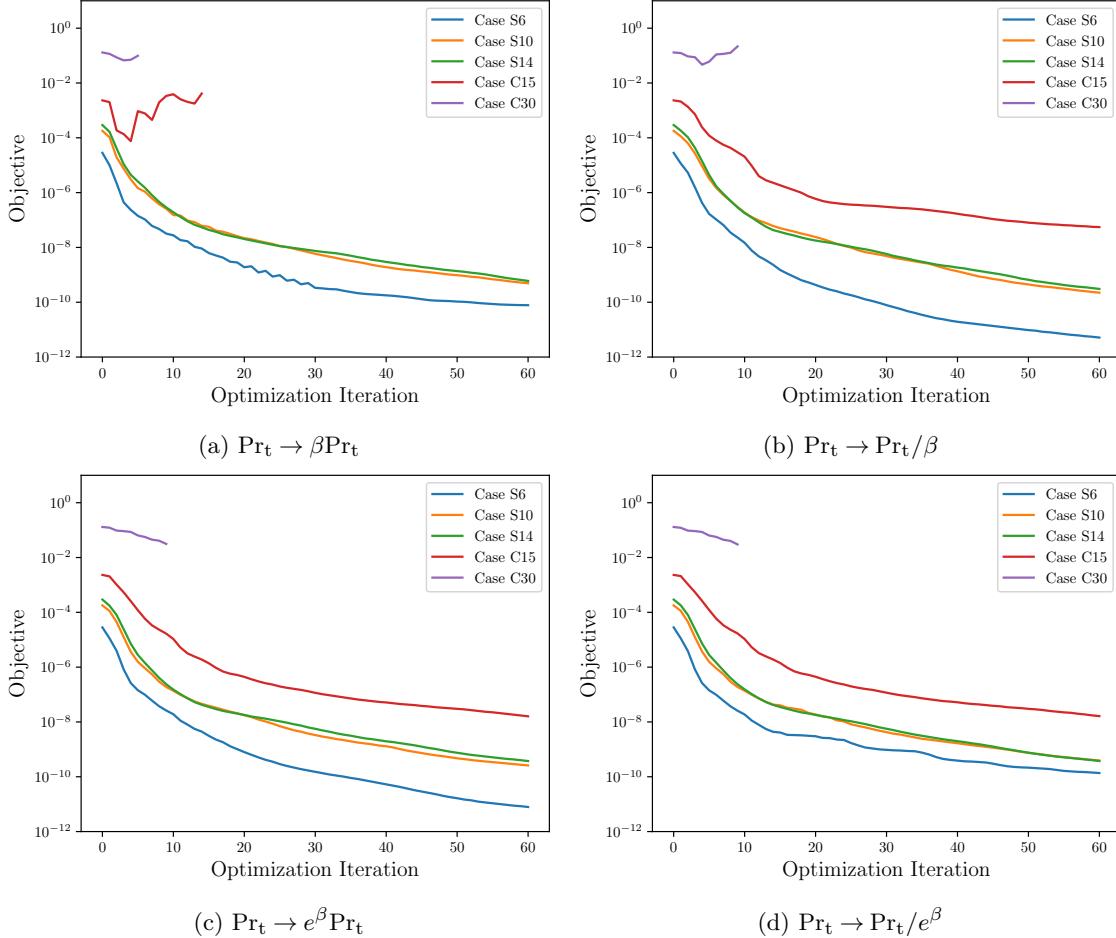


Figure 15.1: FI objective training history for each augmentation form for inferring the reference data constructed for all training cases.

Findings from the idealized learning problem

To better understand the full learning problem and develop feature design guidelines for it, we compare the training and prediction performance of various Pr_t augmentations, employing different feature-space representations and features based on the viscosity ratio ν_t/ν . Specifically, we compare fully connected neural networks with piecewise multilinear maps (PMMs) for the feature-space representation and a simple feature (ν_t/ν) with a wrapped $(\frac{\nu}{\nu_t+\nu})$ and a logarithmed $(5 \tanh\left(\frac{\log_{10}(\frac{\nu_t}{\nu} + \varepsilon)}{0.75 \cdot 5}\right))$ feature. We refer the reader to Section 3.6.2 and Appendix B of Gupta and Duraisamy (2025) for a detailed description of PMMs, which are an efficient way of representing a nonlinear multivariate function by storing its values at discretized nodes in a closed input space.

We find that, for this specific problem, PMM augmentations are easier to train and can reduce the objective function in significantly fewer steps than neural networks. Additionally, for neural networks, random parameter initialization affects training consistency, convergence speed varies across trials, objective values do not always decrease monotonically, and final models behave inconsistently. These findings suggest that PMMs are more robust than neural networks, as their behavior remains consistent and predictable outside the training regime. In contrast, neural networks exhibit inconsistent extrapolation behavior. This can be seen in Figure 15.2, which shows the

| Feature | Lower Bound | Upper Bound | # Nodes |
|---|-------------|-------------|---------|
| $5 \tanh\left(\frac{\log_{10}(\nu_t/\nu + \varepsilon)}{0.75 \cdot 5}\right)$ | -5.0 | 5.0 | 41 |
| $1.5 \tanh\left(\frac{1}{1.5} \frac{-\nabla T \cdot \hat{n}}{\bar{T}/d_{\text{wall}}}\right)$ | -1.5 | 1.5 | 41 |

Table 15.2: Features used and PMM configuration for the preliminary model.

inferred $\Pr_t(\nu_t/\nu)$ -functions for all neural networks (a) and PMM models (b) using the logarithmed feature. Therefore, we choose PMMs as the feature-space representation for the preliminary model to predict the experimental data.

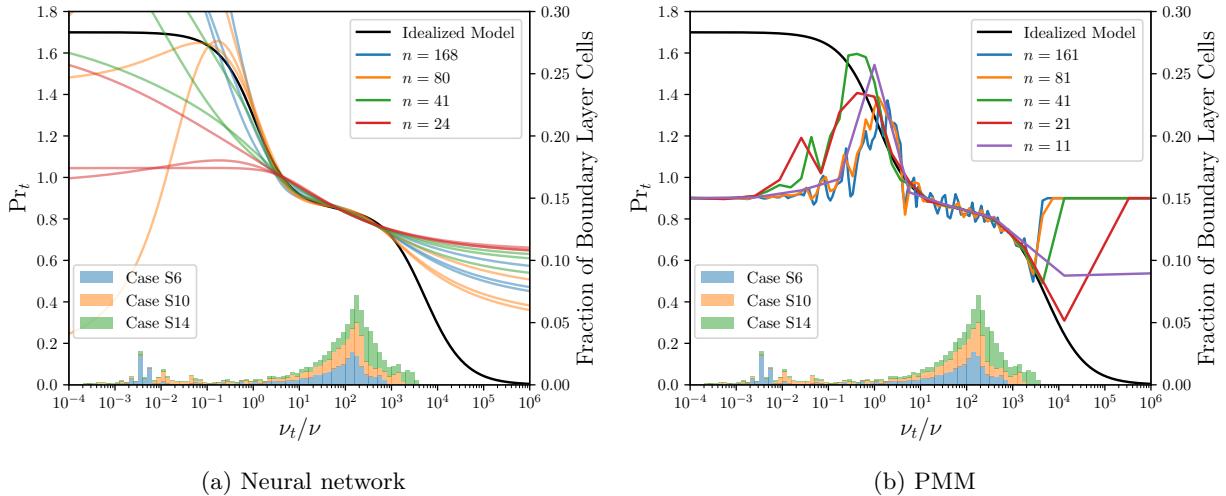


Figure 15.2: Inferred $\Pr_t(\nu_t/\nu)$ functions for augmentations using the logarithmed feature (number of model parameters n distinguished by colored lines) compared with the function of the idealized model along with the distribution of ν_t/ν -values in the boundary layer for each training case (distinguished by colored bars).

Preliminary model and prediction

Using the PMM as our feature-space representation, Table 15.2 describes the features and discretization of the preliminary augmentation. We use two features to gradually increase model complexity and gain insight on how interactions between features affect model performance. We prioritize using flow quantities that have the greatest significance for turbulent heat transfer and apply the process used to solve the idealized learning problem to analyze their values and obtain the final mathematical forms. The first feature is the same logarithmed feature used to recreate the idealized model. The second is based on the wall-normal temperature gradient, a quantity that has been shown to help distinguish between wall heating and cooling (Parish et al., 2024).

Figure 15.4 shows the predicted heat transfer distributions in terms of the Stanton number $St = \frac{q_w}{\rho_\infty u_\infty c_p (T_r - T_w)}$ for the training cases. We train our model on the Mach 5 oblique shock cases (Cases S6, S10, and S14), which are weighted equally. Figure 15.3 shows the training history of the preliminary model, and Figure 15.4 shows the predicted heat transfer distributions for the training cases.

The model successfully infers an improved heat transfer distribution. Quantitatively, the objective function is reduced by over two orders of magnitude for each case. However, there are

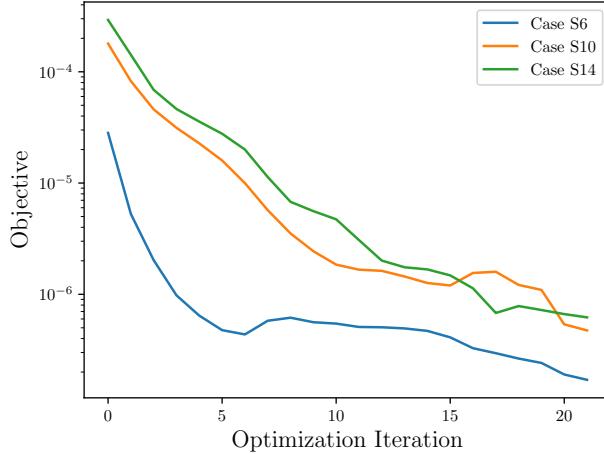


Figure 15.3: Objective training history of the preliminary model for inferring the reference data constructed for the Mach 5 oblique shock cases.

two limitations to this learned model: there is some oscillatory behavior in the predicted heat transfer profiles just downstream of the interaction locations, and the undisturbed boundary layer upstream of the interaction is modified. Since our reference data is designed to minimize augmentation influence in this region, such changes indicate that at least one additional feature is needed to differentiate between undisturbed and disturbed regions in the final model.

To assess the generalizability of this preliminary model, we test it on the $\alpha = 0.26$ forward-facing curved-wall case from Nicholson et al. (2024). It has an inflow Mach number of 4.9 and includes flow separation, similar to Cases S10 and S14. However, the geometry and shock interaction mechanism differ from the training cases. Instead of an externally generated oblique shock that impinges on a flat plate boundary layer and forms a reflected shock, it features a single shock induced by the wall geometry and a curved expansion corner. It also has a higher wall-to-recovery temperature ratio of 0.93 compared to 0.73 for the training cases. Thus, beyond evaluating geometric extrapolation, this test assesses the model's robustness when encountering unseen physical phenomena.

Figure 15.5 shows the predicted wall heat transfer distribution. The preliminary model improves heat transfer predictions, demonstrating the potential of our proposed methodology for developing turbulence augmentations. The existence of the expansion fan does not cause any spurious predictions, which can be attributed to the localized learning enabled by the PMM. However, like the training cases, there is a slight modification to the boundary layer upstream of the interaction. Future work will include investigating and adding more features to prevent this.

15.4 Application to smooth body separation prediction

Most one- and two-equation RANS turbulence models have also been developed in the context of simple shear flows without adverse pressure gradients (APGs). Therefore, for flows that exhibit smooth body separation due to APGs, they are unable to accurately predict the location and extent of the separation region (Tracey et al., 2013; Gao, 2014; Jespersen et al., 2016; Singh et al., 2017a). While several data-driven modeling frameworks have been successfully applied to improve RANS separation prediction (Ling et al., 2016; Wang et al., 2017; Wu et al., 2018; Singh et al., 2017b,a), they require improvements with respect to their interpretability and robustness during extrapolation to be used during engineering design and analysis. Since our proposed methodology seeks to address this, we apply it to this modeling problem with the goal of comparing the resulting model with those

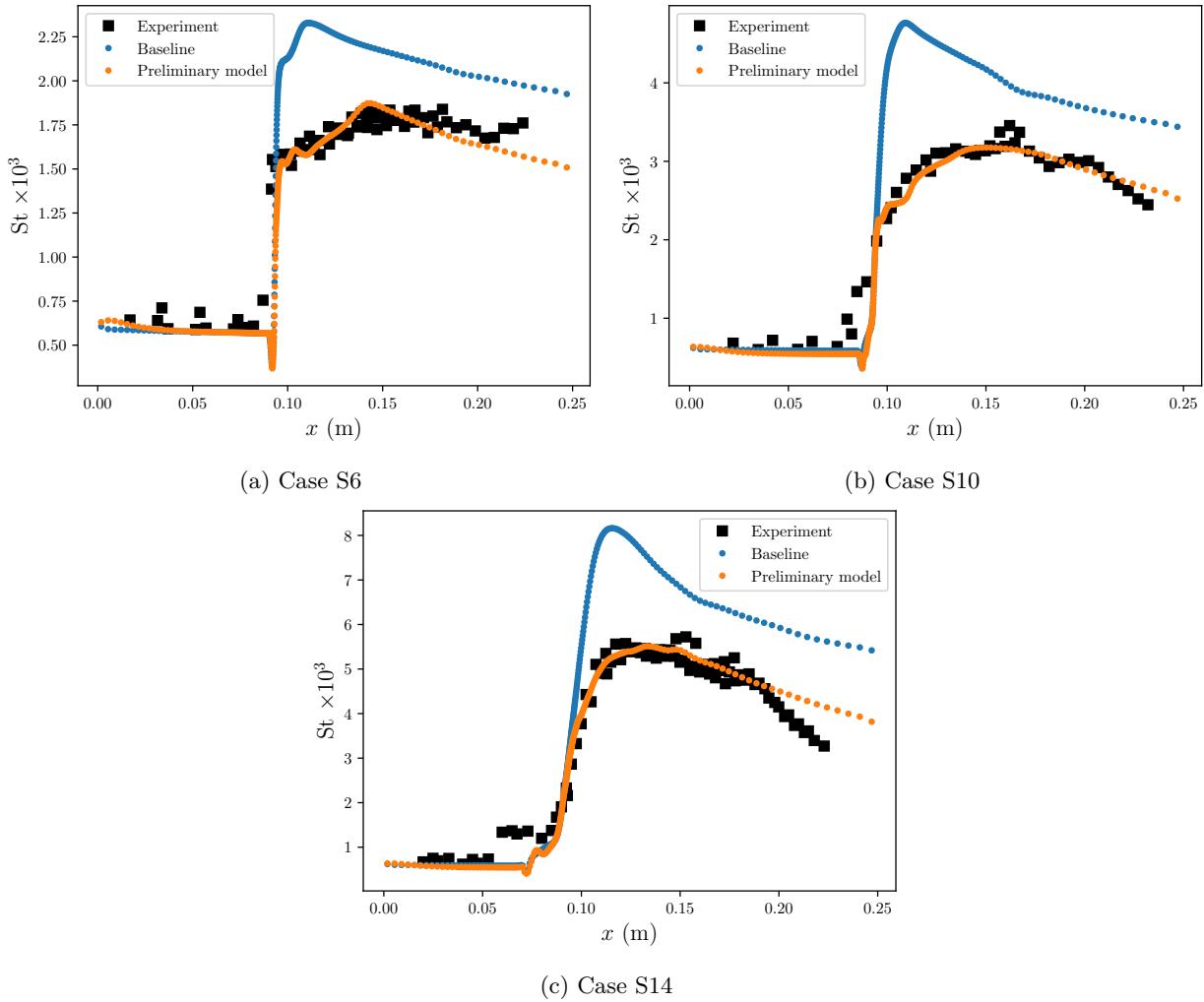


Figure 15.4: Heat transfer profiles of the preliminary model compared with the experimental data and baseline model for the training (Mach 5 oblique shock) cases.

