October 2021

NUMERICAL MODELING OF ADVANCED PROPULSION SYSTEMS

Peetak P. Mitra
University of Massachusetts Amherst

NUMERICAL MODELING OF ADVANCED PROPULSION SYSTEMS

A Dissertation Presented

by

PEETAK P. MITRA

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2021

Mechanical and Industrial Engineering

NUMERICAL MODELING OF ADVANCED PROPULSION SYSTEMS

A Dissertation Presented

by

PEETAK P. MITRA

Approved as to style and content by:

_____
Prof. David P. Schmidt, Chair

_____
Prof. Blair Perot, Member

_____
Prof. Jimi Oke, Member

_____
Prof. Sundar Krishnamurthy,
Chair of the Faculty,
Department of Mechanical and Industrial Engineering

DEDICATION

This thesis and the modest, incremental scientific contributions therein, are dedicated to my mother for without her sacrifices I would not be here. Thank you mom!

ACKNOWLEDGMENTS

In the spirit of the ancient African proverb, '*It takes a village to raise a child*', I would like to acknowledge the support of my *village* in getting me to this point. Firstly, I would like to acknowledge the immense support and contributions, in shaping my research career, of my doctoral advisor Prof. David P. Schmidt. He has been an inspirational teacher, guide and collaborator, and I will look back very fondly on the time that we spent together working on some challenging engineering problems. David provided me with unparalleled opportunities to explore my interests in machine learning and I would forever be thankful to him for that. It is truly an honor to have worked and learnt from you.

I would like to also acknowledge the support of my dissertation committee - Prof. Blair Perot and Prof. Jimi Oke, in the successful completion of my doctoral degree. Thank you very much for your participation, your probing questions and for your guidance throughout. It has been a great learning experience for me.

I have been very lucky to have collaborated with some leading scientists during the course of my Ph.D. degree and I would like to acknowledge some of them during this time. Dr. Lyle Pickett (Sandia) for teaching me how to do good science, and for his relentless pursuit to discover the truth. Dr. Christopher Powell (Argonne) and his team, for all his experimental datasets and the engaging intellectual discussions during the early phase of my doctoral studies. Dr. Peter Kelly Senecal (Convergent Science) for keeping me sharp (and in good spirits). Prof. Animashree Anandkumar (Caltech/NVIDIA) and her team, for giving me some big breaks in machine learning as I was starting in this new area. Friends-turned-collaborators, Dr. Gavin Portwood (LLNL), Dr. Arvind Mohan (LANL), Dr. Mateus Dias Ribeiro (DFKI), Dr. Tan

Nguyen (UCLA) and Dr. Niccolo Dal Santo (MathWorks) for their friendship and amazing research projects that I have been lucky to work on with them. I would like to acknowledge the support and help from Dr. Majid Haghshenas during the final months of my doctoral degree. It has been a pleasure to closely work and learn from him.

I also would like to acknowledge the support of my colleagues from the Multiphase flow simulation laboratory including Hannah, Sampath, Chinmoy, Gabe, Michael, Alden, and Brandon during the course of the doctoral degree.

My mentors from India, including Prof. Kannan Iyer (IIT B), Prof. Shivasubramaniam Gopalakrishnan (IIT B), Prof. Rajneesh Bhardwaj (IIT B) and Prof. Somnath Roy (IIT Kharagpur) deserve special mention for their guidance in the early phase of my career that provided me the opportunity to pursue my doctoral degree in the U.S.

The work in this thesis was mainly possible due to the generous financial support from industry partners. I would like to thank the partners from the Spray Combustion Consortium and the ICEnet consortium for their financial support and for steering the broader direction of my doctoral research. The support of NSF-funded XSEDE computing resources and the Leadership Computing Facility at Argonne, for running CFD simulations, is greatly appreciated.

Last, but not the least, I would like to acknowledge the sacrifices and support of my family in getting me here. Firstly, immensely thankful to my partner Julia, especially for her constant support during the final phase of my dissertation defense. My brother Shounak, for always showing me the way and for his constant encouragement throughout. And lastly, I would like to acknowledge the unparalleled sacrifices of my mother and for teaching me to never give up in the face of insurmountable adversities, qualities that helped me navigate difficult phases during my doctoral studies.

I truly could not have done this without each and every one of you, and I am eternally grateful to you all.

ABSTRACT

NUMERICAL MODELING OF ADVANCED PROPULSION SYSTEMS

SEPTEMBER 2021

PEETAK P. MITRA

B.E., BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor David P. Schmidt

Numerical modeling of advanced propulsion systems such as the Internal Combustion Engine (ICE) is of great interest to the community due to the magnitude of compute/algorithmic challenges. Fuel spray atomization, which determines the rate of fuel-air mixing, is a critical limiting process for the phenomena of combustion within ICEs. Fuel spray atomization has proven to be a formidable challenge for the state-of-the-art numerical models due to its highly transient, multi-scale, and multi-phase nature. Current models for primary atomization employ a high degree of empiricism in the form of model constants. This level of empiricism often reduces the art of predictive modeling, in the case of sprays to a mere data-fitting exercise to experimental observations by tuning model constants.

In this research, first a series of full three-dimensional (3D) Computational Fluid Dynamics (CFD) studies are presented that examine the factors affecting spray atomization, including the effects of nozzle geometry, transient injector needle motion

and in-cylinder thermodynamic conditions on the spray atomization behavior. Informed by this study, a novel reduced order model - Eulerian Lagrangian Mixing Oriented (ELMO), with experimentally-informed inputs is presented that aims to replace current primary atomization models such as the Kelvin-Helmholtz Rayleigh-Taylor (KHRT). The ELMO fuel spray atomization model is coupled to a full 3D gas phase solver, and validation studies on single/multi-hole injector configurations and spray conditions (diesel/gasoline) are presented. In addition, validation studies for ELMO are presented for a standard diesel engine. In the last section of this thesis, the applicability of data-driven machine learning (ML) models are explored. First, a ML based turbulence closure is presented for the four-stroke Darmstadt engine and performance against experimentally validated ground-truth data measured. Sensitivity studies indicate the data-driven model preserves the functional characteristics of the turbulence closure, consistent with theory. In the final study, a machine learning based surrogate model for estimating the discretization error of coarser meshes is developed and integrated into OpenFOAM. This surrogate model enhances the quality of coarse-mesh CFD simulations by adding necessary source terms to retrieve the 'lost' resolved scale information, thereby providing a higher fidelity solution at almost the same compute cost. This model is tested on different OpenFOAM solvers, including for an engine cold flow setting, and the generalizability of the data-driven framework commented upon.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

Low-mach number terrestrial propulsion systems such as the Internal Combustion Engine comprise a major proportion of the current fleet of on-road vehicles, a trend that is expected to continue well into the mid-21st century [101, 196, 181]. In the last few decades, major research emphasis has ensured these advanced propulsion systems are increasingly efficient and reduce their environmental footprint in terms of greenhouse gas emissions thereby adhering to or exceeding strict regulatory standards set by the Environmental Protection Agency (EPA) in the US. A study by Joshi et al. [101] reports light duty vehicles will have to continue to improve fuel economy by at least 3% annually in the near future to meet these evolving standards. This improvement in the fuel economy will be primarily driven by inventing newer, cleaner combustion technologies such as ducted fuel injection [148] as well as incrementally improving the state of the internal combustion engines. Numerical modeling is expected to play an important role in building reliable predictive research tools that help in evaluating new design concepts as well as exploring the parameter space, to develop cleaner, more efficient engines.

The Internal Combustion Engine (ICE) is a marvel of modern engineering. It has revolutionized the way the modern world functions. At the center of the ICE is the process of intermittent combustion caused by oxidizing energy-dense fuel, usually with air, in a combustion chamber. This process releases gases (as a by-product of the combustion), that forces mechanical components of the chamber to move, thereby converting chemical energy to useful work. For a detailed review of the ICE technology

and the various hydraulic and thermodynamical processes that affect it, the interested reader is advised to refer to an excellent treatise on this topic [99].

The intermittent fuel combustion process that drives the energy conversion is complicated and depends on a multitude of factors. In brief, the fuel combustion is the final step of a sequence of overlapping physical processes including fuel injection, atomization, mixing of the atomized fuel with the gas phase in the combustion chamber, followed by the energy release during the injection cycle [49]. Among these, spray atomization is a critical limiting process, which determines the rate of fuel-air mixing, and defines the characteristics of the combustion itself. Atomization helps in distributing the liquid fuel into smaller droplets thereby providing larger surface area for the liquid phase fuel and the gaseous air to mix, promoting vaporization of the droplets thereby driving the combustion process itself. In advanced propulsion systems, the fuel is injected into the combustion chamber in the form of sprays. Therefore, understanding the physics of sprays in the context of these propulsion systems will play a critical role in helping design better, cleaner engines and will remain the focal point of the majority of this thesis.

1.1 Challenges in multi-dimensional modeling for ICEs

As the primary area of focus of this thesis is in modeling sprays, we briefly present the main challenges in numerical modeling. These challenges are generally consistent for other problems in the world of multi-dimensional, multi-scale CFD. Spray atomization is defined as the process of breaking liquid jet into small droplets. This complex phenomenon occurs in a liquid jet when the velocity difference of this jet with respect to the surrounding gas is very high and droplets are formed in the close vicinity of the nozzle. Different regimes are distinguished based on the velocity of the liquid jet, and will be discussed in Chapter 2 of this thesis. The major challenges in understanding and thereby modeling the physical process of atomization include:

- **The Multi-scale nature**: The process of atomization, by definition, involves the disintegration of liquid jet into small droplets, spanning multiple scales. In a typical spray simulation, the mesh resolution varies from an order of a few micrometers (corresponding to the radii of the nozzle exit) in the near-field region to larger cells at the far-field locations.

- **Multiple time-scales**: The fact that multiple scales are involved means the complexity in accurately modeling the phenomena of atomization involves working with multiple timescales. The droplet-specific phenomena are very transient in nature (5ms), and processes such as coalescence, breakup often operate at different timescales.

- **Multi-Phase**: The process of atomization is a result of the interaction between the liquid jet and the immiscible gas phase, and, due to in-nozzle cavitation (and flash-boiling in case of gasoline sprays), the nature of the physics is inherently multi-phase.

Coupling these multi-scale, multi-phase phenomena makes the challenge in developing predictive multi-dimensional models particularly difficult. The subsequent chapter will provide a brief overview of some of the common methods used in the literature. It will also identify some current challenges and thereby knowledge gaps this thesis will aim to fill.

1.2 Scope of the thesis

The main contributions of this thesis are in extending our understanding of transient spray physics phenomena such as cavitation that affects atomization. Following a brief literature review in Chapter 2, studies using existing in-house Eulerian CFD solver HRMFoam are presented for identifying and characterizing internal flow behavior for single and multi-hole spray nozzles. The CFD results are validated using

state-of-the-art experimental studies [141, 131, 245, 164, 161]. Within this scope, factors causing phase change (cavitation) are studied and using statistical methods effect of vortex interactions between nozzles are reported. In Chapter 5 a new primary atomization model is presented, inspired by experimental observations and theory, for external spray modeling based on mixing-limited physics relevant to diesel-like fuels. In the final part of this thesis, data-driven methods have been explored to build turbulence closure models and discretization error surrogate models to improve the cost-accuracy tradeoff for general CFD simulations.

CHAPTER 2

BRIEF LITERATURE REVIEW: CHALLENGES AND OPPORTUNITIES

The focus of this chapter will remain on spray atomization as it is a central theme of this thesis. This brief review provides guidance on the current challenges and opportunities in multi-dimensional modeling of sprays in particular, and CFD in general. Spray atomization is an active area of research and investigation. The multi-scale, multi-phase nature of the physics are some of the primary challenges in building reliable spray models. This therefore requires an informed synergy between the theory, and high-fidelity experiments that are crucial to guide, tune and investigate the numerical models developed using first principles [shown in the schematic in Figure 2.1]. In the next sections, a brief theory of spray atomization will be introduced, following which the overview of various modeling approaches, ranging from multi-dimensional models, reduced order models and physics-guided data-driven models are discussed. The discussion of the experimental methodologies and investigations, are beyond the scope of this chapter and the reader is referred to expansive studies in the literature [c.f. [220, 50, 203]]. This review would help in identifying research gaps and opportunities to make further contributions, which will form the basis of the research in this thesis.