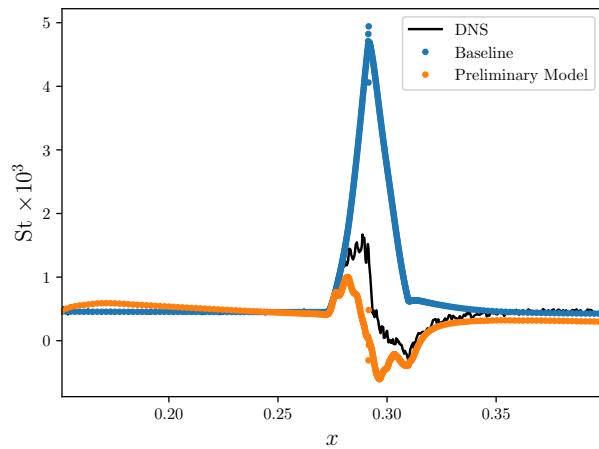


Figure 15.5: Predicted wall heat transfer distribution of the preliminary model compared with the baseline model and DNS results for the validation case.

existing in the literature and further refining it. This section outlines the preliminary work done during the reporting period to develop such a model that will be continued during the upcoming period.

Baseline model

The baseline model for this effort is the Wilcox 2006 $k - \omega$ turbulence model, given by

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho u_j k)}{\partial x_j} = P - \beta_W^* \rho k \omega + \frac{\partial}{\partial x_j} \left[\left(\mu + \sigma_k \frac{\rho k}{\omega} \right) \frac{\partial k}{\partial x_j} \right], \quad (15.10a)$$

$$\frac{\partial(\rho \omega)}{\partial t} + \frac{\partial(\rho u_j \omega)}{\partial x_j} = \frac{\gamma \omega}{k} P - \beta_W \rho \omega^2 + \frac{\partial}{\partial x_j} \left[\left(\mu + \sigma_\omega \frac{\rho k}{\omega} \right) \frac{\partial \omega}{\partial x_j} \right] + \frac{\rho \sigma_d}{\omega} \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j}. \quad (15.10b)$$

Constant definitions can be found in Wilcox (2008). Production is approximated by

$$P = \nu_t \bar{S}_{ij} \bar{S}_{ij}, \quad (15.11)$$

where \bar{S}_{ij} is the traceless rate-of-strain velocity tensor, to better align with the results posted on the NASA Turbulence Modeling resource.

Augmentation choice and model structure

We augment the production term in the ω -equation by directly multiplying the augmentation field to yield the modified term

$$\beta(\boldsymbol{\eta}(\tilde{\mathbf{u}}_m)) \frac{\gamma \omega}{k} P. \quad (15.12)$$

We choose to augment this term since it has shown prior success in improving separation prediction (Singh et al., 2017a). Furthermore, due to this success and subsonic turbulent separation being well-studied in the literature, we forgo the step of examining an idealized learning problem. Instead, we focus more on optimal feature design and selection to refine our proposed methodology.

Our chosen model structure is a PMM of the same form used in the hypersonic modeling effort. As discussed in Section 15.3, a PMM is chosen over a neural network since it can offer improved robustness and generalizability on unseen cases outside of the training regime.

Objective function and data

To assess model performance we compare the skin friction predicted by RANS to reference data using the objective function

$$\mathcal{J} = \sum_{i=1}^{n_x} \left(c_{f,m}(x_i) - c_{f,ref.}(x_i) \right)^2. \quad (15.13)$$

The augmentation model is trained on the experimental skin friction data for the “2D NASA Wall-mounted Hump Separated Flow Validation Case” (2DWMH) as provided on the NASA turbulence modeling website (Rumsey, 2021). To focus on learning the flow separation and reattachment downstream of the hump and due to concerns with the noise of the data along the hump, the reference data is a concatenation of the baseline turbulence model’s skin friction data upstream and along the hump ($x < 0.7$) and the experimental data downstream of this point.

For model validation, we use DNS data provided by Uzun and Malik (2022) for the two-dimensional Boeing speed bump with $Re_\infty = 2 \times 10^6$. While this case, like the 2DWMH training case, is also an incompressible boundary layer flow exhibiting smooth body separation downstream of a wall protrusion, it has two notable differences: the freestream Reynolds number is twice as large, and the boundary layer height is only a fraction of the bump height. In the training case, the boundary layer height is roughly the same as the bump height. Therefore, this case serves as an excellent test of the robustness and generalizability of the trained models.

Features examined

To determine the optimal input features to the augmentation model, we examine a set of seven features, which are listed in Table 15.3. The quantities used in the features are defined as

$$Re_t = \frac{\nu_t}{\nu}, \quad (15.14)$$

where ν and ν_t are the laminar and turbulent kinematic viscosities,

$$Re_\Omega = \frac{d_w^2 \Omega}{\nu}, \quad (15.15)$$

where d_w is the distance to the nearest wall and Ω is the vorticity magnitude,

$$Re_k = \frac{d_w \sqrt{k}}{\nu}, \quad (15.16)$$

where k is the turbulent kinetic energy,

$$r^* = \frac{S}{\Omega}, \quad (15.17)$$

where S is the strain-rate magnitude, and

$$Q_{cri_{sc}} = \frac{1 - r^{*2}}{1 + r^{*2}}. \quad (15.18)$$

Figure 15.6 shows contour plots for each of these features for the training case. As can be seen in the plots, features with similar motivations have differing abilities to highlight regions of interest in the flow.

Results

We perform two investigations on optimal feature design and selection. The first is to determine which triplet of features performs best during training and validation, and the second the impact of scaling a feature to change how it fills its output range.

Optimal feature triple

For this effort, six different feature triples are tested: (1,3,5), (2,3,5), (1,4a,5), (1,4a,6), (2,4a,5), (2,4a,6). Each triple is trained using 21^3 , 16^3 , and 11^3 parameters in the feature space to determine effect of feature-space discretization on performance. The optimization procedure for all augmentations is a modified form of the BFGS method in which a conservative step size is explicitly given instead of being determined using a line search for ease of implementation.

| # | Feature definition | Lower Bound | Upper Bound | Motivation |
|----|-----------------------------------|-------------|-------------|--|
| 1 | $\frac{Re_t}{1+Re_t}$ | 0 | 1 | Distinguish between boundary layer and freestream |
| 2 | $\frac{Re_t}{100+Re_t}$ | 0 | 1 | Distinguish between boundary layer and freestream, slower growth within boundary layer |
| 3 | $\frac{Re_t}{Re_t+0.01Re_\Omega}$ | 0 | 1 | Capture recirculation and non-equilibrium behavior |
| 4a | $\frac{Re_t}{Re_t+Re_k}$ | 0 | 1 | Capture recirculation and non-equilibrium behavior |
| 4b | $\frac{Re_t}{Re_t+0.01Re_k}$ | 0 | 1 | Capture recirculation and non-equilibrium behavior, fill feature's range better compared to 4a |
| 5 | $\frac{r^*}{1+r^*}$ | 0 | 1 | Distinguish between strain dominated and vorticity dominated regions |
| 6 | $\tanh(Q_{crisc})$ | -0.76 | 0.76 | Distinguish between strain dominated (positive) and vorticity dominated (negative) regions |

Table 15.3: List of features examined for the augmentation model.

Table 15.4 shows the final objective values obtained during training for all model configurations. The initial objective value of the baseline model is 1.98×10^{-7} . Figure 15.7 compares the inferred skin friction and pressure coefficients for each augmentation with $n = 16^3$ parameters with the experimental data and baseline model. Figure 15.8 compares their inferred augmentation fields. Across all feature combinations, several common trends are observed. Almost all models are able to qualitatively improve the predicted skin friction profiles significantly and match the size of the separation region and reattachment location. They also improve the pressure coefficient predictions downstream of the hump, but are unable to obtain the desired pressure in the separation region. All triples primarily focus on reducing the production of specific dissipation in the region of the flow just downstream of the separation point on the wall. None augment the turbulence model in the freestream or in the region of the boundary layer in the middle of the hump above the near-wall region. However, all modify the turbulence model in the undisturbed boundary layer upstream of the hump, indicating that at least one new feature is needed for all models to isolate this region.

Despite having similar wall distributions, Figure 15.8 reveals the differences between the feature triples. (2,3,5) and (2,4a,5) exhibit more regular and less oscillatory behavior that would make them more physically likely. (1,3,5) contains less modification in the undisturbed boundary layer upstream of the hump and contains more localized regions of significant deviation from the baseline model. Therefore, just from this examination of the training performance, it is ambiguous which triple is optimal.

Figure 15.9 compares the predicted skin friction and pressure coefficients for each augmentation

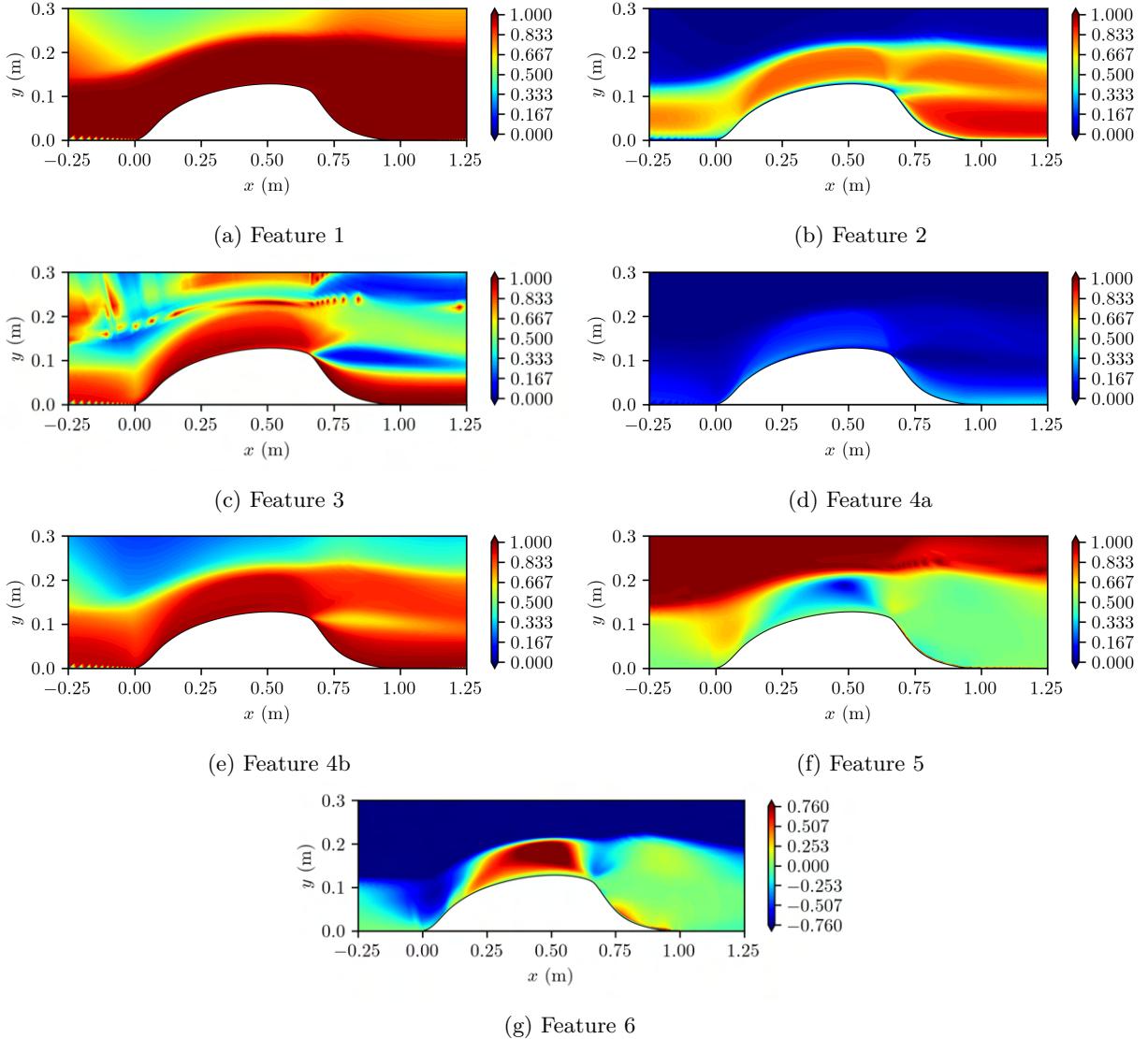


Figure 15.6: Contour plots of the examined features on the 2DWMH training case.

| | (1,3,5) | (2,3,5) | (1,4,5) | (1,4,6) | (2,4,5) | (2,4,6) |
|------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| $n = 21^3$ | 3.35×10^{-8} | 6.90×10^{-8} | 3.55×10^{-8} | 2.53×10^{-8} | 5.81×10^{-7} | 1.74×10^{-7} |
| $n = 16^3$ | 2.42×10^{-8} | 1.01×10^{-8} | 1.83×10^{-8} | 1.94×10^{-8} | 1.00×10^{-8} | 5.50×10^{-9} |
| $n = 11^3$ | 3.89×10^{-8} | 5.63×10^{-9} | 5.27×10^{-8} | 3.84×10^{-8} | 4.26×10^{-8} | 8.60×10^{-8} |

Table 15.4: Final objective function values obtained for all augmentations trained to infer the reference data constructed for the 2DWMH case. The initial objective value of the baseline model is 1.98×10^{-7} .

with $n = 16^3$ parameters with the DNS data and baseline model for the two-dimensional Boeing speed bump validation case. Figure 15.10 compares their predicted augmentation fields. Since the $n = 16^3$ discretization yielded the best results for the training case, it is used for the validation case. Unlike for the training case, the wall distributions vary significantly based on the feature triple used. Only three, triples (2,3,5), (1,4a,5), and (1,4a,6) are able to predict the skin friction reduction.

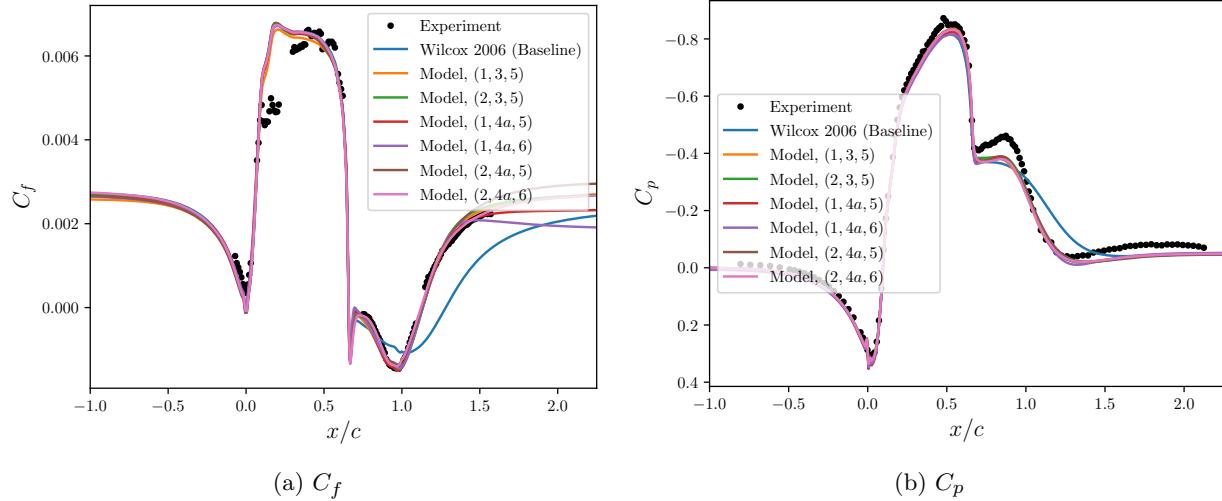


Figure 15.7: Inferred wall distributions for each feature triple using $n = 16^3$ parameters on the 2DWMH training case.

Among these, feature triple (1,4a,5) matches the size of the separation region and the undisturbed region downstream of the hump the closest. Its pressure coefficient prediction is comparable to that of the other two. Examining the augmentation fields reveals why these three feature triples perform the best on the validation case. They are the only ones able to predict a region of decreased turbulence dissipation just downstream of the separation point on the hump. Overall, they match the augmentation fields of the training case the closest. Furthermore, the triple (1,4a,5) has a smoother predicted augmentation field compared to the triple (1,4a,6). Therefore, based on its performance on both the training and validation cases, we determine that feature triple (1,4a,5) is the optimal feature triple based on the ones examined.

Optimal feature scaling

The strong performance of feature triple (1,4a,5) is somewhat surprising considering that Feature 4a is not optimally designed. As can be seen in Figure 15.6d, its effective output range on the training case does not fill the theoretical output range which is discretized in the PMM representation of the feature space. To determine the effect this has on performance, we also examine Feature 4b, which is the same feature except for a different constant value to better fill the output range. Figure 15.6e shows the increase in detail it provides about the flow. Therefore, it is expected that the triple (1,4b,5) will outperform (1,4a,5) for both the training and validation cases. To test this hypothesis, we train three more models using the feature triple (1,4b,5), one for each discretization level, and compare the results with those using the triple (1,4a,5).

Figure 15.11 compares the training history between the two sets of models, Figure 15.12 compares the wall skin friction and pressure coefficient distributions, and Figure 15.13 compares the inferred augmentation fields. The models using Feature 4b reach a lower objective value in the first ten steps due to the increased fidelity. The ability to capture more detail also results in increases in the objective value before the forward simulations converge, whereas Feature 4a results in sudden failure. For all configurations, the final skin friction profiles perform comparably, shown quantitatively by the training history plots and qualitatively by the wall distributions. The differences between the two feature sets can be seen in the augmentation fields. For the two higher discretization levels, the models using Feature 4b have more localized changes with fewer extreme

Development, Application, and Refinement of a New Data-Driven Turbulence Modeling Methodology

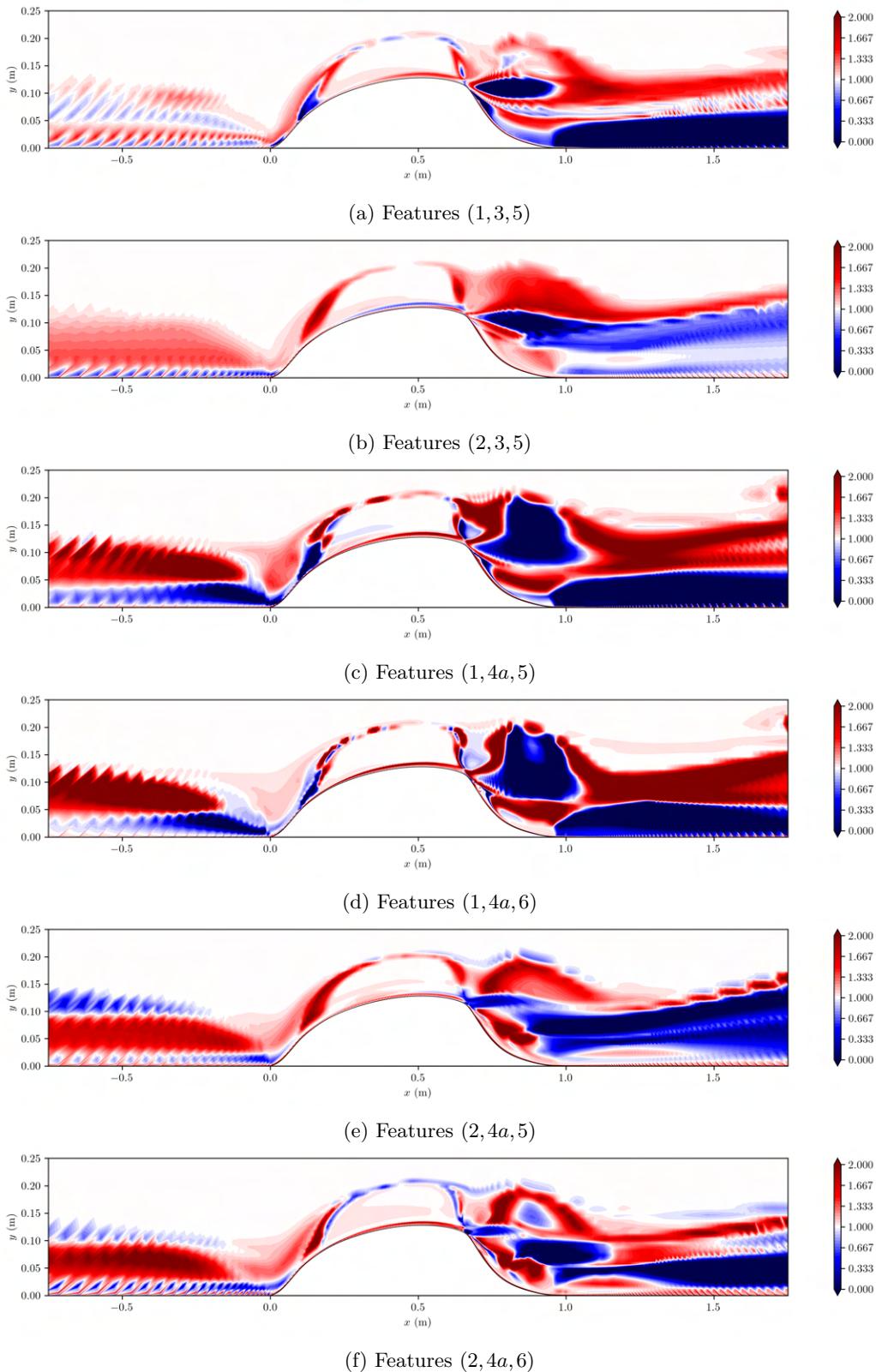


Figure 15.8: Inferred augmentation fields for each feature triple using $n = 16^3$ parameters on the 2DWMH training case.

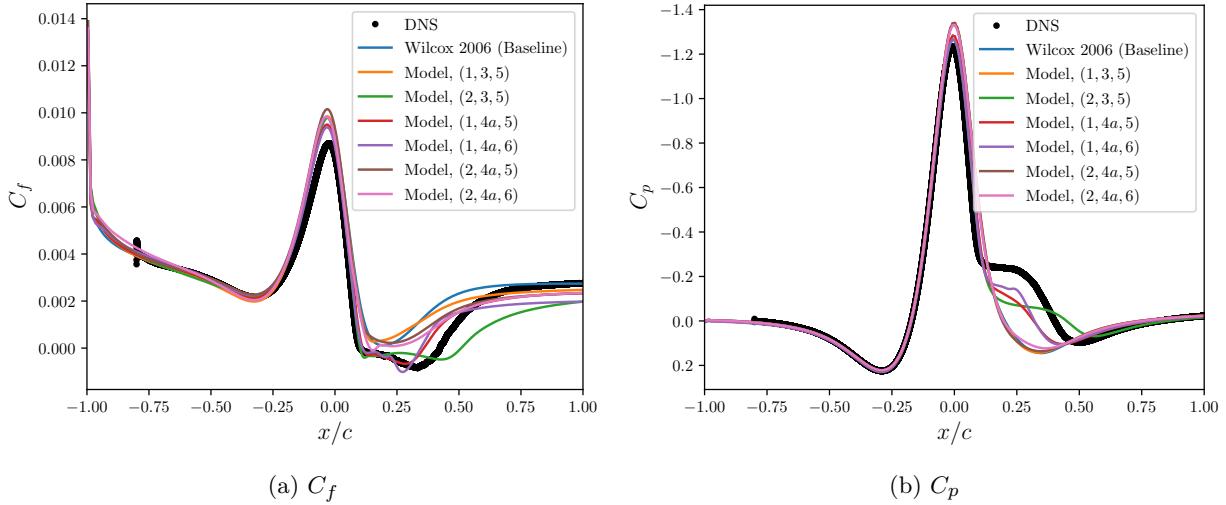


Figure 15.9: Inferred wall distributions for each feature triple using $n = 16^3$ parameters on the Boeing speed bump validation case.

values and modify the undisturbed boundary layer regions by a lesser degree. The reverse is true for the coarsest $n = 11^3$ discretization, but this can be attributed to the optimization procedure for feature triple (1, 4b, 5) lasting for three times as many steps. Regardless of the discretization level, all models using Feature 4b are able to identify the same region of decreased turbulence dissipation production just downstream of the hump. Therefore, due to the simplicity and physical realizability of the augmentation fields, Feature 4b appears to be more optimal than Feature 4a.

Figure 15.14 compares the predicted wall skin friction and pressure coefficient distributions for the $n = 16^3$ discretization level on the Boeing speed bump validation case, and Figure 15.15 compares the predicted augmentation fields. Both show that, unlike for the training case, Feature 4b degrades performance for the validation case. It results in only a minor improvement to the prediction of the separation region for both skin friction and the pressure coefficient. This can be attributed to the inability of the model using Feature 4b to adequately identify the region of decreased turbulence production, which can be seen in the augmentation fields. Unlike the model using the triple (1, 4a, 5), the degree of deviation from the baseline is less and the region of decreased turbulence production is smaller. This results in less change from the baseline in the wall distributions.

One reason for the decreased performance on the validation case could be that with the higher fidelity of Feature 4b, the model was able to overfit to the training data, reducing its generalizability. Feature 4a can be considered equivalent to using Feature 4b with a coarser discretization for the feature, and it is known that coarser discretizations increase the generalizability for PMMs (Srivastava and Duraisamy, 2021; Gupta and Duraisamy, 2025). Therefore, it is possible that there exists a discretization level where the using the triple (1, 4b, 5) yields better results than using (1, 4a, 5). This is something we plan to investigate in the upcoming year. Regardless, it is clear that having more detailed features does not always result in improved model performance, so more sophisticated feature design processes and metrics are required.

Development, Application, and Refinement of a New Data-Driven Turbulence Modeling Methodology

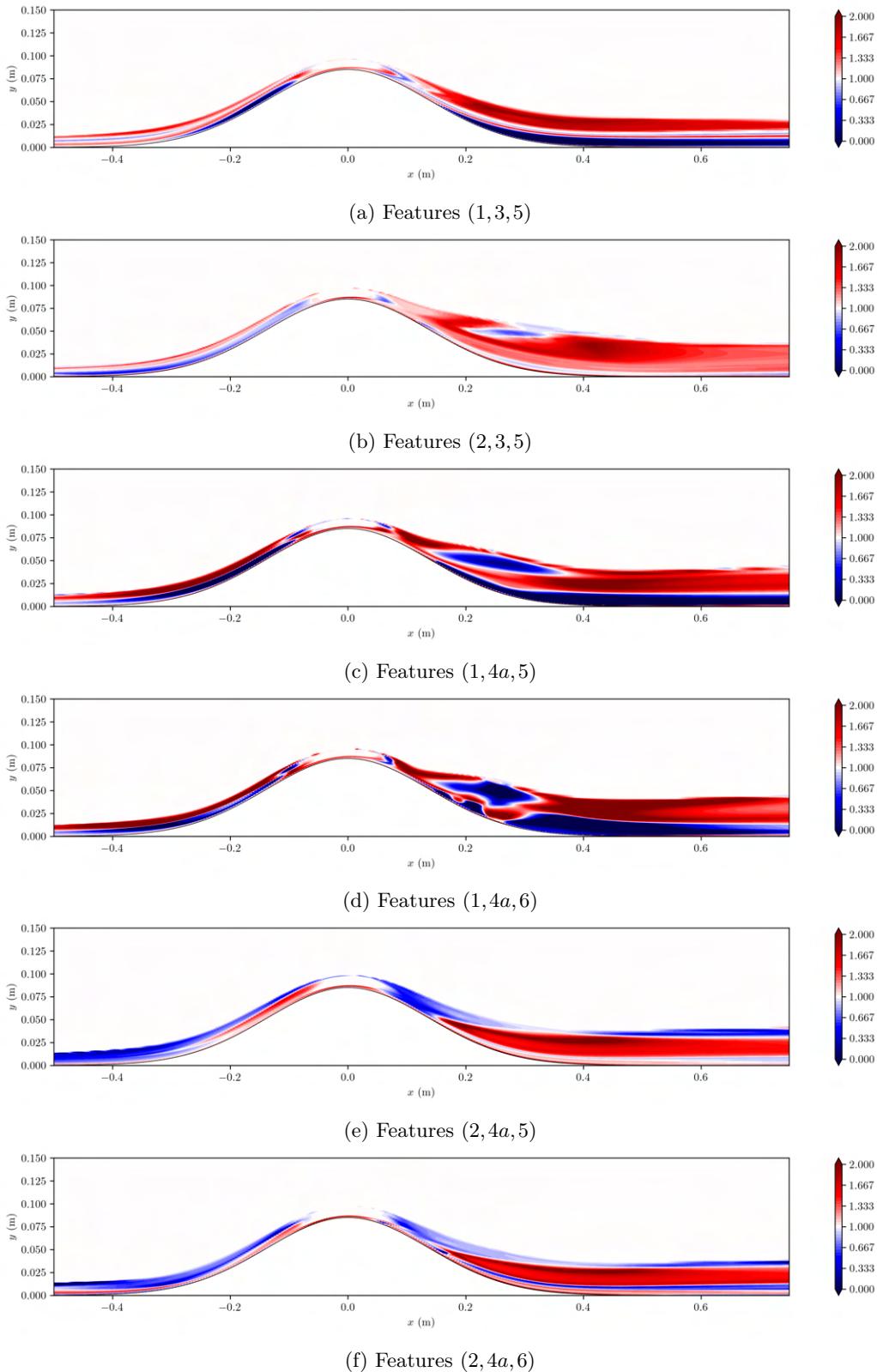


Figure 15.10: Inferred augmentation fields for each feature triple using $n = 16^3$ parameters on the Boeing speed bump validation case.

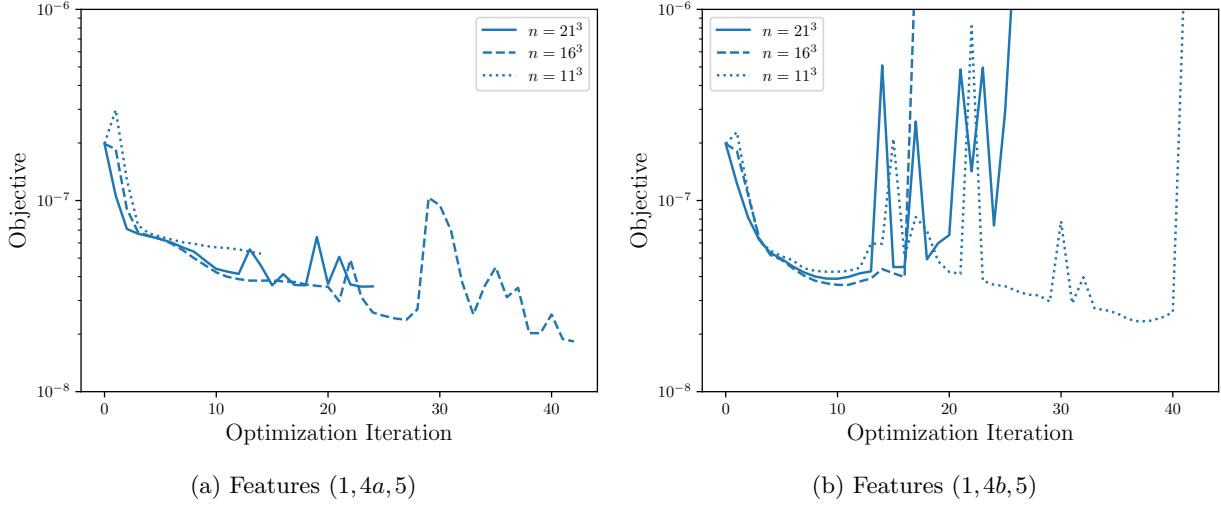


Figure 15.11: Comparison of training history between models using input features (1, 4a, 5) and models using (1, 4b, 5) for all discretization levels for inferring skin friction data for the 2DWMH case.

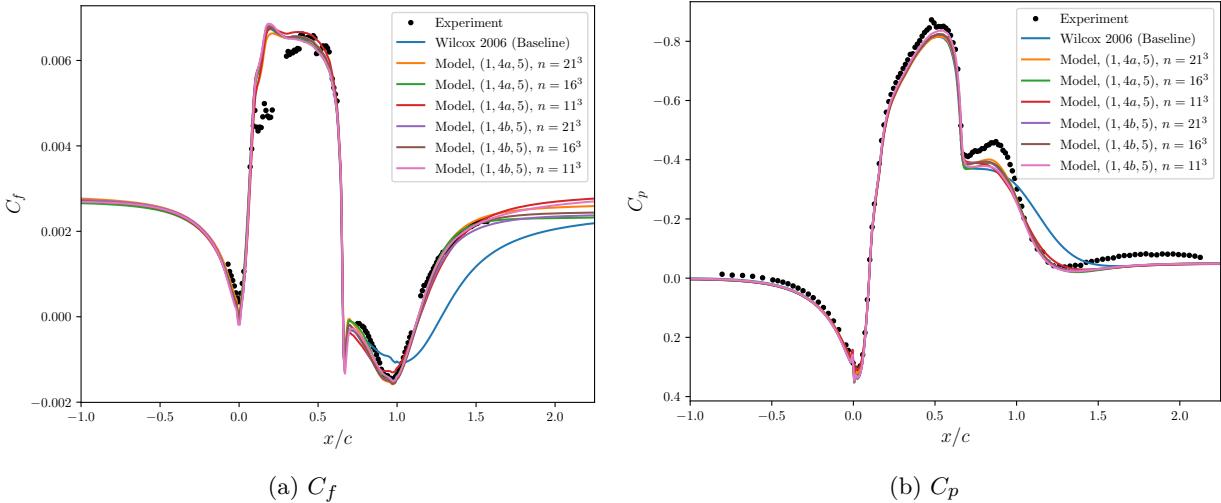


Figure 15.12: Inferred wall distributions for each feature triple using $n = 16^3$ parameters on the 2DWMH training case.