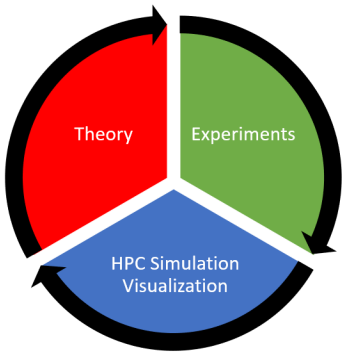

Figure 2.1. Synergistic approach in developing predictive multi-dimensional models where the theory helps in designing experiments and the experiments inform model design and tuning. A well-designed model then improves our understanding of these fast, transient physics.

2.1 The physics of sprays

An important step in improving efficiencies of engines is to understand the precise distribution and characteristics of fuel and air in the chamber throughout the combustion cycle. This is particularly true for Direct Ignition (DI) engines where the atomization of the fuel plays a critical role in the combustion process. Though experiments in this area of research has improved our understanding, it is difficult to obtain an accurate three-dimensional characterization of the combustion chamber because the spray is optically dense, length scales are small (about 100 μm), time scales are short (about 5ms) and the geometries complex. Therefore, multi-dimensional numerical modeling becomes an important tool to use for prediction as well as discovering previously unknown physics. Before we delve deeper into the different approaches for modeling these highly transient, multi-scale and multi-phase sprays, this section will echo some important concepts with regards to the physics of the sprays. A schematic of a typical injection process is shown in Figure 2.2, and it shows the dense liquid core region near the nozzle exit within the breakup region. It is this region, which is optically dense and creates unique challenges for experimental measurements and

visualizations. Due to the process of spray breakup including atomization, the intact liquid core breaks up into very small droplets at a distance downstream, introducing the multi-scale nature of the problem. The complexity of the breakup process is further compounded by the unusually large number of parameters that influence the spray behavior including the details of the design of the nozzle, liquid jet, coflow air velocity, turbulence, cavitation and the physical and thermodynamical properties of the liquid and the gas phase.

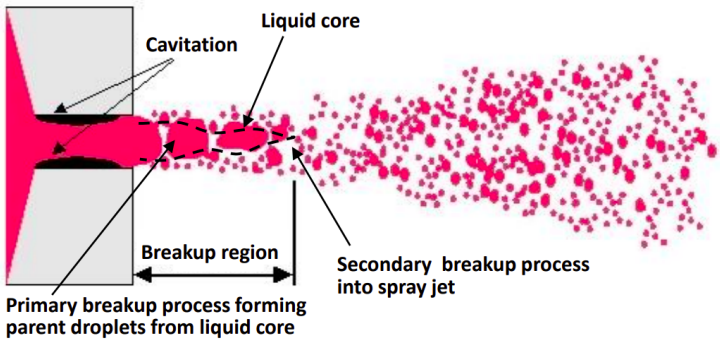

Figure 2.2. Schematic of a typical spray injection showing the different modes of break-up and cavitation within the injector geometry [figure adopted from Ashgriz [15]]

The dispersion of spray drops is important to bring about efficient mass and heat transfer between liquid and gas phases. Excellent reviews on different aspects of liquid atomization can be found in the literature [182, 85, 83, 135, 86, 243, 89, 183, 121]. These studies have established that spray properties are influenced by conditions in the flow at the nozzle exit, including turbulence and cavitation effects. Most pertinent research in this area has been conducted on understanding the mechanism of breakup of liquid jets injected from a single hole orifice into stagnant and co-flowing gaseous media [182, 89, 183]. The cases of stagnant and co-flowing gases need to be studied separately because in case of a gas coflow, the momentum carried by the gas flow can lead to liquid thinning and stretching and formation of membranes in the injected liquid [182]. The jet breakup phenomena have been divided into four breakup regimes

based on the operating conditions that correspond to different combinations of liquid inertia, surface tension and aerodynamic forces [182]. These regimes are due to the action of dominant forces on the jet leading to its breakup, and they have been identified in Figure 2.3

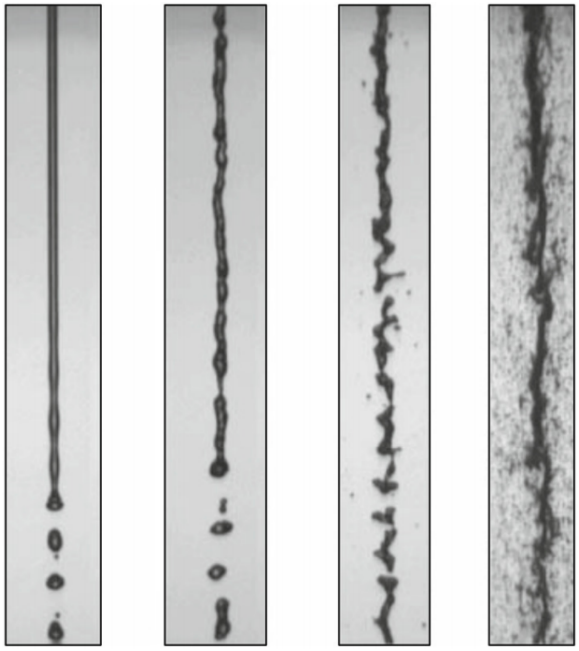

Figure 2.3. [L-R]: Cylindrical jet spray regimes. (a) Rayleigh regime at Re=790 We=0.06. (b) first wind-induced regime at Re=5,500 We=2.7. (c) second wind-induced regime at Re=16,500 We=24. (d) atomization regime at Re=28,000 We=70. [Images adopted from Leroux et al. [118]]

These different regimes of spray breakup are characterized further below and shown [from experimental images] in Figure 2.3:

- **Rayleigh jet breakup:** This is caused by the growth of axisymmetric oscillations of the jet surface, induced by surface tension. Drop diameters exceed the jet diameter.

- **First-wind induced breakup:** The surface tension effect is now augmented by the relative jet velocity between the jet and the ambient gas, which produces a static pressure distribution across the jet, thereby accelerating the break up

process. As in regime 1, the breakup occurs many jet diameters downstream of the nozzle. Drop diameters are about the same as the jet diameter.

- **Second-wind induced jet breakup:** Drops are produced by the unstable growth of short wavelength surface waves on the jet surface caused by the relative motion of the jet and the ambient gas. This wave growth is opposed by surface tension. Break up occurs several diameters downstream of the nozzle exit. Average drop diameters are much less than the jet diameter.

- **Atomization:** The jet is disrupted completely at the nozzle exit. Average drop diameters are much less than the jet diameter.

At lower jet velocities, the growth of the small disturbances on the liquid surface due to the interaction between the liquid and ambient gas is believed to initiate the breakup process [122, 182]. The outcome of the jet breakup process is also dependent on the initial state of the jet as it emerges from the nozzle exit in each of the four regimes. This influence grows as the jet velocity increases. Figure 2.4 shows a classification of the different break-up regimes as a function of the jet speed and further details can be found in the literature (c.f. [182, 122, 118]). The curve in the region A-B-C is marked as the drip-flow regime where only a train of droplets are formed as the spray speed is low, corresponding to the left-most panel in Figure 2.3. At region C, when the jet speed is sufficiently high, the jet is formed and the jet length increases linearly (almost) with the jet velocity, until the first-wind induced break-up point (D) [middle panel of Figure 2.3]. The region C-D is defined as the Rayleigh break-up region, where the surface tension forces are the dominant cause the jet break-up. The drop diameter is of the order of the jet diameter. At the critical jet velocity, V_c, the longest possible jet is formed. After V_c any further speed increase results in a shorter break-up lengths, until point F is reached. This region D-E-F is defined as the wind induced break-up region [refer to second and third panel in Figure 2.3]. Beyond point

F, the atomization region begins, characterized by large ambient interactions where only the dispersed spray is detected [right most panel in Figure 2.2]. The droplets appear to be stripped off from the surface rather then pinched off by segments. As the jet velocity increases beyond point G the entire jet is completely atomized except near the nozzle tip. The average radius of the droplets in the spray decreases with the inlet jet velocity.

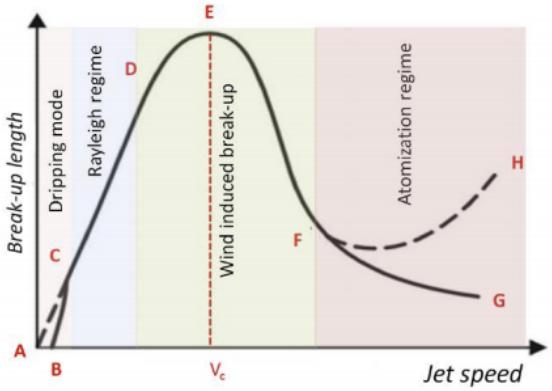

Figure 2.4. Classification of spray break-up regimes as a function of jet velocity [Adopted from [122]]

2.2 Nozzle Non-dimensional Coefficients

Spray nozzles are often characterized by non-dimensional coefficients, that determine the performance of the nozzle, and give a sense of the physics in the internal flow [161, 163]. For example, in case of a cavitating nozzle, the C_d or Coefficient of Discharge is expected to have a lower value than a corresponding non-cavitating nozzle. Studies by Payri et. al. [161] suggested, spray momentum controls both fuel stream penetration inside the combustion chamber and the simultaneous mixing with air. Furthermore, such processes have a direct impact on the characteristics of the flame that forms after the onset of combustion reactions, both in terms of local temperatures and pollutant formation. Thus, momentum flux, spray tip penetration, mixture composition and temperature are closely linked. Below are the mathematical

representations of important functional relationships between flow rate and various coefficients.

$$\dot{\mathrm{m}} = \int \rho \mathrm{udA} \tag{2.1}$$

$$\dot{M} = \int \rho u^2 dA \tag{2.2}$$

$$C_d = C_a C_v \tag{2.3}$$

$$C_a = \frac{\dot{m}^2}{\dot{M}^2 A \rho} \tag{2.4}$$

$$U_b = \sqrt{\frac{2\Delta P}{\rho}} \tag{2.5}$$

$$U_n = C_v U_b \tag{2.6}$$

$$\bar{U} = C_a U_n \tag{2.7}$$

$$K = \frac{P_1 - P_v}{P_1 - P_2} \tag{2.8}$$

where $\dot{m}$, is the mass rate, $\dot{M}$ is the momentum rate (flux), ρ is the fuel density, u is the nozzle exit velocity, A is the nozzle area, C_d is the coefficient of discharge, C_v is the velocity coefficient, C_a is the area coefficient, U_b is the Bernoulli velocity, U_n is the normal velocity, ΔP is the pressure difference characterized as $\Delta P = P_{injection} - P_{ambient}$, K is the cavitation number, P_1 is the local pressure, P_2 is the reference pressure, and P_v is the vapor pressure. These non dimensional spray coefficients will be used throughout Chapters 3, 4 and 5 extensively to characterize nozzle performance and its characteristics.

2.3 Factors Influencing Spray Behavior

The behavior of a spray is often influenced by various factors including ambient gas temperature, the nozzle geometry due to imperfections in the manufacturing process, as well as the in-cylinder thermodynamic conditions.

Before we delve deeper into this topic, a few relevant terms for characterizing spray behavior are introduced. **Liquid Length** is defined as the maximum axial penetration distance of the liquid phase fuel and is considered one among the most important metrics with respect to optimizing in-cylinder processes in engines. **Spray angle** or dispersion angle, is defined as the growth rate of the spray caused by entrainment of ambient air. **Vapor Penetration** is defined as the maximum axial penetration distance of the vapor phase fuel. Typically for a reacting flow condition, the vapor penetration is longer than the liquid length and is an important criterion for spray modeling. In the following section the effect of ambient conditions and nozzle geometry are discussed as these have leading order effect on the spray behavior including its propensity to cavitate.

The injector dynamics are often transient, and is influenced by the motion of the injector needle itself. The needle lift has a major influence in the fluid behavior within the nozzle and hence the ensuing downstream spray. At low needle lift conditions, the spray tends to exhibit the Coanda effect in sticking to the needle surface and creates a pressure gradient within the sac. This causes the spray characteristic at the exit of the nozzle to exhibit higher vorticity, thereby exhibiting transient characteristics [141, 165]. During this time, the area in the needle seat region is lower or equal to the nozzle exit area. As the needle lift furthers, and the needle seat region becomes larger than the exit area, the flow stabilitizes and the injection reaches a quasi-steady state condition. Similarly as the needle lift slowly ramps down, the swirling effect is observed at the nozzle exit and the flow becomes transient again. The quasi-steady

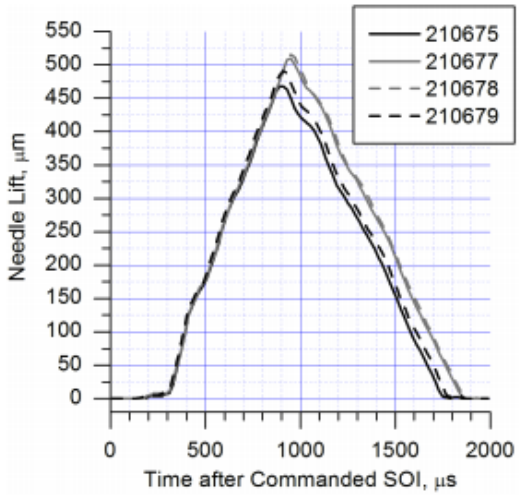

Figure 2.5. A typical diesel spray injector trajectory shows the magnitude of variation in the needle motion. This needle dynamics have a significant impact on the spray growth and development. [Image adopted from [105]]

period of injection is the largest phase of injection in the spray. A typical time varying needle lift curve for a converging nozzle injector is shown in Figure 2.5.