15.5 Conclusions

This report describes the work done over the reporting period to develop, apply, and improve a systematic methodology for developing data-driven turbulence model augmentations. This methodology is developed with the goal of examining modeling choices and implementation details often overlooked in existing literature. The steps of the methodology can be summarized as: 1) identify an augmentation by analyzing critical assumptions in the low-fidelity model, 2) evaluate the viability of the augmentation through field inversion, 3) devise and solve an idealized learning problem to better understand the impact of modeling decisions, and 4) incrementally improve the model by progressively adding features.

We outline how this methodology is developed by summarizing the modeling effort on turbulent heat transfer for hypersonic flows. We identify that the constant turbulent Prandtl number assumption is a key source of error for hypersonic flows and show, through field inversion, that

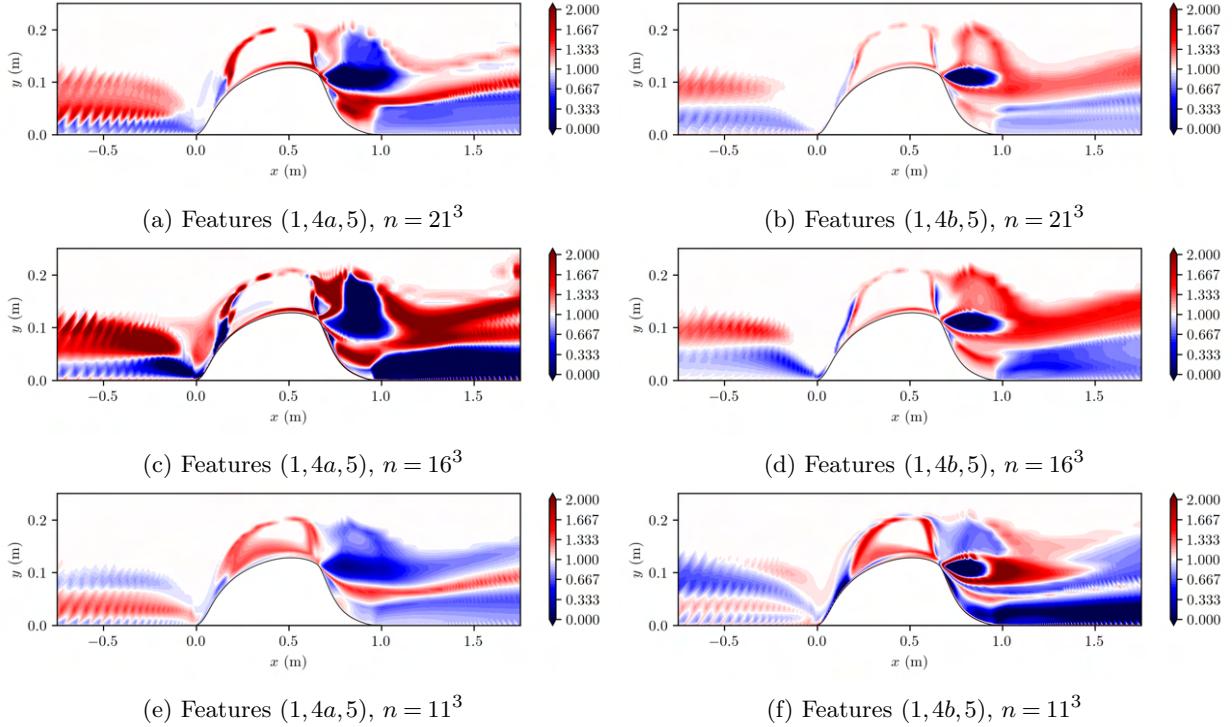


Figure 15.13: Inferred augmentation fields for feature triple (1,4a,5) compared with (1,4b,5) for each discretization level on the 2DWMH training case.

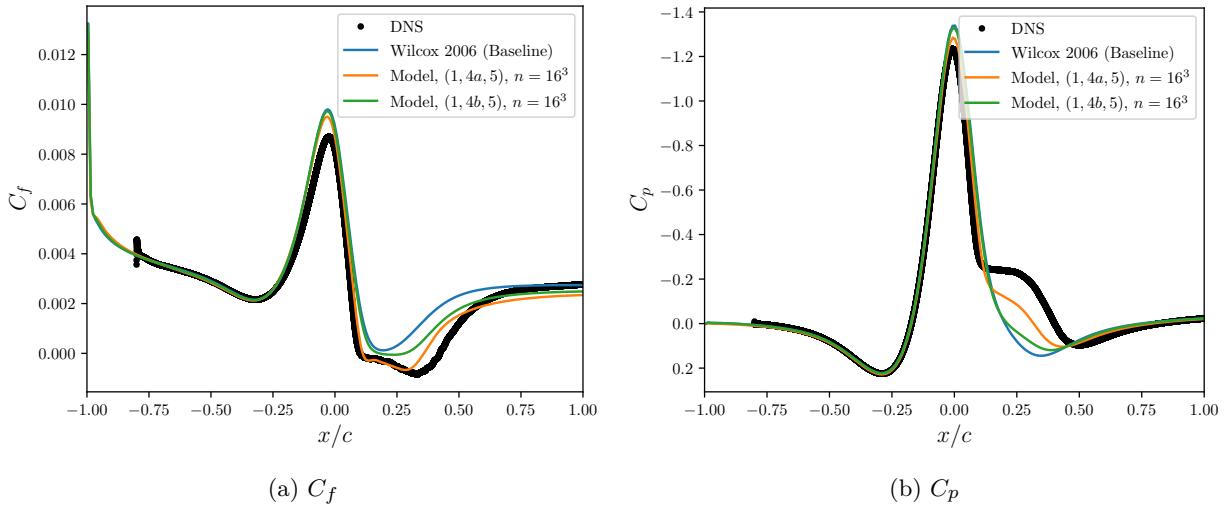


Figure 15.14: Inferred wall distributions for each feature triple using $n = 16^3$ parameters on the Boeing speed bump validation case.

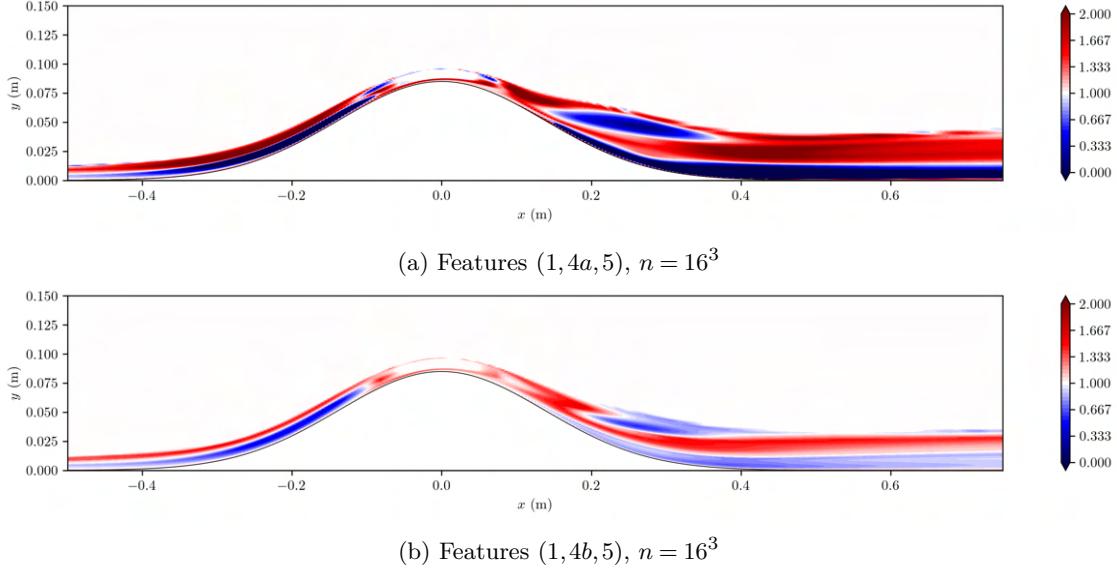


Figure 15.15: Inferred augmentation fields for feature triple (1,4a,5) compared with (1,4b,5) for each discretization level on the Boeing speed bump validation case.

a spatially varying model can address the overprediction in heat transfer due to shock-boundary layer interactions. We devise an idealized learning problem to learn an existing variable turbulent Prandtl number relation and use it to assess the impact of the chosen feature-space representation and input feature form on augmentation performance. We demonstrate that piecewise multilinear maps, over neural networks, can match target behaviors more accurately, require fewer parameters, and remain robust when extrapolating beyond training data. With this knowledge, we develop a two-feature preliminary model with input features based on the scaled laminar-to-turbulent viscosity ratio and the wall-normal temperature gradient. It successfully trains on experimental data for a Mach 5 oblique-shock impinging flow and generalizes well to a more complex geometry with a different shock mechanism, providing physically realizable fields while maintaining accurate predictions. However, several improvements are still required for the augmentation to reduce its region of influence to only regions affected by shock interactions and generate more accurate predictions.

We then apply this methodology to a different modeling problem of improving smooth body separation prediction, focusing mostly on the feature design process. We choose to augment the production term for turbulence dissipation in the Wilcox 2006 $k-\omega$ model for its known ability to improve separation prediction. We analyze a set of seven features to determine what characteristics an optimal combination of features has and how scaling of an individual feature affects augmentation performance. We determine that, while some insight can be gained from analyzing the quality of augmentation fields on the training cases, augmentation behavior on validation cases provides the most information. Performant feature combinations yield augmentations that are able to identify similar flow features across the training and validation cases and have smooth behavior in the augmentation field for the validation cases. For feature scaling, while a more suitably scaled feature that can provide higher fidelity leads to improved performance during training, it can lead to overfitting and worse performance on the validation case despite better theoretical capabilities.

We plan to focus on improving the feature design process and developing more flexible feature-space representations. Specifically, we will develop quantitative metrics, likely based on information theory, to compare (combinations of) features. Such metrics should provide explanations for the findings on optimal feature combinations discussed in this report. We also plan to develop an

adaptive piecewise multilinear map implementation that can consider the availability of training data in the feature space to strike a better balance between coarse discretizations for generalizability and fine discretizations for training accuracy.

References

E. Barnard and L. F. A. Wessels. Extrapolation and interpolation in neural network classifiers. *IEEE Control Systems Magazine*, 12(5):50–53, 1992.

G. T. Coleman and J. L. Stollery. Heat transfer from hypersonic turbulent flow at a wedge compression corner. *Journal of Fluid Mechanics*, 56(4):741–752, 1972.

L. Duan, I. Beekman, and M. P. Martín. Direct numerical simulation of hypersonic turbulent boundary layers. part 3. effect of Mach number. *Journal of Fluid Mechanics*, 672:245–267, 2011.

K. Duraisamy. Perspectives on machine learning-augmented Reynolds-averaged and large eddy simulation models of turbulence. *Physical Review Fluids*, 6:050504, 2021.

K. Duraisamy, G. Iaccarino, and H. Xiao. Turbulence modeling in the age of data. *Annual Review of Fluid Mechanics*, 51:357–377, 2019.

Y. Fang, Y. Zhao, F. Waschkowski, A. S. H. Ooi, and R. D. Sandberg. Toward more general turbulence models via multicase computational-fluid-dynamics-driven training. *AIAA Journal*, 61(5):2100–2115, 2023.

A. Ferrero, A. Iollo, and F. Larocca. Field inversion for data-augmented RANS modelling in turbomachinery flows. *Computers & Fluids*, 201:104474, 2020.

F. Gao. *Advanced Numerical Simulation of Corner Separation in a Linear Compressor Cascade*. Theses, Ecole Centrale de Lyon, 2014.

M. B. Giles and N. A. Pierce. An introduction to the adjoint approach to design. *Flow, Turbulence and Combustion*, 65(3):393–415, 2000.

M. B. Giles, M. C. Duta, J.-D. Muller, and N. A. Pierce. Algorithm developments for discrete adjoint methods. *AIAA Journal*, 41(2):198–205, 2003.

N. Gupta and K. Duraisamy. Computational and physical considerations for the development of machine learning augmented turbulence models. *International Journal of Heat and Fluid Flow*, 2025. Under Review.

J. R. Holland, J. D. Baeder, and K. Duraisamy. *Field Inversion and Machine Learning With Embedded Neural Networks: Physics-Consistent Neural Network Training*. 2019.

C. C. Horstman and F. K. Owen. Turbulent properties of a compressible boundary layer. *AIAA Journal*, 10(11):1418–1424, 1972.

D. C. Jespersen, T. H. Pulliam, and M. L. Childs. Overflow turbulence modeling resource validation results. Technical report, NASA, 2016. NAS-2016-01.

W. M. Kays and M. E. Crawford. *Convective Heat and Mass Transfer*. McGraw-Hill, 3rd edition, 1993.

J. N. Kutz. Deep learning in fluid dynamics. *Journal of Fluid Mechanics*, 814:1–4, 2017.

J. Ling, A. Kurzawski, and J. Templeton. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, 807:155–166, 2016.

G. L. Nicholson, L. Duan, and N. J. Bisek. Direct numerical simulation database of high-speed flow over parameterized curved walls. *AIAA Journal*, 62(6):2095–2118, 2024.

E. Parish, D. S. Ching, C. Jordan, G. Nicholson, N. E. Miller, S. Beresh, M. Barone, N. Gupta, and K. Duraisamy. Data-driven turbulent Prandtl number modeling for hypersonic shock–boundary-layer interactions. *AIAA Journal*, 63(5):1671–1692, 2024.

References

E. J. Parish and K. Duraisamy. A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics*, 305:758–774, 2016.

S. Priebe and M. P. Martín. Turbulence in a hypersonic compression ramp flow. *Physical Review Fluids*, 6:034601, 2021.

P. Raje, E. Parish, J.-P. Hickey, P. Cinnella, and K. Duraisamy. Recent developments and research needs in turbulence modeling of hypersonic flows. *Physics of Fluids*, 37(3):031304, 2025.

J. C. Rotta. Heat transfer and temperature distribution in turbulent boundary layers at supersonic and hypersonic flow. *AGARDograph*, 97:41–63, 1965.

C. Rumsey. 2D NASA wall-mounted hump separated flow validation case. NASA Langley Research Center Turbulence Modeling Resource, 2021.

E. Schülein. Skin friction and heat flux measurements in shock/boundary layer interaction flows. *AIAA Journal*, 44(8):1732–1741, 2006.

A. P. Singh and K. Duraisamy. Using field inversion to quantify functional errors in turbulence closures. *Physics of Fluids*, 28(4):045110, 2016.

A. P. Singh, R. Matai, A. Mishra, K. Duraisamy, and P. A. Durbin. Data-driven augmentation of turbulence models for adverse pressure gradient flows. In *23rd AIAA Computational Fluid Dynamics Conference*, 2017a. AIAA 2017-3626.

A. P. Singh, S. Medida, and K. Duraisamy. Machine-learning-augmented predictive modeling of turbulent separated flows over airfoils. *AIAA Journal*, 55(7):2215–2227, 2017b.

J. Sirignano, J. F. MacArt, and J. B. Freund. DPM: A deep learning PDE augmentation method with application to large-eddy simulation. *Journal of Computational Physics*, 423:109811, 2020.

A. J. Smits and J.-P. Dussauge. *Turbulent shear layers in supersonic flow*. Springer Science & Business Media, 2006.

P. R. Spalart. Philosophies and fallacies in turbulence modeling. *Progress in Aerospace Sciences*, 74:1–15, 2015.

V. Srivastava and K. Duraisamy. Generalizable physics-constrained modeling using learning and inference assisted by feature-space engineering. *Physical Review Fluids*, 6:124602, 2021.

B. Tracey, K. Duraisamy, and J. Alonso. *Application of Supervised Learning to Quantify Uncertainties in Turbulence and Combustion Modeling*. 2013.

A. Uzun and M. R. Malik. High-fidelity simulation of turbulent flow past Gaussian bump. *AIAA Journal*, 60(4):2130–2149, 2022.

J.-X. Wang, J.-L. Wu, and H. Xiao. Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on dns data. *Physical Review Fluids*, 2:034603, 2017.

D. C. Wilcox. Formulation of the $k-\omega$ turbulence model revisited. *AIAA Journal*, 46(11):2823–2838, 2008.

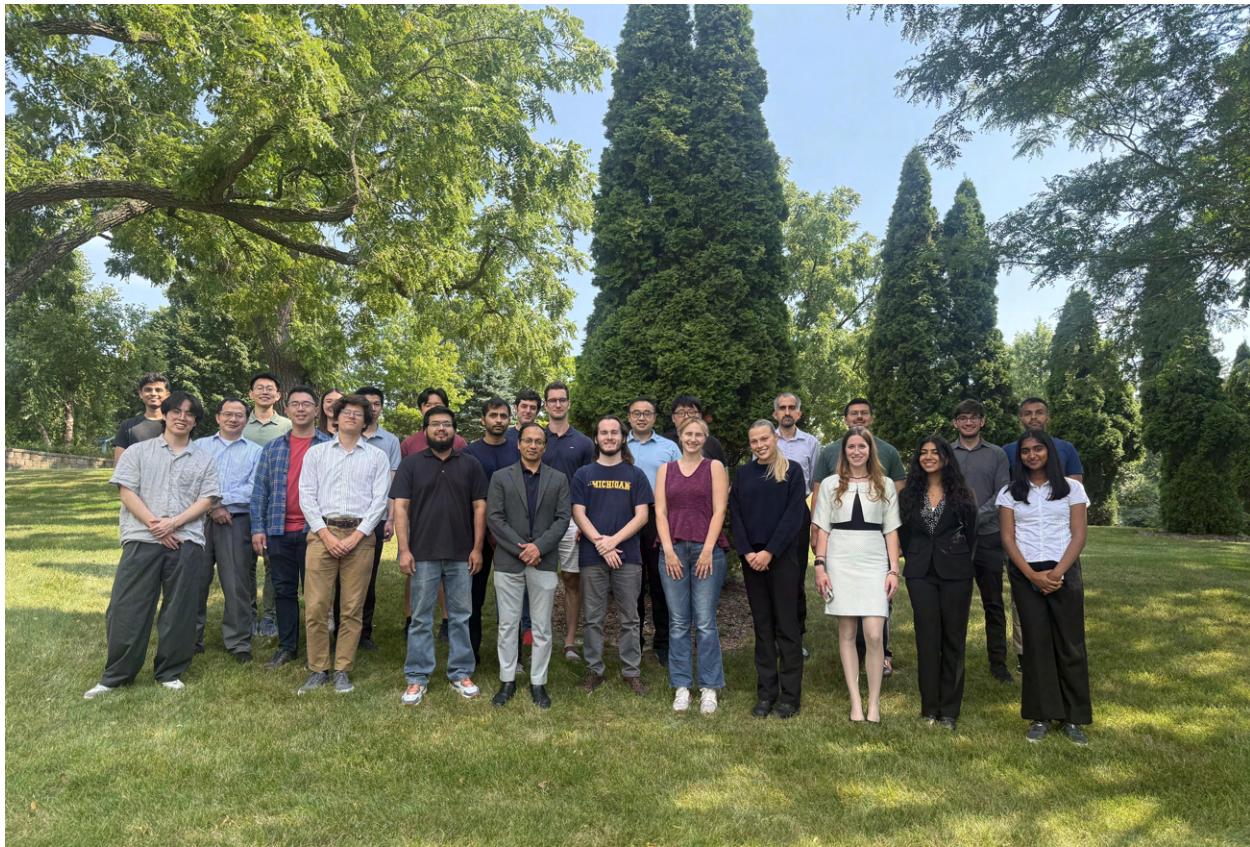
J.-L. Wu, H. Xiao, and E. Paterson. Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework. *Physical Review Fluids*, 3:074602, 2018.

M. Yang and Z. Xiao. Improving the $k-\omega-\gamma$ -ar transition model by the field inversion and machine learning framework. *Physics of Fluids*, 32(6):064101, 2020.

M. Zhu, H. Zhang, A. Jiao, G. E. Karniadakis, and L. Lu. Reliable extrapolation of deep neural operators informed by physics or sparse observations. *Computer Methods in Applied Mechanics and Engineering*, 412:116064, 2023.

CNRE Undergraduate Summer Research Experience 2025

This article summarizes the research performed during the 2025 Undergraduate Summer Research Program that extended from May 27 to August 7, 2025. A total of 13 undergraduate students worked with postdoctoral researcher, senior graduate student and faculty mentors, and Navy collaborators on a range of topics that are detailed below. The entire team was made up of members of the Mechanical, Electrical, Aerospace, NA&ME, Computer Science, and Math departments. The undergraduate students were educated on cutting-edge naval and marine problems and taught to conduct research using state of the art approaches. An exciting ten weeks culminated in a well-received final presentation workshop.



A group picture after the Final Presentation Workshop on August 7, 2025.

The projects are summarized below in alphabetical order based on the last names of the undergraduate researchers. Superhydrophobic surfaces offer the potential to reduce drag, and a recently developed drag model was combined with formal optimization methods to seek superhydrophobic

surfaces with good drag-reducing properties (section 16.1). A custom High-Area Rapid Printing (HARP) system tailored for grayscale, layerless hydrogel printing was studied (section 16.2) as an approach to engineering synthetic soft materials relevant to biofouling. A pneumatic actuation system was improved upon, to enable the rapid control of smart morphable surfaces used for flow control (section 16.3). Numerical simulations were performed (section 16.4) to simulate the growth and collapse of bubbles in cavitation susceptibility meter-relevant geometries. An experimental study on pendulum-based wave energy converters explored (section 16.5) counterweight and negative stiffness mechanisms for resonance tuning. Different synthetic materials were evaluated for future biofouling studies and measurement software for the Skin Friction Flow Facility (SF3) was developed (section 16.6). A recently developed cavitation inception model was applied (section 16.7) to simulate experiments on round jets. The PINN approach was used to compute the pressure at the center of vortices subject to velocity measurements from particles near the outer periphery of the vortex (section 16.8). A data buoy prototype capable of autonomously harvesting both wave and wind energy was tested (section 16.9) for future application to ocean observation. A theoretical framework for modeling electromagnetic scattering from arbitrary, random, rough surfaces was employed (section 16.10) to model scattering from the ocean surface. Deep reinforcement learning (DRL) and its viability for hydrofoil/propeller multi-disciplinary optimization was studied (section 16.11). A framework to streamline and automate differentiation within Monte Carlo simulations was explored (section 16.12) for its viability for systems such as phase transitions in batteries. The performance of superhydrophobic surfaces found in nature was studied by constructing accurate three-dimensional representations of their microstructures and by computing how liquid interfaces interact with them (section 16.13).

16.1 Gradient Based Optimization of Superhydrophobic Surface Geometry for Drag Reduction

Veer Agarwal, Gao Jun Wu and Krishnan Mahesh

Due to chemical properties and roughness designs at the nano-to-micro scales, superhydrophobic surfaces (SHS) are capable of reducing the drag force on vehicles moving through water. This project constructs a gradient-based constrained optimization framework to find the optimal roughness design of a SHS for drag reduction. The framework is based on a semi-empirical model relating the friction coefficient of SHS to the bulk Reynolds number ranging between 5000 and 50000. The model is calibrated from friction coefficient data of SHS collected in channel flow experiments. The dependent parameters of the model are six different roughness measures of SHS, including Wenzel roughness, RMS of the roughness height and the slope, autocorrelation length, the difference between the mean peak-to-valley height and the mean height, and the projected area normal to the flow direction. The optimization treats these surface roughness characteristics as free design parameters and seeks an optimal combination that minimizes the friction coefficient. Inequality constraints are put in place to keep the surface parameters within a manufacturable range.

The optimization program is implemented in MATLAB and uses its built-in quasi-Newton algorithm to obtain a numerical solution of the optimal surface parameters. Two different formulations are used as the objective function: maximizing the drag reduction at a single Reynolds number and maximizing the drag reduction integrated with respect to Reynolds number across the range of interest.

Both formulations of the objective function find optimal surface parameters that result in a drag reduction of at least 18% compared to a regular smooth surface across the entire range of

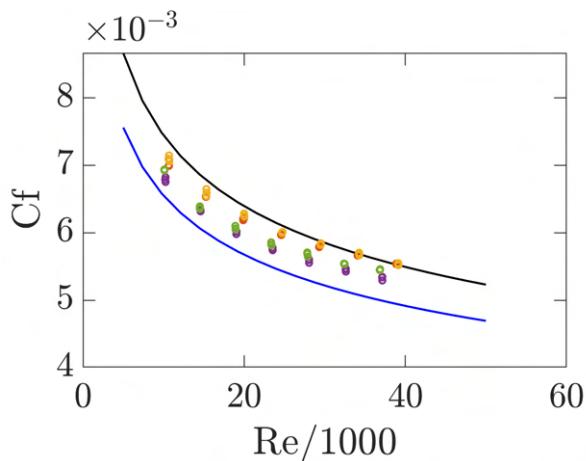


Figure 16.1: C_f with respect to Re_H for a regular smooth surface (black line), baseline SHS (markers), and the optimized SHS (blue line).

Reynolds numbers. The single Reynolds number objective is observed to be more prone to sub-optimal results and thus requires more iterations of the optimization runs at each design Reynolds number. The drag reduction generally decreases as the Reynolds number increases. As opposed to there being a single unique combination, many different combinations of roughness parameters can yield optimal drag reduction. The three most sensitive parameters are - Wenzel roughness, RMS of the roughness slope, and difference between peak-to-valley height and the mean height.

16.2 Towards inverse methods for heterogeneous soft materials: grayscale polymerization of layerless hydrogels

Joseph Beckett, Aryan Shah, Teagan Fynan, Carl Thrasher and Jon Estrada

While vat photopolymerization (VP) 3D printing is increasingly used for precise fabrication of complex polymeric parts, its application for engineering mechanically gradient soft materials is restricted by a lack of standardization and validation procedures. As a result, existing uses of VP have been largely limited to non-critical applications in contrast to metal 3D printing where in-situ process monitoring, quality-controlled feedstocks, and postprocess inspection are established. The lack of validation methods for soft material prints is largely because no single characterization method has emerged as effective for robustly quantifying spatially varying properties. A key limitation of emerging inverse methods for heterogeneous samples is the lack of reliable benchmark materials. Without well-characterized gradient specimens, it is difficult to validate these techniques, yet the inverse methods themselves are needed to produce and confirm these benchmarks. Inverse methods for mechanical property identification traditionally assume homogeneous materials, further complicating efforts to benchmark or validate emerging tools.

To address these challenges, we developed a custom High-Area Rapid Printing (HARP) system tailored for grayscale, layerless hydrogel printing (Figure 16.2). HARP presents advantages over other continuous VP technologies—such as enhanced cooling and resin flow—making it suitable for fabricating specimens with continuous mechanical gradients. The custom printer features a borosilicate glass vat, precision optics, and stepper-motor-driven motion, all coordinated by a microcontroller. Resin formulation was also novel, utilizing high water content and glycerol to tune viscosity, yielding elastic moduli in the 1–10 kPa range. Initial development focused on optimizing

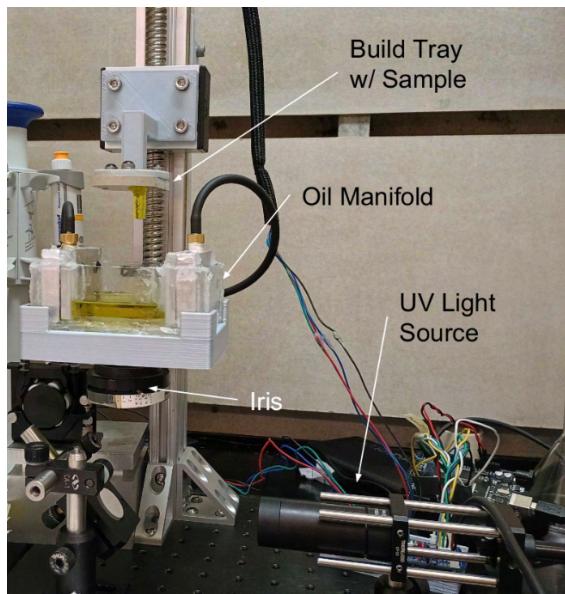


Figure 16.2: 4D "HARP" printer setup.

both the printer and hydrogel resin to eliminate typical VP defects, e.g., layering, cracking at edges, and non-uniform curing. Adjusting print speeds, resin viscosity, and curing parameters permitted successful layerless prints. The optical system's intensity profile was characterized to further refine print accuracy, with additional improvements facilitated by a new precision digital micromirror device (DMD) for grayscale patterning. Mechanical testing of printed hydrogels showed repeatability across specimens and correlation between UV dosage and modulus. Preliminary characterization demonstrated that moduli could be modulated via print parameters. Ongoing work includes integration of the DMD for finer grayscale control, further resin optimization, and extensive validation with localized and volumetric measurements.

16.3 Improvement of a Pneumatic Actuation System for Rapid Control of Smart Morphable Surfaces

Manasa Gollapalli, Sayre Satterwhite and Anchal Sareen

Many aquatic swimmers and avian fliers adaptively morph their fins or wings, enhancing efficiency, maneuverability, stealth, agility, and resistance to biofouling. Inspired by these natural strategies, we developed smart morphable surfaces that dynamically adjust surface texture in response to changing flow conditions, enabling drag reduction over a 3D bluff body or lift generation for maneuvering. However, two key limitations constrain the current approach. First, while dimples can actuate within 1 second, the return to a smooth state requires 10 seconds, limiting applicability in scenarios demanding rapid transitions. Second, to suppress flutter in flow, the skin must be pre-stretched, introducing a baseline micro-dimpled texture instead of a truly smooth reference surface. The objective of the summer undergraduate research project was to address these two challenges.

To overcome these limitations, we designed a new pneumatic system (Figure 16.3) with one pressure line (400 kPa) and one vacuum line (50 kPa). The dual-line setup enables both equalization of internal and external hydrostatic pressure and rapid depressurization/repressurization of the body. The vacuum line is generated via a venturi connected to the pressure line and is supported

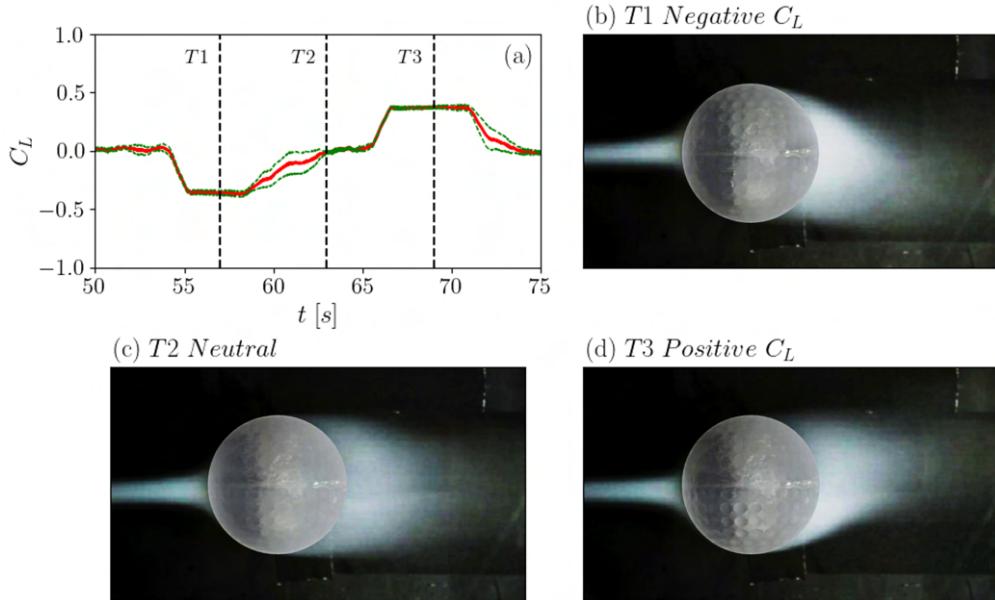


Figure 16.3: Wake deflection and alternating lift generation on a sphere via smart surface morphing. (a) Lift coefficient over one cycle. Smoke visualization at (a) maximum positive lift, (b) zero lift, and (c) maximum negative lift. Smart surface morphing enables controlled wake deflection, producing alternating side forces that could be leveraged for maneuvering.

by 1 L of depressurized air tanks to provide sufficient capacity for air release. Each line is controlled by a solenoid valve that can open or close to decrease, increase, or maintain pressure. A pressure sensor provides feedback to the controller, which governs the solenoid valve states.