2.3.1 Ambient Conditions

The liquid length is required to promote the fuel-air mixing, however if the liquid fuel impinges on the cylinder wall it could lead to greater emissions, especially with respect to soot formation. Siebers et al.[204, 205, 151], conducted seminal studies to better understand the factors affecting liquid length, especially for high We number (diesel like) sprays. In these studies, the authors concluded that the liquid length decreases linearly with orifice diameter and approaches negligible values as the diameter tends to zero. This is further elucidated from the mass flow rate equation (refer to the Eq 2.5 above) where the mass flow rate, $\dot{m}$, is directly co-related to the nozzle area, A. Although surprisingly, the injection pressure has no discernible effect on the liquid length, which can be attributed to the fact that square root of ΔP is directly correlated to the mass and momentum of the spray and therefore a large change will cause only a small change in flow momentum, and thereby minimally

affecting the liquid length. The study shows that the ambient gas conditions have a significant impact on the spray characteristics such as the growth rate, spray angle and the liquid length. Experimental observation shows that spray angle increases and liquid length decreases with elevated ambient density [204, 151, 102, 160, 162]. As the ambient temperature of the gas increases, the vaporization of the fuel is introduced and needs to be accounted for in models. In diesel operating conditions, the ambient temperatures can be as high as 900-1000 K. This is well beyond the boiling point of the diesel fuel, so vaporization occurs quickly. Studies [236, 204] have shown that sprays penetrate much slower and disperse less under vaporizing conditions. Dahms et al.[43] noted that for diesel type injections under some conditions, a distinct gas-liquid interface no longer exists. This leads to diminishing of the surface tension effects, and the jet appears to undergo a continuous change of state with mixing being diffusion dominated. In these conditions, it would appear that atomization is better described as a single-phase, diffusion dominated mixing process, rather than the classic two-phase spray atomization mechanism which shows how single phase Eulerian models are a good bet to model such flows [16]. This is because single-phase Eulerian models (introduced in the next section) such as Σ-Y predict the atomization process as turbulent mixing in a variable density complex multi-phase mixture. While the experimental studies were conducted for single component fuels, the liquid length of a multi-component fuel can often be determined by the behavior of lowest volatility fraction component[204].

Naber et al. [151] found the ambient gas has a major effect on spray penetration as an increase in ambient density leads to larger air entrainment and therefore slower penetration, based on the laws of conservation of momentum. This behavior is further enhanced when the ambient temperatures are higher, which causes higher rates of vaporization of the fuel, thereby reducing the spray penetration, confirming the observations from other studies [204, 43]. Naber et al. [151] reported an increase in

gas density increases the dispersion angle of the spray, which further facilitates higher entrainment, thereby affecting the spray penetration for non-vaporizing sprays. This relationship is slightly more complicated in the case of vaporizing conditions. Experimental studies [84, 183, 85] further report vaporizing conditions reduce the spray angle based on the density ratio of the liquid fuel, ρ_f to the ambient air ρ_a.

The biggest potential factor contributing to the decrease in the dispersion angle by fuel vaporization is expected to be a contraction of the vaporizing spray as a result of cooling of entrained hot gases by fuel evaporation. The most noticeable indication of a contraction by evaporative cooling is the in-cylinder pressure decrease typically noted in engines during the ignition delay period. Siebers [204, 205] noted the vaporization reduces or slows penetration with the effect being most visible at the lower density conditions. This suggests that the practice of using spray penetration data from non-vaporizing sprays to represent vaporizing sprays is only reasonable at the high density conditions.

2.3.2 Effect of Geometry

The nozzle geometry is another important factor that plays a critical role in the development of the spray. The characteristics of a spray emerging from a cylindrical nozzle to one emerging from a converging nozzle is expected to be different. Infact many studies [193, 200, 163, 161], have shown the nozzle geometry has a leading order effect on the propensity for cavitation and have come to the conclusion that converging nozzles tend to suppress cavitation. The degree of convergence is determined by an *ad-hoc* non-dimensional parameter, K_0, defined as

$$K_0 = \frac{d_i - d_o}{L_{nozzle}} \tag{2.9}$$

where d_i is the nozzle inlet diameter, d_o is the nozzle outlet diameter, and L_{nozzle} is the length of the nozzle. A nozzle with a positive value of K_0 is said to be converging,

and therefore expected to cavitate less. The reason is the smoother internal flow behavior in a converging nozzle, compared to sharp turns in cylindrical nozzles leading to flow detachment.

Payri et al. [161] experimentally studied five cylindrical and five converging nozzles under diesel operating conditions, to further elucidate this effect. Converging nozzles on average facilitate a smoother transition from the inlet to the outlet, whereas a cylindrical nozzle leads to an abrupt change in the flow direction. The study confirmed that converging nozzles on average suppress cavitation and reported that cavitation in case of the cylindrical nozzles, which noticeably reduces the discharge coefficient. They further reasoned that due to the sudden changes in flow direction due to sharp inlet corners, causing the fluid flow to detach thereby creating a low pressure region around the inlet corner which further promotes cavitation. The experiments indicated that the spray momentum flux is largely unaffected by cavitation, because while the effective area decreases due to cavitation, it leads to a proportional increase in the effective velocity, due to the conservation of mass.

Studies [210, 105] have shown the effect of nozzle geometry on the ensuing spray and reported that conicity and hydrogrinding reduce cavitation and turbulence inside the nozzle orifice. This, they further reported, slows down primary breakup, increases spray penetration and reduces the dispersion angle. All of this leads to a slower vaporization, and fuel/air mixing, thereby affecting the flame lift-off length. In addition to the overall shape of the nozzle inlet, localized defects arising due to fabrication has a great deal of influence in spray growth and development. These localized regions of inhomogeneity creates a conducive environment for cavitation development [193, 105]. A feature of the localized defect is shown in Figure 2.6.

Numerical studies [17, 175, 176] have reported the differences in orifice-to-orifice variability due to the manufacturing defects in multi-hole nozzles. They seem to significantly alter the spray angle as well as onset of cavitation [17]. Duke et al.[55]

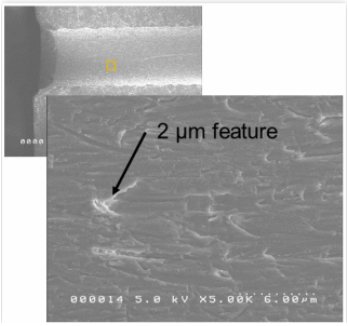

Figure 2.6. A Scanning Electron Microscope (SEM) image obtained from investigating irregularities in the nozzle geometry reveal the small scale local geometric features that have the potential to affect the internal flow in a nozzle [Adopted from [141]].

reported that plume-to-plume variations in the mass fluxes from the holes can cause large-scale asymmetries in the entrainment field and spray structure. Both internal flow transients and small-scale geometric features can have an effect on the external flow. The sharp turning angle of the flow into the holes also causes an inward vectoring of the plumes relative to the hole drill angle, which increases with time due to entrainment of gas into a low-pressure region between the plumes. These factors increase the likelihood of spray collapse with longer injection durations. Kastengren et al.[105, 102] studied single axial hole nozzles and reported the offset of the nozzle holes from the axis of the needle and the sac. Their study further reported that this creates an assymmetry in the inlet condition of the nozzle hole, which varies from nozzle to nozzle. The needle opening and closing also induce transients in the nozzle flow. It has been documented in the past by experiments[172] that needle location impacts the swirling upstream flow and the needle lift and has been discussed in a previous section.

Experiments [200] studied the fuel nozzle geometry effects of cavitation for the Engine Combustion Network (ECN) injectors with a focus on studying the Spray C injector at diesel like operating conditions. To elucidate the effect that this asymmet-

ric inlet condition has on cavitation formation during operating conditions, the fuel flow inside the nozzle hole was observed using high-speed x-ray phase contrast imaging. These images reveal the formation of an asymmetric sheath of fuel vapor that persists throughout the injection event. Complementary x-ray and optical diagnostics of the downstream fuel spray further highlight the effect that this cavitation layer has on the spreading angle of the spray in comparison to the non-cavitating Spray D injector. A few multi-dimensional numerical investigations [224, 75] have confirmed these observations. Magnotti et al. conducted numerical studies [128] demonstrating cavitation and cavitation-induced erosion in fuel injectors in regions of high acceleration and low pressure, thereby causing significant damage to structural integrity of the fuel injector itself.

2.4 Modeling Strategies

While there are a few different approaches currently relevant to identifying and characterizing the physics of sprays, it is primarily dependent on the degree of resolution (or details) that the model aims to resolve, for the given problem. Figure 2.8, shows the plot of energy spectrum, $E(\kappa)$ as a function of the wave number, κ, (more on this can be found in the literature [169, 117]) and regions where different modeling strategies operate. For example, resolving all the scales in a turbulent flow can be achieved by solving the fluid momentum equation, Navier-Stokes, directly. However, as the Re increases so does the computational cost (scales *non-linearly*) for performing such a simulation. For Re of engineering interest, these computations are often intractable. Therefore, modeling frameworks in which the large scales are resolved and smaller scales modeled have grown in popularity. These so called *engineering-level* models are based on the assumption of universal equilibrium range of the small scales [173]. An overview of current methods are presented in this section, first starting with full order models, followed by reduced order representations

and lastly briefly discussing physics-inspired data-driven modeling methods that are growing in popularity. Figure 2.7 details the available methods that will be discussed over the next few chapter and the main emphasis of our contributions in this thesis are in the middle and the right most panels.

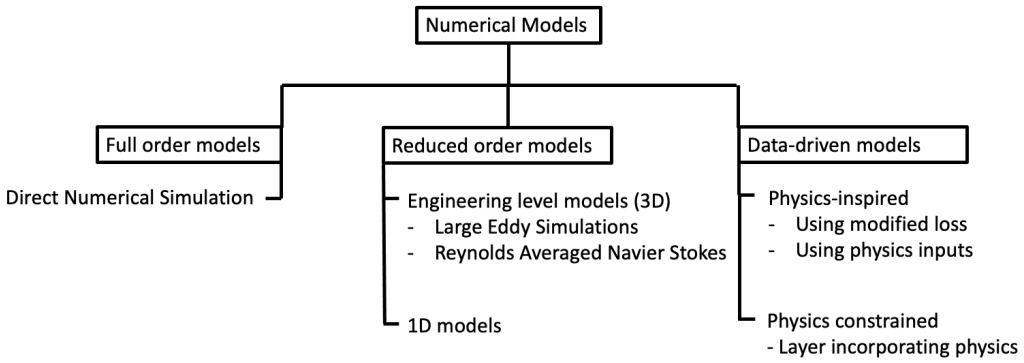

Figure 2.7. Classification of various modeling strategies relevant to ICEs

2.4.1 Full-Order Models

Full Order Models here are defined as models that solve the Navier Stokes equation directly, resolving the large scale features. So in a way, suggesting these methods are models is infact a misnomer. The main advantages of these methods are that these are more accurate compared to reduced order models (introduced later). However, these methods are computationally expensive and often require large High Performance Computing (HPC) resources to successfully run these simulations. These simulations often have a large turnaround time due to higher degrees of freedom and the extremely large number of computations carried out, and require large storage of data depending on the frequency of output. Primary example include Direct Numerical Simuation (DNS). Figure 2.8 provides a summary of the region of applicability of these models based on the energy spectrum - wave number plot.