Using this system, the repressurization time of a test body was reduced dramatically, from more than 8 seconds to less than 1 second. This improvement demonstrates the feasibility of rapid surface morphing. Nonetheless, challenges remain: solenoids could not actuate quickly enough to prevent excess air intake, which led to membrane delamination, and chamber-to-chamber variations in membrane properties required different pressures to achieve uniform dimple depth. Future work will focus on accelerating actuation speed, refining the controller to avoid overpressurization, and developing more consistent membrane materials to ensure reliable performance across all chambers.

16.4 Cavitation-bubble dynamics modeling with application to cavitation susceptibility meters

Katarina Jevtic, Eleanor Anderson-Zych and Eric Johnsen

Cavitation is the phenomenon whereby vapor bubbles form, typically from a gas nucleus, and oscillate due to pressure changes in liquid flows. This process may lead to noise, vibration, reduced efficiency, and even structural damage. Predicting the inception of cavitation is critical to a variety of applications. Cavitation susceptibility meters (CSMs) have been proposed to measure nucleus sizes. Briefly, these devices consist of a nozzle with converging and diverging sections; given a sample of the flowing liquid of interest, the device determines whether nuclei are present and cavitate. However, nucleus sizes and concentrations cannot readily be determined.

Our objective was to use numerical simulations to simulate bubble growth and collapse through CSM-relevant geometries to determine nucleus sizes expected to cavitate during passage through

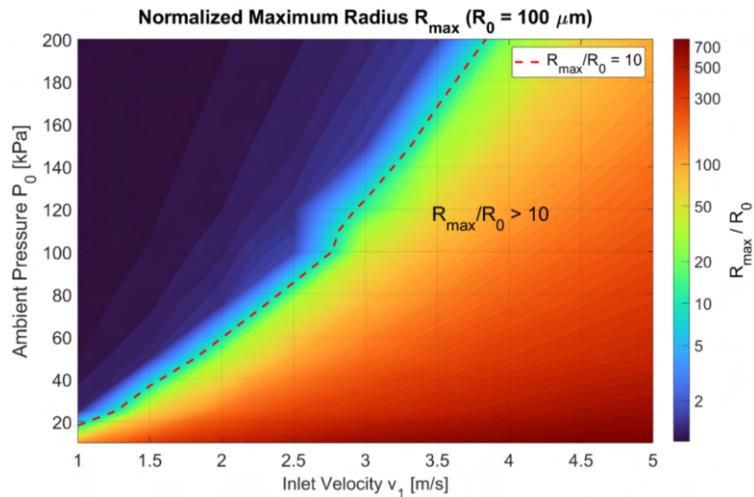


Figure 16.4: Simulated normalized maximum radius as a function of inlet velocity and ambient pressure.

the CSM. From the geometry and given flow conditions (ambient pressure, inlet velocity), the spatial variation of pressure through the device was converted to a temporal variation. The Rayleigh-Plesset equation was numerically solved to model the bubble dynamics subjected to this temporally varying pressure. To investigate particularly violent collapse, we incorporated capabilities for quad-precision calculations. We computed the corresponding growth of nuclei of different sizes during passage through the device. Cavitation was defined to occur when the maximum bubble radius significantly exceeds the initial radius (e.g., tenfold). Maximum bubble sizes, as well as collapse times, were investigated for a wide range of conditions. The result of this study was a heat map for maximum bubble size, for a range of nucleus sizes and inlet velocities (Figure 16.4).

The long-term goal of this work is to assist experimental investigations of cavitation inception with CSMs. For CSM experiments exhibiting cavitation, the plan will be to run the present model for the relevant flow conditions and geometry. Analysis of the bubble growth and collapse will then enable the determination of the expected nucleus sizes. To enhance the fidelity of the model, more sophisticated physics will be incorporated, including compressibility and thermal effects.

16.5 Counterweight and Negative Stiffness Mechanisms for Resonance Tuning in Pendulum-Based Wave Energy Converters

Stacey Jiang, Binh Truong, Vishnu Vijayasankar and Lei Zuo

The sustainability and stability of marine ecological systems depend on accurately monitoring variations in ocean chemical and physical properties, including pH, temperature, and salinity. Current ocean monitoring systems powered by solar energy rely heavily on the weather conditions, while conventional batteries face the challenge of limited lifespan. Integrating sensors with small wave energy converters offers a sustainable and reliable power source, enabling continuous data collection in marine environments. Wave energy converters feature diverse design architectures, including configurations that employ pendulums within floating structures. However, integrating a pendulum into wave energy converters is challenging, since the low frequency of ocean waves requires impractically long arm lengths in traditional designs to form a resonator. To address these obstacles, we present an integrated approach that incorporates both counterweight pendulum and negative stiffness mechanisms.

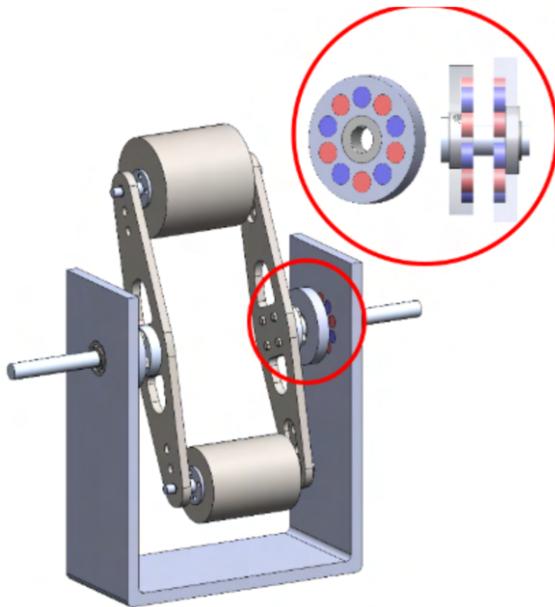


Figure 16.5: Designed prototype of the negative stiffness based Wave Energy Converter.

We investigated the performance of an internal pendulum-based wave energy converter equipped with two frequency tuning techniques, a counterweight with adjustable arms and magnetic softening effects, to approach or achieve resonance and thereby increase output power. The natural resonance frequency of the single pendulum is determined, setting the reference. The effectiveness of each tuning method, both individually and in combination, is then evaluated. The counterweight adjusts the gravity restoring torque, while the magnetic softening component introduces repulsive forces near equilibrium, further reducing the overall system stiffness. As the effective resonance frequency of the pendulum comes closer to the ocean wave frequency, its oscillation amplitude increases, which in turn improves the power delivered to the load.

Simulation results from COMSOL demonstrate that negative stiffness over a wide range of rotation can be obtained by placing two permanent-magnet disks in proximity, with achievable amplitudes up to 90 degrees. This effective operating angle can be adjusted by varying the number of magnets, while the magnetic equivalent torque can be optimized by adapting the gap between the disks. A prototype has been fabricated (Figure 16.5), and the preliminary experimental results indicate considerable advantages of the counterweight pendulum over its conventional single-pendulum counterpart. To quantitatively validate our hypothesis, further rigorous experiments are being conducted. Ultimately, we aim to demonstrate that the integration of counterweight pendulum and negative stiffness concepts can facilitate efficient resonance tuning in small-scale pendulum-based wave energy converters, resulting in substantial improvement in the maximum power harvested from ocean waves.

In summary, this work introduces an alternative strategy for achieving tunable resonance in compact, pendulum-based wave energy converters by combining counterweight and magnetic negative stiffness mechanisms. By addressing frequency-matching challenges and enabling sustained, renewable power generation, these innovations expand the potential for reliable, long-term ocean monitoring beyond existing solutions.

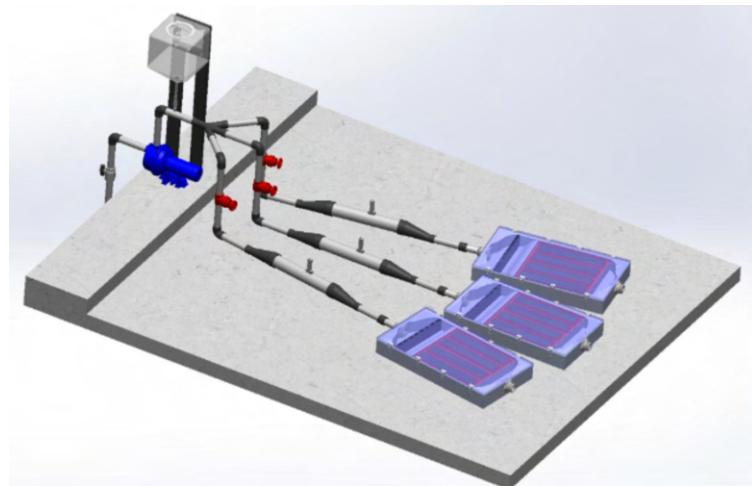


Figure 16.6: Designed growth loop to study biofouling growth in laboratory environment.

16.6 Effects of Drag-producing Biofouling on the Performance of Ships

Bhumi Kumar, Rodrigo Vilumbrales-Garcia and Harish Ganesh

Biofouling, the growth of marine organisms such as algae and barnacles on submerged surfaces, poses a significant challenge for marine vessels. These organisms grow on ship hulls, increasing the surface roughness, making the vessel less hydrodynamic, and requiring additional fuel to go the same distance. Studying live biofouling presents two primary challenges. The first is that living organisms do not want to grow in the same way or amount every time. Second, sloughing—the degradation of a fouling or sample layer during a test—prevents repeatability and invalidates any data collected. The primary objective of this research is to understand and quantify the effects of drag-producing biofouling on the performance of ships.

The research had three key focus areas: hydrodynamic testing for synthetic biofouling, creating a code for the Skin Friction Flow Facility (SF3), and building a growth loop (Figure 16.6). For hydrodynamic testing for synthetic biofouling, three different materials were explored: carrageenan, polydimethylsiloxane (PDMS), and alginate. For PDMS, the study investigated three different ratios of silicone to PDMS (15:1, 7.5:1, 5:1). For alginate, different methods were tested to get it to fully gelify through Calcium ion interactions. To improve the reliability of drag measurements, a new code for the Skin Friction Flow Facility (SF3) was written that simultaneously records pressure at all 10 locations.

We found that carrageenan samples degrade at low velocity values, making it unreliable, as it leads to inconsistencies in data collection. However, it may lead to a way to study sloughing. In the tests conducted in the SF3, the flushed PDMS samples at a 10% plate coverage led to a drag increase of 6% with respect to the smooth plate. The alginate also seems promising, as it was able to pass the survival tests in the SF3. The research successfully evaluated three different synthetic materials for future biofouling studies and defined material synthesis and testing protocols. It also improved the capabilities of SF3 with a new code that will allow for more accurate and reproducible data. The results from the PDMS testing prove the viability of using synthetic materials to simulate the drag increase caused by biofouling. The next steps for the research include exploring different material properties, heights, arrangements, and coverage to better understand and mitigate the effects of biofouling on ship performance.

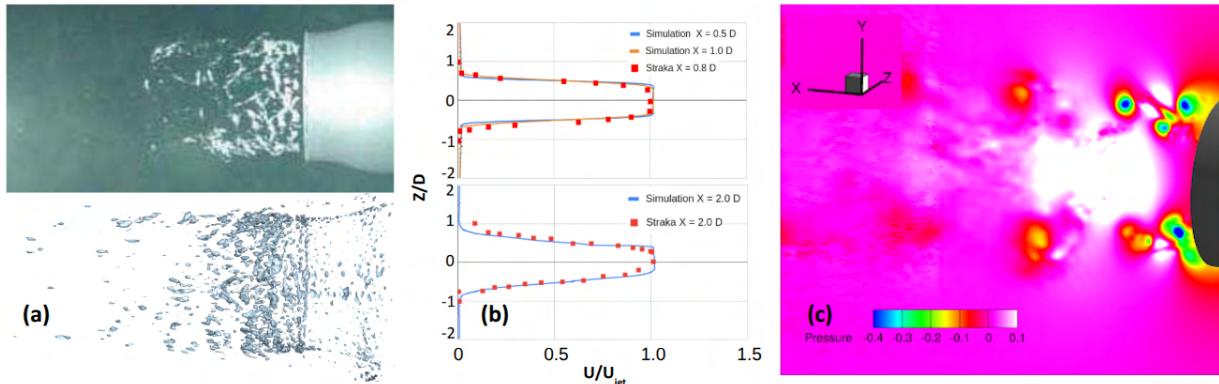


Figure 16.7: (a) Iso-surface of Q-criterion showing similar flow structures in simulation (bottom) compared to experiment (top) around the jet. (b) Velocity profiles at several jet-wake location showing good-agreement with the experiments. (c) Normalized pressure contour downstream the nozzle showing low pressure vortices where cavitation inception occurs.

16.7 Simulating Cavitation Inception in Turbulent Jets

Kian Montazery, Mehedi H. Bappy and Krishnan Mahesh

Turbulent jets play a crucial role in numerous engineering applications, but the cavitation phenomena associated with them present significant practical challenges. The flow fields generated by these jets are characterized by unsteady, multi-scale turbulence with complex vortex dynamics, stochastic nuclei distribution, and pronounced scale effects. These features greatly complicate efforts to accurately simulate the onset of cavitation using conventional numerical methods. The objective of this project was to simulate 3D cavitating turbulent jets and apply a recently developed subgrid-scale (SGS) model to estimate cavitation inception at various grid resolutions. The SGS inception model uses unresolved subgrid-scale fluctuations to estimate their contribution to fluctuating pressure and cavitation event rates without substantially increasing computational demands.

The project's scope encompassed the analysis of trends in cavitation inception across several velocity ratio between the jet and ambient water (ranging from quiescent to non-quiescent conditions) and Reynolds numbers. Multiple levels for each parameter were systematically tested to understand their effects on inception behavior. The simulations were validated against experimental data at different velocities and velocity ratios. The results consistently showed the formation of low-pressure regions within turbulent shear layers one to two nozzle diameters downstream from the jet exit—locations particularly susceptible to cavitation. These regions also exhibited the highest prevalence of subgrid-scale fluctuations, highlighting that ignoring such fluctuations in conventional LES may lead to under-prediction of cavitation risk. Increasing the velocity ratio at a fixed Reynolds number led to greater turbulence intensity and lower minimum pressures. Importantly, the radial locations of lowest pressure remained consistent across all test conditions, although the axial location moved downstream with decreasing velocity ratio.

Additionally, a more detailed validation and scaling study has been designed based on the jet flow conditions and nozzle geometry across the full range of Reynolds numbers and velocity ratios considered by Straka et al. (2010). Initial simulations for a specific nozzle geometry and a range of velocities and velocity ratios were performed, and the results indicated agreement in the observed trends for cavitation inception with the experiment (Figure 16.7). Future work will include a scaling study across a wide range of Reynolds numbers, velocity ratios, and nozzle geometries to comprehensively determine their effect on cavitation inception for different engineering applications.

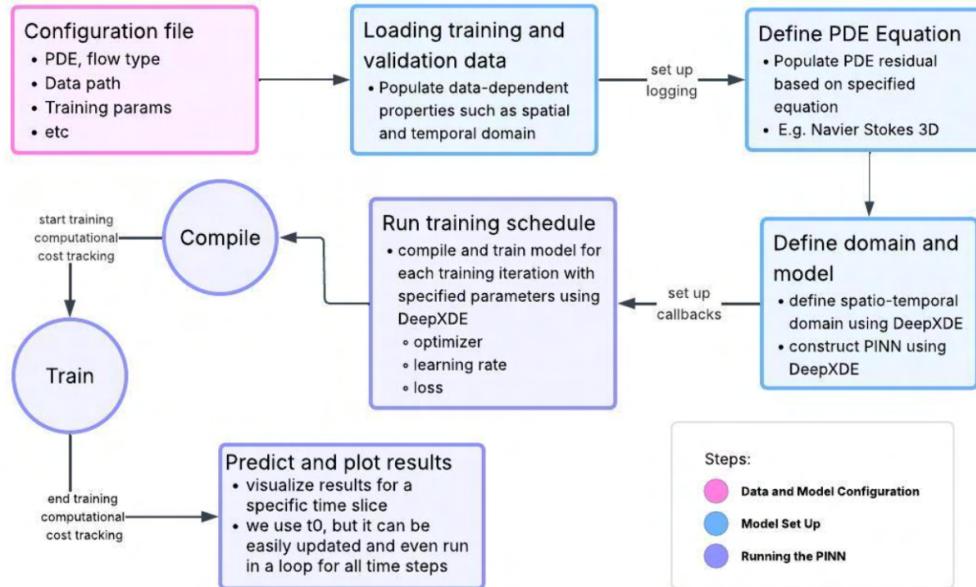


Figure 16.8: Modularized PINNs codebase.

16.8 Machine Learning Aided Flow Field Reconstruction from Sparse Particle Tracks

Daria Skalitzky, Xianzhang Xu and Krishnan Mahesh

Cavitation can erode the blades of marine propellers and increase the vibration and noise of ships. Instantaneous minimum pressure in the flow field drives cavitation inception. However, the minimum pressure is not accurately obtained from sparse experimental data (PIV/PTV) or simulation data. Our goal is to apply a physics-informed deep learning approach to reconstruct the full flow field from sparse particle tracks and accurately capture the minimum pressure.

To implement this model, we developed a modularized Physics-Informed Neural Network (PINNs) codebase using the DeepXDE library. The framework is summarized in Figure 16.8. The codebase dynamically adapts to different vortex models and domains through the configuration parameters, allowing for simple, flexible experimentation. Training is performed using only sparse velocity measurements, and the model predicts the fine-resolution velocity and pressure fields. It was evaluated with data generated using analytical solutions of three canonical vortex flows: 2D Taylor-Green & Lamb-Oseen vortex, and 3D Burgers vortex. Using this framework, we carried out a series of numerical experiments to study how the sampled particles affect the prediction results of the pressure field reconstruction. In particular, we tested how the sampling density and the spatial distribution (random r and uniform r) affect the model's performance. Then, to evaluate the robustness of our PINNs framework under these more realistic conditions, Gaussian noise was added independently to both velocity and spatial particle measurements. Three noise levels were chosen with standard deviations of $\sigma = 0.01, 0.05, 0.1$.

The model was able to accurately recover the full pressure and velocity fields in 2D for the Taylor-Green and Lamb-Oseen flows. The framework was extended to 3D and applied to the Burgers vortex, which was also accurately reconstructed. The results of the sampling experiments showed that the uniform-radius sampling strategy produced more reliable reconstructions and increasing the sampling density further improved the accuracy. The noise injections tests showed that PINNs can tolerate moderate noise and still recover the essential flow features. Though, as ex-

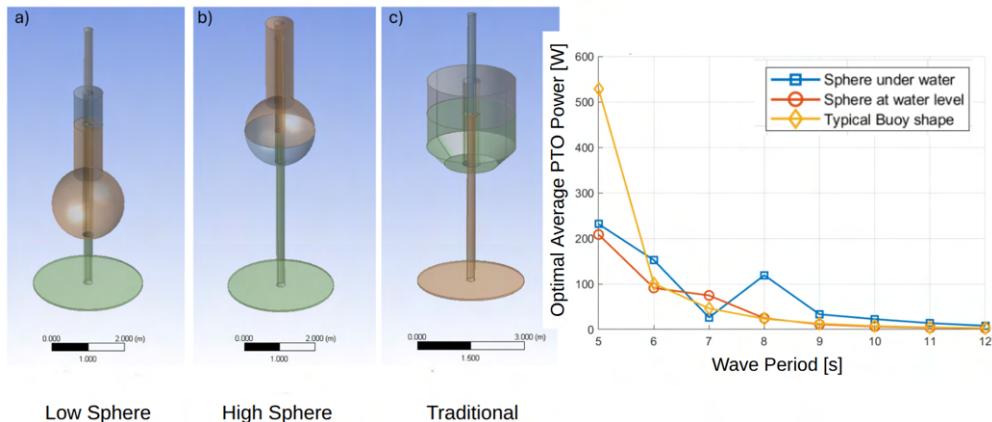


Figure 16.9: Designed wave energy converter and optimal PTO power at different wave periods.

pected, increasing the noise level led to greater uncertainty in the predictions and larger deviations from the ground truth minimum pressure. The Taylor-Green and Burgers vortices were especially robust, with only minor deviations even under heavy noise. Our study showed the importance of sampling strategy for the Lamb-Oseen vortex, with denser and more uniformly distributed training particles improving reconstruction accuracy. We also found that PINNs are robust to moderate levels of noise, though accuracy depends on the type of noise, and the specific vortex.

16.9 Autonomous Hybrid Wind Wave Energy Converter Data Buoy

Justin Song, Jianuo Huang and Lei Zuo

Continuous ocean observation in harsh and ice-covered regions is very costly and requires high maintenance. Traditional ocean-observing platforms depend heavily on solar energy and frequent vessel-based maintenance, making them ineffective in high-latitude winters characterized by prolonged darkness and heavy ice coverage. Remote coastal regions and the Marginal Ice Zone (MIZ) suffer from seasonal data gaps that hinder both scientific understanding and operational decision-making, as instruments lie dormant beneath thick ice or fail outright under crushing loads. These gaps drive up operations and maintenance (O&M) costs, since each re-deployment requires ice-breakers or specialized crews, often in dangerous conditions. By contrast, a buoy capable of autonomously harvesting both wave and wind energy promises a continuous, low-O&M power source that can sustain sensors and communications during ice formation, storms, and periods of darkness, all of which are essential conditions for year-round ocean observation.

Our solution directly addresses this challenge with a hybrid wind-wave energy-harvesting buoy featuring a Mechanical Motion Rectifier (MMR) Power Take-Off (PTO) and ball-screw mechanism that converts bidirectional heave into unidirectional generator rotation. Simultaneously, we integrate a vertical-axis wind turbine (VAWT) that drives the same generator in parallel. This compact system yields a reliable and continuous output of 100-300 W at the wave condition, sufficient to power a comprehensive set of environmental sensors, data processors, and a meshed radio frequency (RF)/acoustic communications network (Figure 16.9). Additionally, a stored battery bank system is integrated for load leveling and blackout resilience. The onboard electronics manage energy flows, prioritize critical payloads, and perform self-diagnostics, while the remote health-monitoring system alerts operators only when maintenance is necessary, thereby reducing expensive vessel visits. Communication is enabled by a mesh network, in which each buoy serves as both a sensor node

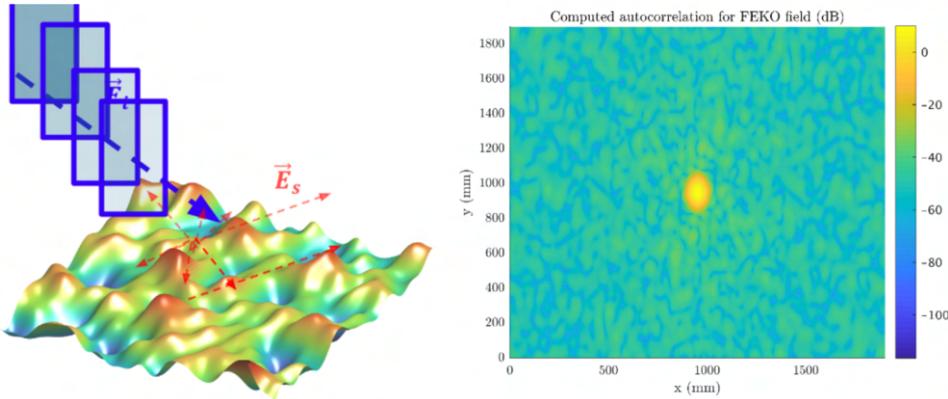


Figure 16.10: EM wave scattering over a randomly generated surface and computed autocorrelation of the scattered waves.

and a relay. Data is dynamically routed via RF or satellite links, to adjust transmission parameters and maintain reliability even under challenging conditions. Local processing handles basic quality control and event detection, such as ice-quake threshold breaches, before forwarding compressed summaries to shore or to neighboring buoys for multi-hop relay. This integrated approach ensures that critical time-series measurements (e.g., temperature, salinity, wave height, ice thickness) remain uninterrupted. A smaller scaled version of the buoy was prototyped this summer, and tested in the Huron river, where it showcased stability and power generation. With this proof of concept, we can move forward with creating a to-scale version of the prototype and run tank tests to validate simulations on power generation, stability, and natural frequency for the purposes of final deployment near the arctic rim.

16.10 Statistical Modeling of Electromagnetic Waves Scattering from Random Ocean Surface

Ryan Strohman, Behzad Yektakhah and Kamal Sarabandi

A theoretical framework for electromagnetic scattering from arbitrary, random, rough surfaces was employed to model scattering from the ocean. Potential applications include weather prediction, remote detection of targets in water, and local vector wind measurements (speed and direction).

Empirical models of the power spectrum of ocean waves were used to generate realistic ocean surfaces in MATLAB for use in full-wave electromagnetic simulations. The high electric permittivity of saltwater at microwave frequencies was leveraged to simplify calculation of the scattered fields under plane wave excitation by considering the ocean as an impedance surface. Since we considered the ocean as a random surface, we generated and simulated 15 different possible ocean surfaces, which allowed the statistical assumptions present in the scattering theory to be valid. Furthermore, the surfaces themselves were required to be extremely electrically large ($80\lambda \times 80\lambda$), which necessarily increased the simulation time and computational load. Simulation of all surfaces took about one month.

To verify the validity of the theory for modeling the ocean, various quantities were computed and compared against theoretical prediction. As an example, for a truly random surface it is expected that the average scattered field (calculated across all surfaces) tends to zero, which we confirmed in our simulations. Another example is that the two orthogonal components of the scattered field (co- and cross-polarization) are statistically independent, which was also confirmed (Figure 16.10).

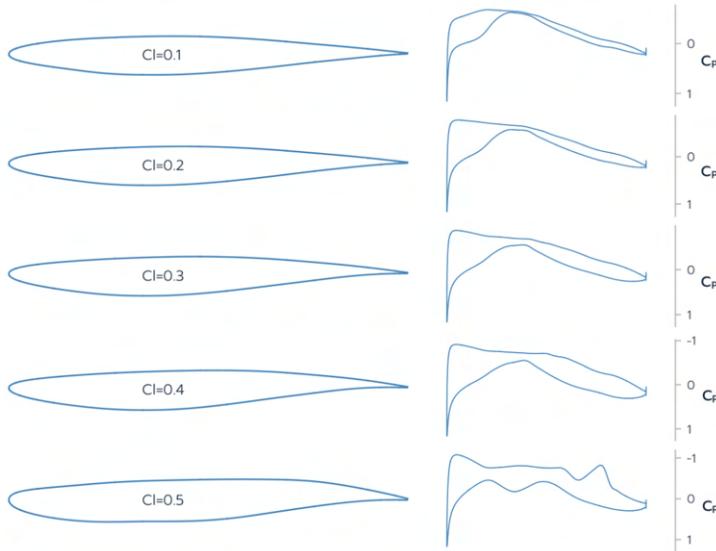


Figure 16.11: Optimal airfoil shapes for different lift constraints.

The final component of the project was using the verified statistical models to completely characterize the total scattered field using only a single measurement - e.g using one co-polarized field component to calculate the other co-polarized field component. This can be done due to the fact that these two quantities are not statistically independent, and their interaction is characterized by a joint probability density function. This part of the project is still ongoing and promising.

16.11 Hydrodynamic Shape Optimization of Hydrofoils and Propellers

Peter Susanto, Mohamed A.S. Abdul-Kaiyoom and Joaquim R.R.A. Martins

The design of hydrofoils and propellers is crucial to the longevity, performance, environmental impact, and efficiency of seafaring vessels. Multidisciplinary design optimization (MDO) aims to find optimal shapes for hydrofoils/propellers for various objectives, such as minimizing drag or maximizing lift.

The University of Michigan's Multidisciplinary Design Optimization Lab (MDOLab) has developed a powerful framework for this task, using gradients from the CFD flow solver (ADflow) to iteratively modify the shape and optimize for a certain objective. However, there are some limitations with this framework, namely the cost per optimization job and the reliance on some in-house MDOLab packages.

Our work this summer involved exploring deep reinforcement learning (DRL) and its viability within hydrofoil/propeller MDO. We built a framework for DRL shape optimization of airfoils using many of the same MDOLab tools. This framework trains an agent to modify the design variables to find an optimal shape for a target lift coefficient. While training, the agent explores different airfoil shapes. During the training process, the agent observes and takes action to learn the design space for mapping. The updated hydrofoil shape for the given action is required to be a valid shape to run the CFD simulation to get the performance metrics. Oftentimes, these shapes are not valid, and the CFD simulation fails. Therefore, we developed several methods to enforce the thickness and curvature constraints at each iteration to obtain a feasible shape (Figure 16.11). This results in more robust training.

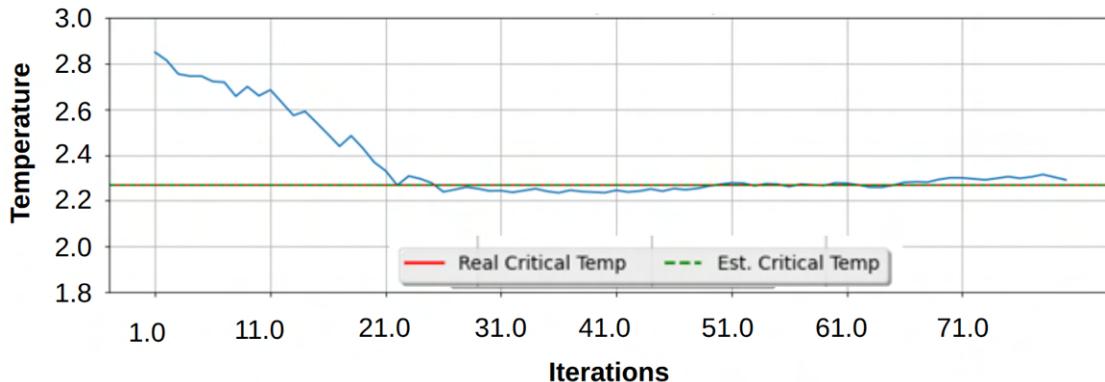


Figure 16.12: Rapid critical temperature optimization via automatic differentiable Monte Carlo simulations. A stochastic gradient descent algorithm was implemented on an differentiable Ising model simulator (50x50 square lattice) to search for the critical temperature at which there is a second order phase transition (with respect to the specific heat). The final calculated critical temperature was 2.267, which agrees closely with the known analytical solution 2.269.

16.12 Differentiable Monte Carlo Simulations

Evan Teal, Chenyang Huang, Andrew Li and Venkat Viswanathan

Monte Carlo (MC) methods are essential computational techniques for sampling high-dimensional spaces and are widely used in the study of physical and chemical systems. Their application is fundamental in modeling processes such as phase transitions while adhering to fluctuation-dissipation relationships. Despite their strengths, optimizing MC simulations is difficult due to inherent stochasticity, making analytical derivations for complex adjoint problems particularly challenging. The primary objective of this project was to explore, implement, and verify an automatic differentiable Monte Carlo (ADMC) framework to streamline and automate differentiation within MC simulations.

Initial efforts focused on reviewing existing methods of traditional & Markov Chain - Monte Carlo (MC) algorithms and implementing the basic Metropolis algorithm for the Ising Model in both conserved & unconserved cases, as well as scenarios involving external energy fields (Figure 16.12). A basic Metropolis algorithm was developed for sampling from a Gaussian distribution. Parallel to these activities, research was conducted into differentiable programming, particularly the capabilities offered by JAX, to support the development of differentiable MC simulations. Several challenges were encountered, particularly on convergence and validation of the ADMC objective. There was limited guidance in existing literature on model training, which necessitated extensive empirical testing and debugging. Compatibility of proposed ADMC objectives with classical MC and stochastic algorithms was verified, establishing the validity of the framework. A robust approach was developed to compute first-order gradients of MC observables, using both gradient computation and averaging over multiple random seeds to ensure consistent gradient descent performance.