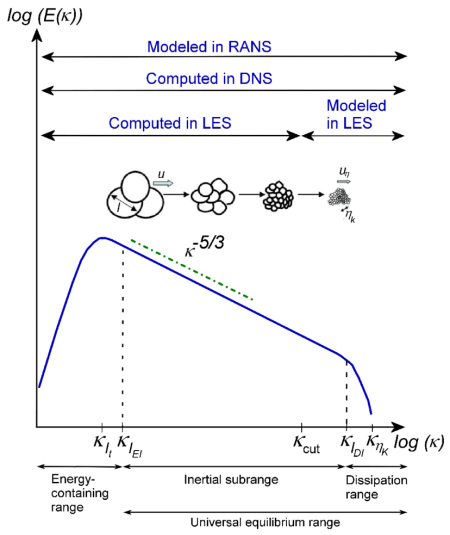

Figure 2.8. The energy cascade figure shows the typical energy-wave number plot and regions where different high-fidelity modeling strategies operate [Adopted from [169]].

2.4.1.1 Direct Numerical Simulation

In a Direct numerical simulation (DNS) the Navier–Stokes equations are numerically solved without any turbulence model. All the spatial scales of the turbulence must be resolved in the computational mesh, from the smallest dissipative scales (Kolmogorov microscales [227]), up to the integral scale L, associated with the motions containing most of the kinetic energy. These solutions in turn have the promise to improve our understanding of the spray physics, including atomization - primary and secondary breakup, droplet dynamics among others. Since most of the atomization occurs at low Mach numbers (ratio of flow velocity to the speed of sound) and often the two fluids are immiscible, the flow is governed by the unsteady variable density Navier-Stokes formulation, within the incompressibility limit. Mathematically the momentum and mass conservation equations can be represented as:

$$\frac{\partial \rho u}{\partial t} + u \cdot \nabla \rho u = -\nabla p + \nabla \cdot (\mu(\nabla u + \nabla^T u)) + \rho g + T_\sigma \tag{2.10}$$

$$\nabla \cdot u = 0 \tag{2.11}$$

here u is the velocity, p, is the pressure, μ is the dynamic viscosity, g is the gravitational acceleration, and T_σ is the surface tension force which is a Heaviside function with values that are non-zero at the phase interface. Although direct numerical simulation (DNS) offers the potential to study the physical processes during primary atomization in detail, thereby supplementing experimental diagnostics, it also introduces severe numerical challenges. The key challenges in using Direct Numerical Simulation (DNS) widely are the numerical requirements, as well as in the case of multi-phase flows the phase interface and the surface tension forces pose a major challenge in wide usage of this approach. However, these limited DNS simulations [12, 13, 149, 72, 7, 45] are often times useful for validating low-compute approximate models for physical processes such as for primary atomization, droplet breakup among others.

2.4.2 Reduced-Order Models

This section will be dividied into two sections. First the engineering level reduced order models will be introduced and discussed, and in the latter part the one-dimensional models would be introduced.

2.4.2.1 Large Eddy Simulation

The principal idea behind the Large Eddy Simulation (LES) modeling paradigm is to reduce the computational cost by modeling the smallest length scales, which are the most computationally expensive to resolve, via low-pass filtering of the Navier–Stokes equations. Such a low-pass filtering, which can be viewed as a time- and spatial-averaging, effectively removes small-scale information from the numerical solution. Performing such filtering divides the Navier Stokes into two parts - resolved and

unresolved terms. It is this unresolved terms that are a subject of great interest by the turbulence community and the problem therein is generally referred to as a turbulence closure problem. This unresolved term information is not irrelevant and its effect on the flow field must be modeled using sub-grid models, one of which is presented below. LES simulations are often computationally less expensive than DNS and the cost scales with the mesh resolution. In the past, LES studies [238, 10, 77, 168] have successfully modeled the phenomena of primary atomization.

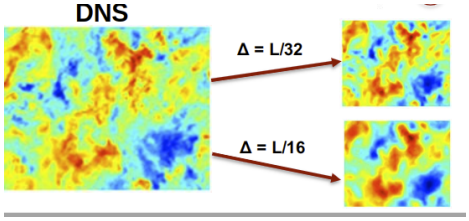

Figure 2.9. The low-pass filtering operation on a high-fidelity simulation is dependent on the grid size. For a larger grid, the details appear to be smeared [Adopted from [4]].

The LES-filtered governing equations (using Favre-averaging) for the balance of mass and momentum are as below:

$$\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial}{\partial x_j}(\bar{\rho}\tilde{u_j}) = 0 \tag{2.12}$$

$$\frac{\partial(\bar{\rho}\tilde{u_i})}{\partial t} + \frac{\partial \bar{\rho}\tilde{u_i}\tilde{u_j}}{\partial x_j} = \frac{\partial}{\partial x_j}[\bar{\rho}\bar{\nu}(\frac{\partial \tilde{u_j}}{\partial x_i} + \frac{\partial \tilde{u_i}}{\partial x_j}) - \frac{2}{3}\bar{\rho}\bar{\nu}\frac{\partial \tilde{u_k}}{\partial x_k}\delta_{ij} - \rho\tau_{ij}^{sgs}] - \frac{\partial \bar{p}}{\partial x_i} + \bar{p}g_i \tag{2.13}$$

where u represents the velocity, p is the pressure, ρ the fluid density, ν the dynamic the viscosity and τ the subgrid stress term. The effect of the sub-grid scale appears on the right hand side of the governing equations through the sub-grid scale stresses, τ_{ij}, which are modelled using the Boussinesq approximation [214], and the assumption by Smagorinsky that the smallest scales are isotropic [208]. Based on Prandtl mixing length theory, the subgrid viscosity can be derived in terms of characteristic length

and one velocity scale [129] as follows, therefore helping to close the Reynolds stress term

$$\tau_{ij}{}^{sgs} - \frac{1}{3}\tau_{kk}{}^{sgs}\delta_{ij} = -\mu_{sgs}S_{ij} \tag{2.14}$$

There are various sub-grid models to effect this closure and [28] has review of some of these methods for compressible flows. To discuss all of the possible approaches is beyond the scope of this section.

2.4.2.2 Reynolds Averaged Navier Stokes

Reynolds Averaged Navier-Stokes, or RANS, is one of the most popular engineering reduced-order models currently in practice. Contrary to DNS or LES, Time-averaged Navier Stokes, or RANS aims to model the entire range of scales. Therefore, the choice of turbulence models plays an important role in accurately modeling the physics of interest. RANS are time-averaged solutions to the Navier Stokes equation, where the instantaneous quantity is decomposed into its time-averaged and fluctuating quantities - $\overline{u}$ and u' respectively.

$$\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j} = \frac{1}{\rho}\frac{\partial}{\partial x_j}(-\overline{p}\delta_{ij} + \mu\frac{\partial \overline{u}_i}{\partial x_j} - \rho\overline{u_i'u_j'}) \tag{2.15}$$

where, u is the velocity, ρ is the density, p is the pressure, δ is the Kronecker delta function, μ is the viscosity term. The bar $\bar{()}$ is used to indicate time-averaged quantities. $\tau_{ij}' = -\rho\overline{u_i'u_j'}$ is the Reynolds stress term, that is used to close the system of equations. This stress term is a derived quantity that is obtained from models and assumptions (such as the Boussinesq approximation)[65, 36, 136, 73]. Due to the multi-phase nature of the physics, additional complexity is introduced in the modeling framework. Many simulations have successfully used RANS models to study different aspects of the spray modeling [141, 224, 188].

2.4.2.3 Multi-phase modeling

Previous studies [176, 16] have discussed common multi-phase modeling strategies in great detail, and therefore only key ideas will be presented in this section. Interfacial multiphase flow for immiscible fluids requires an accurate representation of the interface separation between the two fluids, as well as the ability to model the phenomenon of fluid coalescence and breakup, which further increases the computational complexity of these methods. Details about these popular models including volume-of-fluid (VOF), coupled level set volume of fluid (CLSVOF) and front tracking models can be found in the literature [241, 229, 184, 67, 40, 154, 71]. While each method has their advantages and limitations, this section will restrict the discussions to the single-phase Eulerian modeling paradigm.

2.4.2.4 Single Fluid Diffused Interface Model

Although outside the scope of this brief review, modeling paradigm such as Lagrangian-Eulerian (LE) has found a great deal of interest in the CFD community. These models however have some massive shortcomings. In case of the LE models, some of the shortcomings include:

- Due to the spray dense core region, it is non trivial to model the liquid phaseas Lagrangian parcels. Most models make an (ill-defined) assumption that theliquid is injected as droplets that are the size of the nozzle diameter [201]. Thisinhibits the applicability of the LE approach in a spray-to-tip fully predictive-modeling paradigm

- The computational complexity increases non-linearly with the number of droplets,due to its varied interactions. Some studies have attempted to couple similardroplets together into one greater computational parcel, however it has beenshown that thisad-hocmeasure affects the model predictions [133, 67]

- Due to the coupling mechanism between the Eulerian phase (local gas velocity)to the droplet formulation, there is a severe mesh dependency that cause overprediction of droplet drag [244]. This mesh dependency manifests itself in theturbulence intensity and diffusivity as well, for larger mesh cells [266]

Further, Schmidt etal. [231, 232] shows a nonlinear growth rate for the mean expected error for a simulation where parcel sizes are held constant and the Eulerian mesh is subsequently refined, which leads to lack of convergence for these LE methods. Improved convergent diesel spray LE simulations using an improved particle-tracking algorithm [12]has been successfully demonstrated for large engine simulations [237, 238], however these methods add to the complexity of the setup.

While the numerical limitations seen for LE is minimized by the EE paradigms. Pai et al. [156] studied the two popular statistical representations of multi-phaseflow in Lagrangian-Eulerian and Eulerian-Eulerian frameworks, in an effort to estab-lish the exact form of the relationship between the two statistical representations.Their analysis, based on probability density function or pdf, given the statisticalvariability of the LE and EE methods, showed that the fundamental quantities ofboth representations bear a simple relationship with one another only when the two-phase flow is homogeneous.

The main characteristics of the diffuse interface single-field Eulerian approach is that it makes no attempt to resolve or capture the interface. Instead, the effects of the interface are modeled and a density-averaged (Favre-averaged) flow field is solved on a lower resolution grid. This approach works well for low Stokes number flows, where the dispersed phase naturally follows the streamlines of the continuous phase. It is also a fitting approach for flows with diffusion-dominated mixing [43]. The goal of this model is to realistically describe the physics of the dense core region and its atomization, regions where LE models face significant challenges. In this region,

the explicit resolution of the liquid-gas interface movement can be computationally intractable for engineering applications.

Alternatively, this liquid-gas interface can be resolved using an empiricism-based Eulerian methods, such as in the $\Sigma - Y$ model [232, 35]. The $\Sigma - Y$ model attempts to characterize the liquid spray with two scalar quantities, the liquid volume fraction, $\bar{Y}$ and the interface area density, $\bar{\Sigma}$. The evolution of the scalar quantities are described using transport equations which contain the modelled source terms that describe the atomization process. While the original framework was developed for Reynolds Averaged Navier Stokes (RANS) formulation, it has been extended successfully to Large Eddy Simulations (LES). With the knowledge of $\bar{Y}$ and $\bar{\Sigma}$ it is possible to determine the Sauter mean diameter (SMD) and the droplet density using simple algebraic formulations. For the Eulerian Lagrangian Spray Atomization (ELSA) model [111, 237], the SMD and spatial distribution of droplets from the $\Sigma - Y$ model is used as inputs for the secondary Lagrangian spray model, used in the disperse region downstream of the dense core region. The $\Sigma - Y$ model is based on four main assumptions:

- The large scale features of the two-phase flow are independent of the effects of surface tension and viscosity at infinite Reynolds number (regime for atomization), as they act at small length scales.

- Since it is non-trivial to recover the random velocity field of a two phase flow at each time, it is proposed that the mean values from two-equation turbulence models such as $k - \epsilon$ be used.

- The dispersion of the liquid phase can be computed through a transport equation of the liquid mass fraction containing a turbulent diffusion liquid flux term to account for mixing.

- The mean droplet size of the dispersed liquid phase can be predicted by modeling the mean surface area of the liquid/gas interface per unit volume. This term quantity is also known as Σ.

The momentum equation for large Reynolds number, neglecting viscosity in the liquid and/or the gas phase, can be written as:

$$\frac{\partial \bar{\rho}\tilde{u_j}}{\partial t} + \frac{\partial \bar{\rho}\tilde{u_i}\tilde{u_j}}{\partial x_i} = -\frac{\partial \bar{p}}{\partial x_j} - \frac{\partial \bar{\rho}\widetilde{u_i' u_j'}}{\partial x_i} \tag{2.16}$$

and the continuity can be expressed as:

$$\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial \bar{\rho}\tilde{u_i}}{\partial x_i} = 0 \tag{2.17}$$

the terms have their usual meaning introduced previously.

Similar to the single phase turbulent flow, the momentum equation cannot be solved without modeling the correlation terms. Typically a turbulence model with a Boussinesq eddy viscosity assumption can be used to close the Reynolds stress tensor on the right hand side. While one of the model assumptions states that the resolved momentum for the mixture can be defined by one bulk velocity, the slip velocity can be expressed in the form of the following equations [46]. The mass transport equation is as below

$$\frac{\partial \bar{p}\tilde{Y}}{\partial t} + \frac{\partial \bar{p}\tilde{u_i}\tilde{Y}}{\partial x_i} = -\frac{\partial \bar{\rho}\widetilde{u_i' Y'}}{\partial x_i} \tag{2.18}$$

where $\widetilde{u_i'Y'}$ is the term that accounts for the mixing effect of the relative velocity between the fuel and the gas. The term Y is the liquid mass fraction. The classical closure expression by the analogy with Fick's law yields,

$$-\bar{\rho}\widetilde{u_i'Y'} = \frac{\mu_t}{Sc_Y}.\frac{\partial \tilde{Y}}{\partial x_i} \tag{2.19}$$

where Sc_Y is the Schmidt number, μ_t is the turbulent viscosity. These unknown quantities, $\tilde{Y}$, $\tilde{u_i}$, $\bar{p}$, $\bar{\rho}$ require one for equation each for closure. This is obtained by the relationship between $\tilde{Y}$ and density. Under the assumption of an immiscible mixture, the relationship can be modeled as follows:

$$\frac{1}{\tilde{\rho}} = \frac{\tilde{Y}}{\rho_l} + \frac{1-\tilde{Y}}{\rho_g} \tag{2.20}$$

the suffixes l, g indicate liquid and ambient gas quantities.

As per the fourth assumption, the mean droplet size of the dispersed liquid phase can be modeled by the balance equation of Σ, the mean surface area of the liquid/gas interface per unit volume. This balance equation was developed by [232, 26].

$$\frac{\partial \bar{\Sigma}}{\partial t} + \frac{\partial \bar{\Sigma}\tilde{u_i}}{\partial x_i} = \frac{\partial [D_s(\frac{\partial \bar{\Sigma}}{\partial x_i})]}{\partial x_i} + (A+a)\bar{\Sigma} - V_s\bar{\Sigma}^2 \tag{2.21}$$

where D_s is an appropriate diffusion coefficient, Σ is the interfacial area density term, $1/A$ and $1/a$ are two different production time scales, V_s is a destruction coefficient with the same dimension as of velocity. The physical phenomenon associated with the term A is the stretching of the liquid-gas interface by the mean velocity gradients. This term can be modeled to use the same time scale as in the production term for the turbulent kinetic energy, and represented as:

$$A = \alpha_0^{'} \frac{\widetilde{u_i^{'} u_j^{'}}}{\tilde{k}} . \frac{\partial \tilde{u}_i}{\partial x_j} \tag{2.22}$$

where $\alpha_0^{'}$ is a modeling constant. k is the kinetic energy. The term a is associated with the effect of the turbulent flow field on the stretching of the liquid-gas interface and can be modeled as the integral characteristic time scale of turbulence itself, as follows

$$a = \alpha_1 \frac{\tilde{\epsilon}}{\tilde{k}} \tag{2.23}$$

where ϵ is the dissipation rate and α_1 is a constant. For more details about the $\Sigma - Y$ model, the interested reader is referred to the seminal work in the literature [232, 35, 26, 27].

2.4.2.5 Lagrangian Switch

While the $\Sigma - Y$ model can essentially capture the liquid and vapor penetration as well as the droplet density profile for high Weber number flows [189], the Lagrangian approach is suitable for applications related to combustion modeling. As discussed earlier, the Lagrangian models are superior in regions of low-density, or dispersed spray. The Eulerian formulation is able to describe the regions of disperse spray by accounting for the effects of collision, coalescence, and secondary breakup. However, the modeling of the liquid dispersion is accomplished with a gradient-based closure, which is not necessarily valid in the disperse region of the spray [111, 112, 9]. This makes it pertinent to explore switching to a Lagrangian framework for the dispersed region of the spray.

While there are many methods to switch the formulation, the most popular ones include the Eulerian cell liquid volume fraction. The transition is triggered when

the liquid volume fraction is below a particular threshold, a value that has been empirically ascertained and is sensitive to the fuel/gas properties.

$$\tilde{Y_l}\frac{\bar{\rho}}{\rho_l} \leq \tilde{\phi_l^{crit}} \tag{2.24}$$

where Y_l is the liquid mass fraction, ρ is the equilibrium density, ρ_l is the liquid density, ϕ is the volume fraction, and suffixes l indicates liquid phase. The transition zone is composed of the computational cells that form the border with the dense zone (i.e. the zone where the liquid volume fraction is greater than 0.01) and only one parcel is generated per transition cell and per time step. Further more the ELSA model is used to determine the diameter of the droplets, calculated using the Sauter Mean Diameter formulation

$$D_{32} = \frac{6\tilde{Y_l}}{\rho_l \Sigma} \tag{2.25}$$

D_{32} is the Sauter Mean Diameter (volume/Area of a sphere). The other terms have the same meaning as previously introduced. The number of droplets generated is obtained by mass conservation:

$$n_{drop} = \frac{\bar{\rho} Y_l \tilde{V}_{cell}}{\pi/6 \rho_l D_{32}^3} \tag{2.26}$$

where V_{cell} is the cell volume, and the droplet velocity is extracted from the mean flow itself [119]

2.4.2.6 One-Way coupled Lagrangian approach

In addition to all the above methods, the one-way coupled Lagrangian approach uses as input the results from a well resolved internal flow simulations measured at the nozzle exit to a Lagrangian spray model for external spray simulation. This is because in the case of gasoline conditions (and this can be extended to diesel relevant conditions), internal nozzle flow simulations have shown to capture the phenomena of phase change successfully, as well as capture injector transients due to needle motion, which affects the momentum flux. The idea is to use in-nozzle flow simulations and the results at the nozzle exit can be used to initialize a Lagrangian spray calculation. This coupled approach accounts for the presence of phase-change induced voids at the nozzle exit and hence is expected to be more predictive compared to the standard blob injection model (which does not account for the in-nozzle phenomenon), therefore creating a more predictive tool relevant for modeling transient spray behavior [187, 188]. However this limited approach, suffers from the issues of solution convergence discussed previously.

2.4.3 One-Dimensional models

In the previous section, we discussed some of the popular full order models relevant to spray modeling. However, these reduced order models have inherent assumptions and constants/parameters that are derived empirically or by fitting to data from canonical experiments. Lower dimensional (1D/2D) reduced order spray physics models are simplified versions of the high-fidelity, complicated, multi-dimensional full order models built using first principles that enable a straightforward identification of the influence of boundary parameters on the macroscopic results. As these models are relatively computationally inexpensive, they are useful in the design space exploration and parameter sensitivity studies.

There have been many attempts at building reduced order models including useful correlations for spray physics in the past [85, 80, 47, 235, 11]. However, often times these models were too simplistic, or were only applicable to a very limited set of conditions. The paradigm shifted with some seminal experimental studies conducted by Siebers and co-authors [204, 205]. Siebers et al. [204, 205, 151] conducted experiments to study the nature of diesel sprays, proposing a scaling law for maximum penetration distance of the liquid phase fuel. This scaling law, which was an improvement over previous attempts at defining a correlation [11, 85], is based on applying jet theory to a simplified model of spray. It accounts for the effect of injector characteristics, fuel properties and in-cylinder thermodynamic conditions on the liquid length, and provides a deeper insight into the process of vaporization of single component fuels. Siebers additionally suggested, that this model can be used for multi-component fuels through the use of single component surrogates.

The Siebers model, based on mixing-limited hypothesis, provided excellent agreement with experiments over a wide range of conditions. The mixing limited hypothesis assumes that interfacial details such as interphase mass and energy transport at droplet surfaces are not the limiting factors in the mixing of the fuel and gas. Rather this process was limited by the process of turbulent mixing due to entrainment of high-temperature ambient air throughout the spray cross-section - hence the term, *'mixing-limited'*. The energy for vaporization of the fuel comes mainly from the high temperature entrained air into the spray. This is evident from the experimental observations that the liquid length does not shorten significantly after ignition, and vaporization occurred largely upstream of the combustion zone. The authors further concluded, that the liquid length of the jet is independent of the injection pressure, and is only dependent on the orifice size, density of the fuel and ambient gas as well as the ambient temperature. Building on the jet theory, vaporization limited by mixing

processes yielded the following relationships for mass of ambient air entrained and the fuel mass flow rate during the steady state conditions:

$$\dot{m_a}(x) = f(\sqrt{\rho_a \cdot \rho_f}) \cdot d \cdot U_f \cdot tan(\theta/2)) \tag{2.27}$$

$$m_f(x) = f(\rho_f \cdot d^2 \cdot U_f) \tag{2.28}$$

$$B = \frac{Z_a(T_s, P_a - P_s).P_s.M_f}{Z_f(T_s, P_s).[P_a - P_s].M_a} = \frac{h_a(T_a, P_a) - h_a(T_s, P_a - P_s)}{h_f(T_s) - h_f(T_f, P_a)} \tag{2.29}$$

$$L = \frac{b}{a}\sqrt{\frac{\rho_f}{\rho_a}}\frac{\sqrt{C_a d}}{tan(\theta/2)}\sqrt{(\frac{2}{(F/A)_{liq}} + 1)^2 + 1} \tag{2.30}$$

where a and b are constants with suggested values of 0.41 and 0.66 respectively. The suffixes f and a denote fuel and ambient air respectively. The terms Z_f and Z_a are vaporized fuel and ambient gas compressibilities respectively, and M denotes molecular weights, P is pressure, T is temperature. The unknown T_s can be solved iteratively given the fuel and ambient gas properties as well as the initial conditions. Once obtained, T_s defines the pressures, temperatures, and enthalpies of the fuel and ambient gas at the liquid length location. B is the fuel and ambient gas mass flow rate ratio. For the full derivation, please refer the original manuscript [204, 151, 205]. Figure 2.10 shows the schematic of the mixing limited physics for high We flows. For a detailed analysis please refer to the original manuscript [205].

In spite of the successes of this model in validation against experimental observations, there are certain inherent limitations of this framework. While the spray angle remains constant during the quasi-steady phase, during early/late transients Siebers reported deviations from the constant spread. This transient effect is not accounted for by the current approach. Additionally, this model operates under the assumption of non-vaporizing sprays and constant spray angle. This reduced order model is in itself a standalone framework and not coupled to a 3D Navier Stokes solver to account for gas effects.

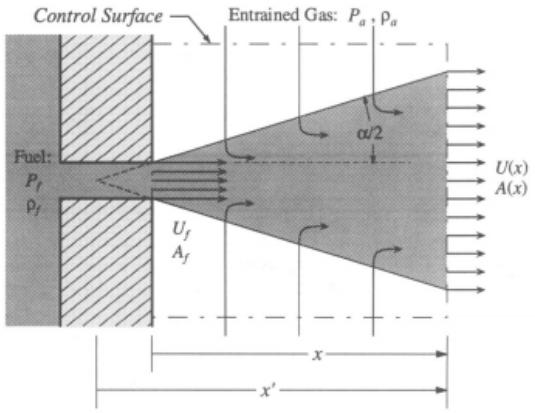

Figure 2.10. Schematic of the mixing limited liquid length model shows the simplified control surface that accounts for the entrained ambient gas and the subsequent turbulent mixing with the liquid phase fuel that acts as a limiting factor in the spray behavior [Adopted from [205]].

Some studies that built on the success of the Siebers framework include the Virtual Liquid Source (VLS) model [6] for vaporizing diesel sprays. The VLS model takes advantage of some important conclusions from the Siebers study, including that the liquid length is affected by the vaporization instead of the energy released due to combustion. Using that knowledge the authors do not explicitly model the liquid phase of the fuel, instead use the liquid length inputs from the Siebers approach in the VLS model. However, a severe limitation of this model comes across for cases where the injected fuel remains in liquid phase for a significant fraction of the total injection duration, such as in the case of low temperature combustion. There is no explicit treatment prescribed for such a scenario.

Other studies built on the Siebers 1D model for liquid penetration and used that as an input to their 3D calculation by imposing conservation of mass, momentum and energy transfers from the liquid to the gas phase applied in the region from the nozzle exit to the liquid penetration length, as predicted by the Siebers model. Since most of the mesh resolution is required near the dense core region, this approach helps to bypass the issue entirely by using a reduced order model for the most compute

intensive region of the model, thereby saving compute time. However like the VLS model study, this model is expected to have severe limitations to only being used for sprays that are valid for the mixing limited characteristics and does not impinge onto the cylinder bowl surface.