Over the course of the project, significant progress was made in implementing and testing core MC and ADMC concepts. The work successfully demonstrated convergence-checking criteria for the Ising Model and Metropolis algorithms, as well as general training strategies for ADMC objectives. These results confirm the viability of ADMC frameworks for optimizing MC simulations and set the stage for future research into more complex applications, such as Kinetic Monte Carlo methods and advanced automatic differentiation techniques.

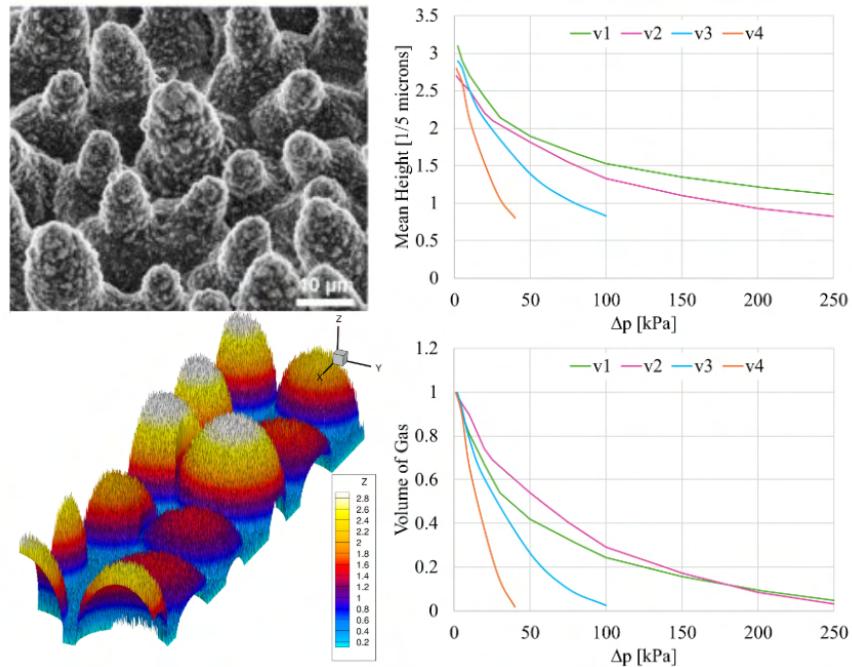


Figure 16.13: *Left:* An SEM image of lotus leaf (top) & modeled surface representation (bottom) with 50 nm nano-structure spacing (v1). *Right:* Normalized mean interface height (top-right) and volume of trapped gas (bottom-right) over surfaces with different nano-structure spacings (v2: 100 nm, v3: 200 nm, v4: 400 nm) at different interfacial pressure differences (Δp).

16.13 Modeling and Performance Prediction of Natural Superhydrophobic Surfaces

Sofie Vos, Mehedi H. Bappy and Krishnan Mahesh

Superhydrophobic surfaces (SHS) are materials that repel water strongly due to surface roughness and low surface energy. Found in nature on leaves such as lotus, rice, and salvinia, these surfaces allow for unconventional drag reduction. This project focused on modeling and predicting the performance of natural SHS by constructing accurate three-dimensional representations of their micro- and nano-structures and studying how liquid interfaces interact with them under different environmental conditions.

A variational level-set method was used to generate representations of natural surfaces, beginning with simple cylindrical geometries and progressively adding ellipsoidal and fractal features to mimic realistic roughness. Simulations were conducted for different canonical SHS and natural SHS with micro- and nanoscale roughness at multiple interfacial pressure differences ranging from 2 to 300 kPa. For each case, interface statistics such as mean and RMS height, trapped gas volume, Wenzel roughness, solid fraction, and autocorrelation length were calculated to quantify the response of the surface.

The results (Figure 16.13) showed that the nanostructures are critical for SHS to trap air and reduce drag. Lotus leaves sustained air layers at pressures up to 300 kPa, while rice leaves experienced an earlier interface collapse. Closer nanostructure spacing was found to improve performance by increasing trapped gas volume and maintaining interface height, thereby enhancing superhydrophobicity. These findings suggest that nanoscale spacing and density of structures play a key role in drag reduction and provide design principles for bioinspired SHS in engineering applications.

Publications and Presentations

Publications and conference presentations resulting from CNRE research are listed below.

1. L. Silwal, R. Vilumbrales-Garcia, A. Sareen. (2026). Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow. *International Journal of Heat Fluid Flow*, 117, 110071.
2. Z. Wei, P. Dong. (2025). An Analytical Structural Strain Method for Low-Cycle Fatigue Evaluation of Thin-Walled Pipes Under Bending. *Journal of Thin-Walled Structures*, 114148.
3. N. Gupta, K. Duraisamy. (2025). Computational and physical considerations for the development of machine learning augmented turbulence models. Accepted *International Journal of Heat and Fluid Flow*.
4. B. D. Truong, L. Zuo. (2025). On the Optimal Performance of Oscillating Surge Wave Energy Converter. Submitted.
5. Z. Wei, P. Dong, S. Wu. (2025). A Load-Path Nonproportionality Parameter for Time-Domain and Frequency-Domain Multiaxial Fatigue Evaluation. Submitted to *International Journal of Fatigue*.
6. A. Szabo, K. Mahesh. (2025). Investigation of the asymmetric flow field around the prolate spheroid using global stability analysis. Submitted to *Journal of Fluid Mechanics*.
7. M. H. Bappy, K. Mahesh. (2025). A Dynamic Framework for Simulating Cavitation Inception with Large Eddy Simulation. Submitted to *International Journal of Multiphase Flow*.
8. G. J. Wu, S. Anantharamu, K. Mahesh. (2025). Adjoint-Based Assimilation of External Data with High-Fidelity Simulations of Complex Flows. Submitted to *Journal of Scientific Computing*.
9. X. Xu, D. Skalitzky, K. Mahesh. (2025). Machine Learning–Aided Estimation of Minimum Pressure from Sparse Velocity Data in Vortex Flows. Submitted to *Physical Review Fluids*.
10. M. Saja, J. Martins. (2026). Multipoint versus UQ-based optimization. Abstract accepted to AIAA Conference 2026.
11. S. Kim, M. Bappy, K. Mahesh. (2026). Compressible LES of hydrofoil sheet-to-cloud cavity transition and the characterization of erosion. Abstract accepted to 36th Symposium on Naval Hydrodynamics, Busan, Korea.
12. T. Leasca, M. Bappy, A. Madabhushi, K. Mahesh. (2026). Ducted propulsor and twin vortex cavitation inception and the development of a subgrid model. Abstract accepted to 36th Symposium on Naval Hydrodynamics, Busan, Korea.

13. R. Vilumbrales-Garcia, M. Elsouht, J. Estrada, H. Ganesh. (2026). Understanding the drag of heterogeneous and sparse soft biofilms. 14th International Symposium on Turbulence and Shear Flow Phenomena (TSFP14).
14. M. Elsouht, R. Vilumbrales-Garcia, V. Molina, M. First, J. Lee, H. Ganesh. (2026). Understanding the drag of soft biofilms: A multidisciplinary experimental approach. Abstract accepted to 36th Symposium on Naval Hydrodynamics.
15. A. Li, V. Viswanathan. (2025). Differentiable Programming/Physics for Phase Field Model Parameter Optimization. CHiMaD Phase Field XVIII, Evanston, IL, USA.
16. B. Yektakhah, K. Sarabandi. (2025). Statistical modeling of electromagnetic wave scattering from the ocean surface. IEEE International Symposium on Antennas and Propagation (APS/URSI), Ottawa, Canada.
17. B. D. Truong, J. Huang, L. Zuo. (2025). Power optimization for an oscillating surge wave energy converter with motion amplitude constraints. *Proc. SPIE 13432, Active and Passive Smart Structures and Integrated Systems XIX*, 134321B, 5 May 2025.
18. W. Lin, J. Huang, B. D. Truong, L. Zuo. (2025). System Modeling and Power Optimization of a Single-body Point Absorber Wave Energy Converter. *Proc. EWTEC*, vol. 16, Sep. 2025.
19. Z. Wei, P. Dong. (2025). Loading-path dependent time- and frequency-domain multiaxial variable amplitude fatigue analysis for marine structural reliability. 16th International Symposium on Practical Design of Ships and Other Floating Structures (PRADS 2025), October 19–23, Ann Arbor, MI, USA.
20. Z. Wei, N. Vladimir, P. Dong. (2025). Wet vibration analysis and uncertainty quantification of ship plates. 16th International Symposium on Practical Design of Ships and Other Floating Structures (PRADS 2025), October 19–23, Ann Arbor, MI, USA.
21. L. Silwal, A. Sareen. (2025). Effect of Surface Dimples on Wake Dynamics of a Pitching Foil in Quiescent Flow. The 4th Colloquium on Vortex Dominated Flows, June 3-6, 2025, Bozeman, Montana.
22. M. H. Bappy, K. Mahesh. (2025). Modeling Cavitation Inception During Vortex-Vortex Interaction due to Crow Instability, 12th International Conference on Multiphase Flow, ICMF 2025, May 2025, Toulouse, France.
23. L. Silwal, R. Vilumbrales-Garcia, R., A. Sareen. (2025). Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow. Remote Colloquium on Vortex Dominated Flows Series, Sept 19, 2025.
24. M. H. Bappy, K. Mahesh, Dynamic Cavitation Inception Modelling in the Subgrid Scale of Large Eddy Simulation, US-Japan seminar on Two Phase Flow Dynamics, June 2025, Tokyo, Japan.
25. A. Szabo, K. Mahesh. (2025). Low Reynolds number modal stability analysis of the 6-1 prolate spheroid. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.

Publications and Presentations

26. M. H. Bappy, K. Mahesh. (2025). A Dynamic Approach to Model Subgrid Scale Cavitation Inception. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
27. X. Xu, D. Skalitzky, K. Mahesh. (2025). Machine Learning–Aided Flow Field Reconstruction and Minimum Pressure Estimation from Sparse and Noisy Particle Measurements. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
28. K. Montazery, M. H. Bappy, K. Mahesh. (2025). Simulation and Modeling of Cavitation Inception in Turbulent Water Jets. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
29. D. Gola, S. Prajapati, K. Mahesh. (2025). Two-way FSI Modeling of Biofouled Surfaces in Turbulent Flow. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
30. N. Gupta, K. Duraisamy. (2025). Enforcing Learnability and Model Consistency in Data Driven Turbulence Modeling – MS273 Advancing Turbulence Modeling with Scientific Machine Learning Part II of II, SIAM Conference on Computational Science and Engineering, March 3–7, Fort Worth, Texas.
31. G. J. Wu, S. Anantharamu, K. Mahesh. (2025). Parallel Implementation and Assessment of Adjoint-Based Data Assimilation with the Hybridizable DG algorithm. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
32. V. Agarwal, G. J. Wu, K. Mahesh. (2025). Gradient Based Optimization of Superhydrophobic Surface Geometry for Drag Reduction. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
33. R. Vilumbrales-Garcia, L. Silwal, A. Sareen. (2025). Surface Roughness Influences Vortex Interactions and Jet Stability in Pitching Foils in Quiescent Flow. Remote Colloquium on Vortex Dominated Flows Series, Sept 19, 2025.
34. R. Vilumbrales-Garcia, C. Sun, B. Kumar, J. Estrada, H. Ganesh. (2025). Towards inferring drag produced by patchy, heterogeneous soft biofilms. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
35. M. Elsouht, E. G. Callison, R. Vilumbrales-Garcia, H. Ganesh, S. L. Ceccio. (2025). Roughness and the effect of streamers motion on drag in biofilm-like surfaces. 78th Annual Meeting of the APS Division of Fluid Dynamics, November 23–25, Houston, Texas.
36. M. H. Bappy, A. Madabhushi, K. Mahesh. (2024). Modeling Incipient Cavitation During Vortex-Vortex Interaction due to Crow Instability. 77th Annual Meeting of the APS Division of Fluid Dynamics, Nov 24-26, Salt Lake City, Utah.
37. G. J. Wu, S. Anantharamu, K. Mahesh. (2024). An Adjoint-Based Data Assimilation Algorithm using the Hybridizable Discontinuous Galerkin Framework. 77th Annual Meeting of the APS Division of Fluid Dynamics, Nov 24-26, Salt Lake City, Utah..
38. R. Vilumbrales-Garcia, L. Silwal, A. Sareen. (2024). Effects of Surface Roughness on a Pitching Foil in Quiescent Flow Conditions. American Physical Society, Division of Fluid Mechanics, , Nov 24-26, Salt Lake City, Utah.

2024-2025 Roster

The CNRE roster is provided below. Postdoctoral fellows, graduate students, faculty and undergraduate student researchers are separately listed along with their contacts and research topics.

Postdoctoral Scholars/Graduate Students*

| Name | Email | Research Area |
|----------------------------|--------------------|-------------------------------------|
| Mehedi H. Bappy | mehedib@umich.edu | Multiphase flow modeling |
| Daniel Esterkin* | dester@umich.edu | Kirigami materials |
| Divyanshu Gola | divyans@umich.edu | Biofouling simulation |
| Niloy Gupta* | niloy@umich.edu | ML & data-driven modeling |
| Mohamed A.S. Abdul-Kaiyoom | arshsaja@umich.edu | Optimization & AI |
| Seungnam Kim | skimnaoe@umich.edu | Cavitation simulation |
| Andrew Li | andreli@umich.edu | Battery simulation |
| Eva Pontrelli* | epontrel@umich.edu | Vapor jet printing & hydrophobicity |
| Lokesh Silwal | lokeshsi@umich.edu | Experimental flow control |
| András Szabó | andrasz@umich.edu | Flow stability & asymmetry |
| Binh Truong | drtruong@umich.edu | Wave energy converter design |
| Rodrigo Vilumbrales-Garcia | rodrigga@umich.edu | Biofouling experiments |
| Zhigang Wei | zgwei@umich.edu | AI & modular construction |
| Gao Jun Wu | gaojunwu@umich.edu | Adjoint-based data assimilation |
| Xianzhang Xu | xianzhx@umich.edu | ML-enabled simulation |
| Behzad Yektakhah | byekta@umich.edu | Electromagnetics modeling |

Faculty

| Name | Department |
|------------------------|---|
| Steve Ceccio | NA&ME, Mechanical Engineering |
| Pingsha Dong | NA&ME |
| Karthik Duraisamy | Aerospace Engineering |
| Jon Estrada | Mechanical Engineering |
| Harish Ganesh | NA&ME |
| Eric Johnsen | Mechanical Engineering |
| Krishnan Mahesh | NA&ME |
| Joaquim R.R.A. Martins | Aerospace Engineering |
| Kamal Sarabandi | Electrical Engineering & Computer Science |
| Anchal Sareen | NA&ME |
| Max Shtein | Materials Science |
| Venkat Viswanathan | Aerospace Engineering |
| Lei Zuo | NA&ME |

Undergraduate Students

| Name | Email | Research Area |
|----------------------|--------------------|--------------------------------------|
| Veer Agarwal | veeraga@umich.edu | Optimization of drag reduction |
| Amy Fahmy | amyfhmy@umich.edu | Denticle-inspired control |
| Teagan Fynan | tfynan@umich.edu | Soft hydrogels |
| Manasa Gollapalli | manasavg@umich.edu | Actuation for morphable surfaces |
| Sunaad Gurajada | sunaadg@umich.edu | Bubble collapse modeling |
| Katarina Jevtic | kjevtic@umich.edu | Cavitation bubble modeling |
| Stacey Jiang | jstace@umich.edu | Wave energy converter |
| Bhumi Kumar | bhumik@umich.edu | Biofouling experiments |
| Jackson Lewis | jadlewis@umich.edu | Turbidity currents |
| Kian Montazery | kianm@umich.edu | Cavitation inception LES |
| Megan Pribak | mmpribak@umich.edu | Underwater current energy harvesting |
| Vaibhav Ramaseshadri | vaibhavr@umich.edu | Cloud cavitation |
| Daria Skalitzky | dskalit@umich.edu | ML & data assimilation |
| Justin Song | jjiasong@umich.edu | Wind-wave observation buoy |
| Ryan Strohman | strohman@umich.edu | Electromagnetics modeling |
| Chelsea Sun | sunchel@umich.edu | Synthetic algae for biofouling |
| Peter Susanto | psusanto@umich.edu | ML & optimization |
| Evan Teal | evanteal@umich.edu | Monte-Carlo simulations |
| Sofie Vos | sevos@umich.edu | Natural superhydrophobic surfaces |
| Jack Zaeske | jzaeske@umich.edu | Physics-informed neural networks |