The Siebers model's success in predicting the liquid length at steady state conditions, along with its independence of injection pressure, implies that diesel fuel vaporization is controlled by mixing, rather than atomization processes. However, the experimental data used for comparison with model predictions only went as low as 700 K, and Siebers notes that departures from model predictions tend to occur at low-temperature, low-density conditions, which are prevalent during early injection Low Temperature Combustion (LTC). Pickett et. al. [167, 164, 166] observed that under these conditions, liquid lengths are unusually long and therefore have the potential to hit the piston, which leads to detrimental engine performance. Pickett further shows that for an impinging jet, a steady liquid length may not be obtained, and therefore the quasi-steady liquid length formulation of Siebers may have limited relevance. The authors formulated a time to reach the steady state (peak injection) that would help in design decisions.

$$t_{ss} = f(\frac{L}{U_f} . \frac{1}{(F/A)}) \tag{2.31}$$

where L is the liquid length, F, A refer to fuel phase and gas phase respectively. This equation states the mixing times after Start of Injection may increase or decrease based on whether L or $(F/A)_{liq}$ is dominant. Like the steady-state liquid length formulation, t_{ss} depends on the ambient temperature, ambient density, injector orifice (and therefore mass flow rate of fuel), and the fuel itself. This result was specifically useful in determining the effect of injection characteristics (for example injection duration, injected mass) on the possibility for wall impingement at early-injection and low temperature conditions. Pickett suggested multiple injections for short duration

to limit the liquid penetration, although the second injection will tend to penetrate longer than the first as the latter leaves a cool, fuel laden mixture in its wake. Since liquid length is directly proportional to orifice diameter, Pickett suggested using small diameters to limit the liquid penetration for short injection durations.

Musculus and Kattke [150] observed increased mixing after the end of injection in low temperature combustion (LTC) diesel engines, which they termed as an 'entrainment wave'. The authors noted that this wave traveled downstream at twice the initial jet propagation rate and increases mixing by at least a factor of three. They observed that this causes over-mixed regions and spatial shift in onset of soot formation, among other observations. To understand this process better, they proposed a one dimensional model to study transient diesel jets. This model is built on the mixing limited assumption of Siebers. Musculus et al.[150] faced the same limitations in using the Siebers model for modeling end of injection transients. Here they proposed modifying the Siebers model by discretizing the domain in the axial direction and adding transient terms for the jet momentum and fuel mass transport equations in each discretized control volume. Expanding on the steady-jet development of Naber and Siebers [151, 205], the transient transport equations for mass and momentum are solved for the jet geometry in Figure 2.11. The Musculus Kattke model however still assumes non-vaporizing conditions, under the assumption that the original Siebers model predictions were reasonable. Unlike Siebers, the MK model assumes a radial profile for the transverse distribution of mass in the spray. This radial profile of the fuel volume and velocity were based on a parameter α which evolved from a top hat shape ($\alpha = \infty$) at the nozzle to a fully developed (almost parabolic) shape ($\alpha = 1.5$) at some distance downstream. This is similar to the Gaussian error function shape, observed from experiments [52]. The quantities such as mass, momentum, velocity are calculated at the center of the axial midpoint of each control volume and then using the radial profile projected onto the spray cone frustum. While the approach

here is similar to the studies by Pastor et. al.[159, 49, 48], this model is relevant to the transient phase of spray injection. The entrainment rate given by:

$$\frac{\partial \dot{m}_e}{\partial z} = \rho_a \frac{\partial}{\partial z}(\bar{u}A) \tag{2.32}$$

where the $\bar{u}$ predictions are obtained from the model after the end-of-injection, ρ_a is the ambient air density, $\dot{m}_e$ is the mass of the entrained fuel and z is the axial location of the spray, and A the projected surface area.

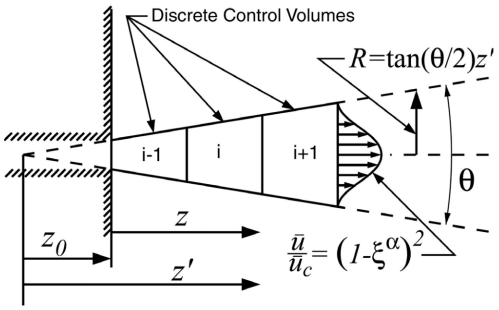

Figure 2.11. Schematic of the One-dimensional discrete control volume transient diesel jet model [Adopted from [150]]

The successes of these so called mixing limited approach for building simplified reduced order models for high We number flows, necessitates the question as to why this regime/assumptions work so well compared to the other ideas proposed in the past. One of the reasons could be that the current engine technologies (high boost and injection pressure, and small nozzle hole diameter informed by the relevant experimental studies [166, 167]) have resulted in a complete atomization regime inside the spray very near the nozzle exit. Furthermore, even under non-evaporating conditions, in which the problem is undoubtedly a two-phase flow, the spray can be analyzed from a point of view of gas jet theory, since droplets upon formation almost immediately reach a dynamic equilibrium with the surrounding gas phase.

Desantes et al. [49] proposed a one-dimensional evaporative spray model relevant under diesel engine conditions, based on the mixing limited hypothesis making use of the non-ideal gas state relationship descriptions for the whole spray. Unlike previous attempts, this model has the ability to model the physics of fuel vaporization. This model couples the non-ideal gas state relationships that adequately describe the thermodynamic conditions inside the diesel spray with the one-dimensional spray model [205, 48, 159] to predict the evolution of the evaporating diesel spray, under constant and variable rate of injections [49]. The model features are divided into two separate zones, as shown in Figure 2.12. The zone 1 is where the evaporation is taking place and includes the liquid phase fuel, the vapor phase fuel and the ambient gas. And Zone II is the region where the evaporation is complete, and only vapor fuel and ambient gas are present. The atomization effect can be neglected in Zone I based on the mixing limited assumptions. The model has two major differences from the Siebers approach, including:

- The thermodynamic state relationships, under an equilibrium assumption, are derived independently of the spatial evolution of the spray and they serve as the equation of state of the mixture, so if the fuel mass fraction is known at a certain point in the spray, any thermodynamic property can be known.
- The spray model acts as the conservation vehicle and is coupled to the state relationships which helps to identify the spatial distribution of the spray properties.

For a full derivation, the reader is referred to the original manuscript [49]. Although this evaporative spray model has many advantages over the original mixing limited model, this approach can only be used under inert conditions.

Pastor et al. [159, 48], developed a one-dimensional spray model (refer to Figure 2.13 for a schematic) that is capable of prediction of spray behavior under transient

Figure 2.12. Schematic of the one-dimensional mixing limited vaporizing spray model, dividing the injection into the two separate zones in evaporation (Zone I) and vaporization (Zone II) [Adopted from [49]].

conditions for inert or reacting conditions, based on the mixing-controlled hypothesis and locally homogeneous flow. Although in nature, it seemed inspired by the Siebers model, it differed from it in the sense that it required a uniform discretization (δx)of the domain in the axial direction, with each cell limited by inlet and outlet sections such $x_{i+1} = x_i + \delta x$. The tip penetration is defined as the farthest cell from the nozzle where is the inlet velocity is different from zero and the outlet velocity is zero. The conservation requirement is imposed on each cell and the symmetry on the spray axis is assumed. Further differentiating with the Siebers model, the authors assumed a fully developed turbulent profile, in the form of a Gaussian radial distribution profile. Under the conditions for non-vaporizing sprays, where energy exchanges can be ignored, the local density is assumed under ideal mixing to be:

$$\rho(x,r) = \frac{1}{(\frac{f(x,r)}{\rho_{f,0}} + \frac{1-f(x,r)}{\rho_{a,\inf}})} \tag{2.33}$$

where $\rho_{f,0}$ and $\rho_{a,inf}$ correspond to the pure fuel and pure air densities. Under vaporizing conditions, further assumptions are made to the model including solving an additional equation for conservation of energy (along with mass and momentum) in every discretized cell. The local enthalpy at any given location is calculated by the means of

$$f(x,r,t) = \frac{h(x,r,t) - h_{a,inf}}{h_{f,0} - h_{a,inf}} \tag{2.34}$$

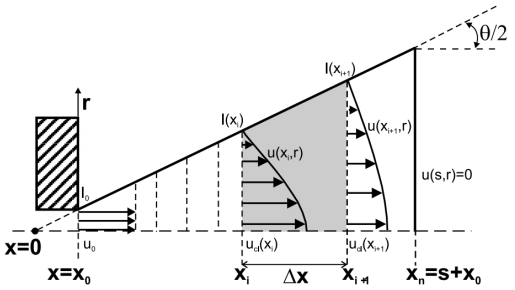

Figure 2.13. Discrete one-dimensional model for reacting flows, showing the similarity in its formulation with the Siebers model except for a few key differences in assumed profile and spatial discretization [Adopted from [159, 48]].

If the local enthalpy and local composition of the fuel mass fraction (fuel includes liquid phase and vapor phase) is known, the local temperature can be obtained assuming an ideal mixture

$$h(T, f) = \sum (Y_i.h_i(T)) \tag{2.35}$$

Finally local density can be obtained for the whole range of fuel mass fraction by

$$\rho(x, r) = \frac{1}{\sum (\frac{Y_i(x,r)}{\rho_i(x,r))}} \tag{2.36}$$

Pastor further extended the model for reacting sprays, based on the above work, by slightly modifying the assumptions to include effects of combustion such as the non-conservative nature of the fuel mass fraction in each cell and the combustion-induced changes in local conditions modifying the state relationships. The discussions for this framework is beyond the scope of this current paper and the reader is referred to the original manuscript [159, 48].

These mixing-limited physics based reduced order models have found great success in characterizing spray statistics under engine-relevant conditions. While these reduced order spray models have proved to be an important tool in the design and discovery of the physics of the sprays, they have some key limitations:

- Most of the mixing-limited reduced order spray models [205, 159, 49, 48, 150] assume constant spreading angle. In reality, experiments have shown that a typical spray injection event has varying spreading angles depending on the needle lift that these family of models fail to take into account.

- These spray models are stand-alone tools to predict liquid length among other characteristics, and are not implemented into a CFD code to incorporate effects of the gas phase source. One expected improvement would be to develop a framework that can model the two way effects of the exchanges between the fuel spray and the ambient gas.

- Many underlying assumptions are made in the spray model especially regarding the entrained air behavior such as a purely radial flow, which full 3D CFD has shown to severely under predicts the entrainment.

- The Siebers [205] model does not take into account the varying rate of injection profiles, which the other models Musculus et al. [150] and Pastor et al. [159], have attempted to resolve somewhat, although most comparisons for validating the model are made to the quasi-steady state behavior.

2.4.4 Physics Inspired Data driven Modeling

Following the discussions in the previous sections about full-order and reduced-order models, in this section we explore the progress in the area of developing data-driven modeling with a focus on fuel spray applications. While the other models rely on physics and conservation laws to build intuition and modeling framework based on first principles, they often have a few different constants that rely on empirical derivation or by fitting to high-fidelity numerical experiments. Since there is an element of data-fitting to the exercise of model building, we explore the potential in using a physics-inspired data-driven modeling in this section. The development of these

physics models were either based on the first principles, relying heavily on the physics of the problem, or is based on simply emulating data as in the case of curve fitting or identifying constants for models, that are purely data-driven. However physics inspired data driven modeling is somewhere in between the two extremes. In this paradigm, the model is rooted in the first principles for its development, identifying important parameters based on the relevant physics of the problem. Furthermore, the relationship between the parameters is determined via data, hence the adage 'physics-inspired data driven methods'. A schematic comparing different approaches can be seen in Figure 2.14.

Scale bridging is a critical need in computational sciences, where the modeling community has developed accurate physics models from first principles, of processes at lower length and time scales that influence the behavior at the high scales of interest. However, it is not computationally feasible to incorporate all of the lower length scale physics directly into up-scaled models. This is an area where machine learning has shown promise, in building emulators of the lower length scale models which incur a mere fraction of the computational cost of the original higher fidelity models. We discuss some of the most recent endeavors in this area below.

Machine learning (ML) can be broadly defined as, the study of computer algorithms that improve automatically through experience[69]. Often times, machine learning is seen as a subset of the broader term, *artificial intelligence* or AI, where a mathematical model is built on a sample representative data, known as training data, to make predictions without being explicitly programmed to do so. For a full review of different machine learning approaches and their effectiveness, the reader is encouraged to review publicly available literature [69, 113, 191].

The interest in ML for scientific applications has seen an uptick in the recent years [143, 37, 100]. Many studies have used ML for fluid-relevant applications

[177, 134, 202, 97, 170] ranging from sub-grid closures for turbulence [78, 171, 134] to purely data-driven models with physics-inspired constraints [170, 202, 177]. Readers are encourgaed to read some excellent reviews on this topic [33, 32, 222, 248, 109]. Scientific ML differs from other applications such as in text, natural language processing, images etc because the scientific datasets are often high-dimensional, complex and sparse. These datasets have underlying physical constraints that need to be met in order to build truly predictive models [217]. A few different categories of the scientific ML work has emerged and these include:

- **Emulator models** for physical processes that build cheap surrogates for complex physical models. These help in exploring the design space for sensitivity analysis.
- Purely **data-driven approach** for model optimization, especially for model constants, relevant for applications such as in fluid turbulence [170].
- Increasingly these approaches have been used for **model discovery**, directly from high fidelity simulation data or from experimental observations, where the physics itself is not well-known.
- **Inverse Design** space exploration for parametric studies on a given set of inputs.

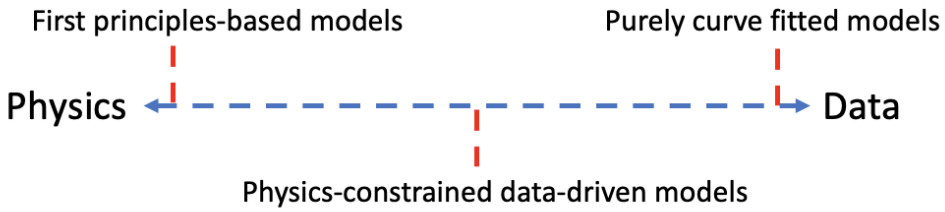

Figure 2.14. The clustering of models on a physics-data curve shows the dominant basis between first principles physics based models and purely data-driven models.

One of the very common approaches in using machine learning for modeling fluid flows is to use a reduced order input such as Proper Orthogonal Decomposition (POD)[39]. Ganti et al. [61, 60] proposed a machine learning surrogate model that was able to predict complicated, spatio-temporal gaseous and spray fields using Gaussian process based machine learning techniques and applied to single and multi-phase flows. The challenges with using a POD based dataset are that these data reduction techniques do not fully resolve the solution and often miss key information. Apart from this, these POD simulations often suffer from a lack of generality. Kuntz [109] further discusses the limitations of the POD models such as the limited ability to capture transient, multi-scale phenomena and their inability to capture invariances. The other challenge is in the fact that the expectation for the model to conserve energy and other critical quantities is not very strictly enforced, thereby limiting its use to only first results for approximation.

Zhang et al. [247], using multi-task neural networks approach, built a predictor and an auto-encoder network to optimize the Kelvin Helmhotz - Rayleigh Taylor (KH-RT) spray break up model constants. The predictor is trained to determine the sub-model parameters θ for a given X and error, ϵ. The optimal θ then can be estimated by setting ϵ as zero. The auto-encoder is used to learn a latent representation of a pair of (X, θ), which is encouraged by a regularization term to share the same latent space as the predictor. For an unseen condition X and estimated optimal ϵ, Zhang et al. showed they can use the auto-encoder to find similar (X,θ) pairs from the training data to interpret the predictor prediction and quantify the uncertainty.

Milan et al. [138] used the dimensionality reduction methods such as proper orthogonal decomposition[39] and autoencoders to evaluate their efficacy in representing a multi-phase flow in a single orifice injector, in a reduced dimensional space based on fuel properties and needle lifts. They determined that an autoencoder with a set of neural networks, in a dense layer encoder-decoder architecture, outperformed the

POD methods and provided a robust representation of the multi-phase flow across different conditions, thereby creating a cheap surrogate to explore the design space.

Although the area of physics-inspired data-driven modeling for sprays is fairly young, it is expected that in the near term more researchers would make use of these data-driven methods for improving the current state of the art in this area. Training a machine learning algorithm is often a stochastic process, due to the sampling of the data during the training process, a multitude of hyper-parameters, and the complicated non-convex loss manifold for high-dimensional scientific datasets [217]. Some critical challenges that need to be explored for a wider adoption of these methods include characterizing the robustness of the machine-learnt solutions, guarantees of convergence, uncertainty quantification of these trained models.

CHAPTER 3

MULTI-DIMENSIONAL INTERNAL FLOW SIMULATIONS FOR CAVITATING NOZZLES

3.1 Motivation

In the previous chapter, various numerical modeling strategies were introduced. In this and the subsequent chapters, we will expand on each one of those ideas to provide examples of original research that used these tools to improve the scientific understanding of the spray physics in the internal nozzle flow, near-field, and far-field regions of spray growth and development.

As discussed previously, the the quasi steady phase of injection is typically the longest phase in an injection cycle. While the transients of the needle motion have an effect on the spray growth and development, it is generally short-lived. Therefore many studies, numerical and experimental, are rightfully focused on understanding and characterizing the quasi-steady phase of injection [141, 225, 131, 245]. To fully understand and characterize the spray, it is important to better resolve the internal flow and thereby the near-field spray region. The injector geometry has shown to have an important effect on the internal flow [105, 128, 161, 193]. Using simplified assumptions for the near-field nozzle exit conditions, as a starting point for external spray modeling often leads to erroneous far-field predictions. Therefore, it is important to use the internal flow simulations in order to inform appropriate boundary conditions fpr external flow simulations. In this chapter, we use high precision nozzle geometry as inputs to the CFD model in an effort to validate the model predictions against experimental observations, providing further insights into the mechanisms

of cavitation within certain nozzles, that hitherto were not possible to extract from highly transient experimental studies (as timescales are on the order of a few μs). In the past, multi-dimensional numerical modeling has relied on simplified geometries that are unable to address the issue of in-nozzle irregularities and lead to predictions that are not accurate. In order to accomplish this, it is imperative the actual nozzle geometry, characterized by high-fidelity X-ray Computed Tomography (CT) scans, be incorporated as inputs to the CFD model. A schematic of our workflow is shown in Figure 3.1.

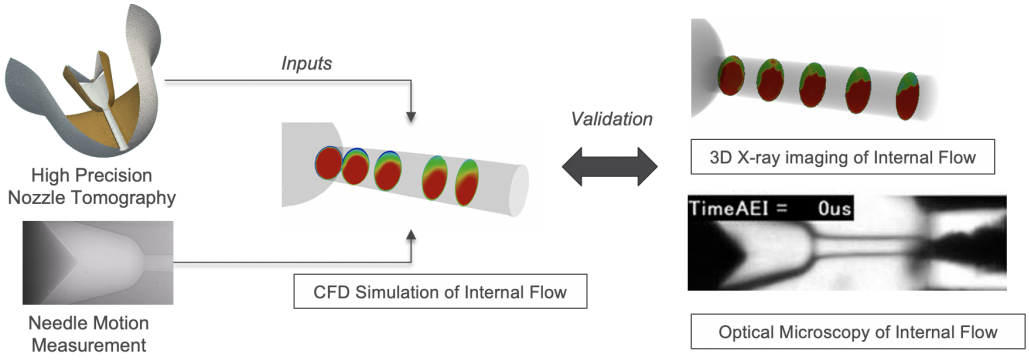

Figure 3.1. A schematic of our workflow in comparing CFD predictions with experimental observations within the nozzle injector (internal flow) as well as in the near-nozzle region. This involves using high-precision nozzle geometry (upto 2 μm spatial resolution) as an input to the multi-dimensional CFD simulation. Image credit: Aniket Tekawade, Argonne National Laboratory

Two separate nozzles, a single-hole cylindrical nozzle and a multi-hole convergent nozzle, are studied to this effect. While fuel injectors in production are generally multi-hole in nature [245], it is prudent to characterize the nozzle performance isolating one individual nozzle. This is done as it is easier to examine the spray characteristics experimentally for a single-hole nozzle, especially if they are cavitating [245].

3.1.1 Nozzles

The fuel injector chosen for this study is the Engine Combustion Network's Spray C#210037 (referred to as Spray C henceforth). The geometry [200, 2] and the rate of injection of Spray C have been measured indicating a lower mass flow rate compared to its counterpart Spray D [3] despite having a larger hole radius[161]. Through high-speed X-ray imaging experiments, Sforzo et al. [200] illustrated that there is strong asymmetric flow-separation and cavitation in Spray C#37 due to its sharp inlet corner. This led to a significant blockage in the flow while no such blockage or cavitation was evident in smooth nozzle Spray D. A subsequent X-ray imaging study coupled with computed tomography resulted in a 3D dataset for visualizing the internal flow in Spray C at maximum needle lift [225, 226]. As these data revealed intricate 3D flow morphology that resulted directly from the geometric features, Spray C made an excellent candidate for testing computational models for cavitating internal flows. The sharp inlet corner is shown in Figure 3.2, and it is shown that it induces severe asymmetry in the internal flow due to flow separation.

In addition, to studying the single-hole Spray C injector, a multi-hole Spray M1 [based on the design of smooth nozzle Spray D] injector has been studied. While the Spray C injector has shown asymmetric behavior due to the cylindrical nature of the nozzle and sharp inlet corners, the Spray M1 based on smooth inlet corner profiles and a converging nozzle design is not expected to cavitate [161]. The M1 nozzle was designed using the same nominal hole profile as the ECN Spray D nozzle [3], but five holes were installed at symmetric clock angles at the side of the injector sac as typical of a production injector. The angle of these holes corresponds to 146° based upon an in-production metal injector included angle. Figure 3.3 shows that the nozzle itself has a sharp inlet corner. Preference for hole length and optical access with back-lighting was given to one hole of interest. More details are discussed in the original work [245].

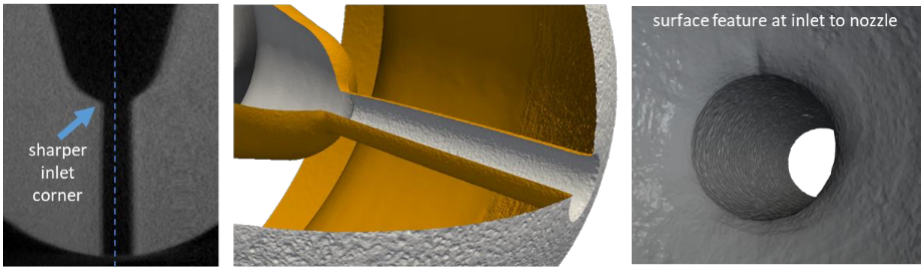

Figure 3.2. The X-ray CT scan shows the region of sharp inlet corner as well as features in the nozzle that have an impact on the spray growth and development. Figure adopted from [225].

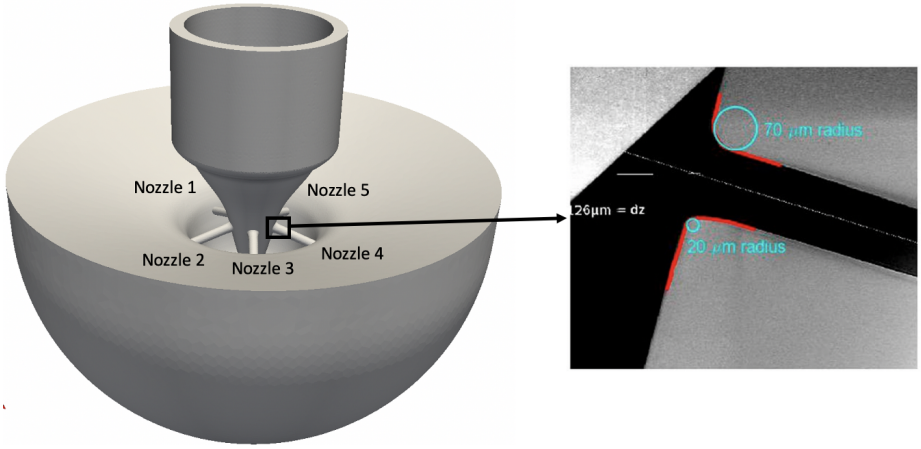

Figure 3.3. The schematic of the multi-hole Spray M1 injector showing the orientation of the nozzles. The inset image expands on one of the nozzles and shows the X-ray CT scan and the measured radii of the inlet corner. Inset figure adopted from [245].

Operating condition	**phase contrast imaging**	**radiography**
Ambient gas temperature (K)	298	298
Ambient gas pressure (MPa)	0.1	2
Ambient gas density ($kg - m^{-3}$)	1.4	22.8
Fuel type	n-dodecane	n-dodecane
Injection pressure (MPa)	150	150
Fuel temperature nozzle (K)	298	298

Table 3.1. Operating Conditions for the Spray C experiment, and used as boundary conditions for the CFD simulations

Operating condition	Optical Microscopy
Ambient gas temperature (K)	293
Ambient gas pressure (MPa)	5
Ambient gas density ($kg - m^{-3}$)	53.5
Fuel type	n-dodecane
Injection pressure (MPa)	50
Fuel temperature nozzle (K)	293

Table 3.2. Operating Conditions for the Spray M1 experiment, and used as boundary conditions for the CFD simulations

3.2 Experimental Method Description

In this section, the experimental methods, used to compare the model predictions, have been described in brief. These experimental studies have been conducted by collaborators and are only presented here briefly, to aid the discussions. Full details can be found in the manuscript discussing the original experimental studies [76, 103, 55, 200, 133, 131, 245, 130]

3.2.1 Geometry and Flow Morphology

. A suite of diagnostic techniques were implemented to characterize the internal geometry and spray behavior of the single-hole diesel injector designated as Spray C #210037 by the Engine Combustion Network (ECN) [1]. The target conditions used for each spray diagnostic technique are summarized in Table 3.1. Spray C is a common-rail fuel injectors which underwent minimal hydro-erosion to help maintain a sharp inlet corner and a cylindrical hole profile. To understand the effect of geometry on cavitation behavior, and by extension, the fuel spray distribution, high-resolution x-ray computed tomography (CT) scans of the injector geometries were obtained at the 7-BM beamline of the APS[103]. The nozzle geometry was reconstructed from a series of 2D projections using TomoPy [76]. The final spatial resolution of the geometry is 1.8 µm with a field of view (FOV) of 2.25 × 1.4 mm, capturing the full hole profile as well as part of the sac region. A more thorough discussion of

the experimental setup, image processing, spatial uncertainty, and geometric analysis may be found in a prior work [133]. Internal nozzle flow visualization for Spray C were achieved with high speed x-ray phase contrast imaging conducted at the 32-ID-B beamline at APS, using similar methodology described in detail in a previous work [55]. Spray C imaging was performed with ambient pressures of 0.1 MPa. A 2D mapping of the time-resolved projected fuel density field were carried out at the 7-BM facility at APS using the x-ray radiography technique [103]. Each injector was horizontally mounted in a chamber, and pressurized to 2 MPa with flowing N2, purging at a flow rate of 4 standard L/min to inhibit droplet accumulation in the field of view during data acquisition. The injectors were driven by a common-rail light-duty diesel injection system, and triggered at a rate of 3 Hz for a commanded injection duration of 725 μs. The beam was focused to a 4×6 μm spot, and the focal point was aligned with the injector tip. The pressure chamber was translated over a grid to build a raster plot of measurement points, capturing the spray evolution up to 25 mm from the nozzle tip. The x-ray intensity at each spatial location was averaged over 16 injection events. X-ray photons undergo photoelectric absorption as they pass through the fuel spray, with the degree of absorption directly related to the line-of-sight density via the Beer-Lambert law [8]. The recorded beam intensity is binned over a single synchrotron orbit, providing a temporal resolution of 3.68 μs. More details are described in the experimental work [133, 200].

In order to make direct comparisons to the multi-dimensional numerical modeling predictions, the flow morphology from the X-ray radiography has to be quantified. While the details of the experimental setup are described in the literature [226, 225], they are briefly discussed here. To further quantify the influence of the nozzle geometry on the flow morphology, the CT data were segmented using a simple Gaussian mixture model based on voxel intensities. The histogram of voxel intensity in the nozzle revealed a two-phase system of light and dark pixels which we have assumed

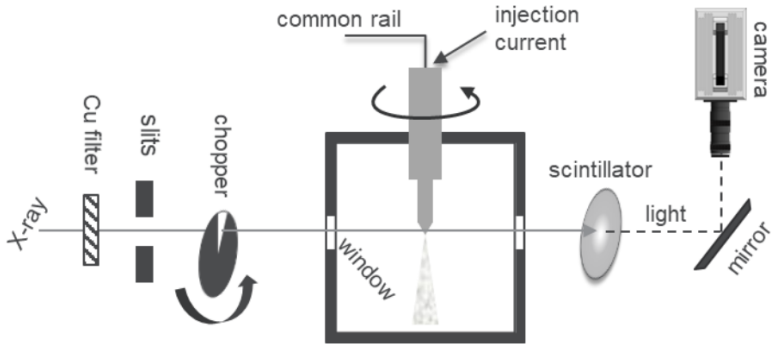

Figure 3.4. Schematic of experiment illustrating the high-speed X-ray imaging setup. Adopted from [225]

correspond to gas and liquid density, respectively. After segmentation, a volume of binary values was created indicating which pixels contain mostly liquid or mostly gas. The boundary between the liquid and gas layer was then used to define a "cavitation layer thickness". The technique achieved a pixel resolution of 2.1 microns per pixel, but the authors report an effective spatial resolution may be 5 microns or worse due to phase contrast effects as well as small motions of the injector body during the time-average snapshot of the injection [226, 225]. A segmentation algorithm was employed to identify and track the liquid-gas interface to enable extraction of quantitative information about the flow morphology. This flow morphology within the internal nozzle is used to validate the predictions from CFD.

3.2.2 Transparent Nozzle Setup

Experiments were conducted using an optically accessible chamber specifically designed for internal flow visualization as well as for carrying out observations in optically transparent nozzles, the schematic of which is shown in Figure 3.5. The vessel is equipped with four 25-mm diameter fused silica windows providing dual or stereoscopic line of sight optical access to the transparent nozzle. The nozzle is placed on a pedestal with open slots on four sides to allow direct visualization of the flow exiting the nozzle as well as the internal flow. The open slots allow for entrainment

of the air into the spray creating realistic in-cylinder conditions. The chamber is operated with nitrogen gas at constant pressure. The orifice is vertically aligned, and a modified Spray A solenoid-actuated injector is mounted atop the transparent nozzle. The n-dodecane fuel that was not degassed has been used in this experiment. The refractive index of n-dodecane is 1.42, while the acrylic nozzle material has a refractive index of 1.51. The refractive indices of the fuel and acrylic are similar enough that the internal nozzle appears optically transparent when filled with liquid; however, the refractive index difference is also large enough to make the internal sidewalls discernable due to the longer path length at the interface. When gas or vapor bubbles are present, the incident light will be scattered resulting in dark zones in the imaging. The fuel was pressurized by a high-pressure syringe pump at 50 MPa while ambient pressure is maintained at 5 MPa. The operating conditions are tabulated in Table 3.2. More details about the setup can be found in the relevant publication [130, 245].

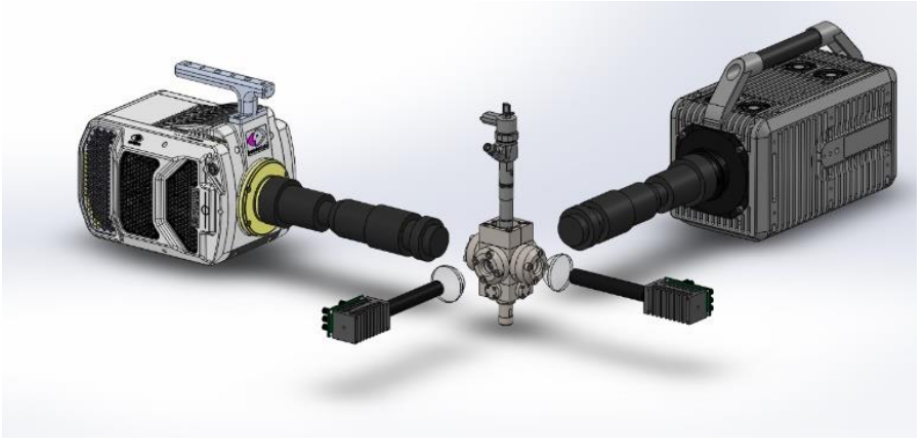

Figure 3.5. Schematic of the optical microscopy experimental setup, involving the use of high-speed cameras. Adopted from [141].

3.3 Numerical Model Description

3.3.1 Governing Equations

The flow is governed by conservation laws of mass and momentum given by:

$$\frac{\partial \rho}{\partial t} + \nabla . \phi = 0 \tag{3.1}$$

$$\frac{\partial \rho \vec{U}}{\partial t} + \nabla . (\phi \vec{U}) = -\nabla p + \nabla \overline{\overline{\tau}} \tag{3.2}$$

where ϕ is the mass flux, $\overline{\overline{\tau}}$ is the viscous stress tensor. The mass and momentum equations were Favre averaged to account for turbulence and variable density without the special notation for the sake of simplicity.

To account for the compressibility, Schmidt et al [195], formulated a compressible pressure equation based on the continuity and the discretized momentum equation given by

$$\frac{1}{\rho} \left.\frac{\partial \rho}{\partial p}\right|_{x,h} (\frac{\partial \rho p}{\partial t} + \nabla \rho p \phi) + \rho (\frac{H(u)}{a_i}) - \rho \nabla . \frac{1}{a_i} \nabla p + \left.\frac{\partial \rho}{\partial p}\right|_{p,h} \frac{Dx}{Dt} = 0 \tag{3.3}$$

where a_i is the coefficient of the contribution from momentum in the cell of interest, $H(u)$ is the sum of contributions from the neighboring cells, and x is the instantaneous quality of the mixture. The simulations in the current study were performed under isenthalpic conditions, which implies that the total enthalpy remains constant thus rendering the energy equation inconsequential. The fluid properties, such as the temperature, quality, densities of liquid and vapor, are interpolated from a lookup table. This lookup table containing the various fluid properties is generated as a function of pressure and specific enthalpy for a specified range using the REFPROP database [116].

3.3.2 Modeling Phase Change using Homogeneous Relaxation Model

Spray C is known to be a cavitating injector [161, 200, 225, 226], involving the phenomena of phase change within the nozzle. This therefore requires the ability to model vaporization within our numerical framework. A vaporization model was added to the $\Sigma - Y$ solver [62] introduced in the previous chapter, and results validated against experimental data. The implementation has been described in full detail in the literature [62, 16] and briefly introduced in this section. A transport equation for the fuel vapor mass fraction is added to the system and in order to conserve the mass, a vaporization sink term is added to the liquid mass transport equation introduced in the previous chapter.

$$\frac{\partial \bar{\rho}\tilde{Y_v}}{\partial t} + \frac{\partial \bar{\rho}\tilde{u_i}\tilde{Y_v}}{\partial x_i} = -\frac{\partial \widetilde{\bar{\rho}u_i{}'Y_v'}}{\partial x_i} \tag{3.4}$$

$$\frac{\partial \bar{\Sigma}}{\partial t} + \frac{\partial \bar{\Sigma}\tilde{u_i}}{\partial x_i} = \frac{\partial [D_s(\frac{\partial \bar{\Sigma}}{\partial x_i})]}{\partial x_i} + (A + a)\bar{\Sigma} - V_s\bar{\Sigma}^2 \tag{3.5}$$

where $\tilde{Y_v}$ is the local vapor fuel mass fraction.

There are a few ways in which fuel vapor can be generated in DI systems. The first is through cavitation and the second through flash-boiling. These occur when the local pressure drops below the vapor pressure of the fluid. Here the latent heat of vaporization is provided for by the sensible heat, which is already available in the fuel. The third way that it can occur is through evaporation, when the downstream ambient conditions are hot and the latent heat of evaporation is provided by the ambient gas. Phase change models for both of these mechanisms have been developed and are described below.

The CFD code HRMFoam, based on the foam extend library of OpenFOAM employs the Homogenous Relaxation Model (HRM) to capture the phase change due

to cavitation. This model calculates the rate at which the vapor mass fraction will move to the equilibrium value [24,25]. The HRM model is governed by the set of the following equations:

$$\frac{Dx}{Dt} = \frac{\bar{x} - x}{\theta} \tag{3.6}$$

$$\theta = \theta_0 \alpha^{-0.54} \Theta^{1.76} \tag{3.7}$$

$$\Theta = \frac{P_{sat} - P}{P_c - P_{sat}} \tag{3.8}$$

The void fraction, α, is determined by the density and saturated liquid and vapor densities,

$$\alpha = \frac{\rho_l - \rho}{\rho_l - \rho_v} \tag{3.9}$$

where x is the instantaneous vapor mass fraction, $\bar{x}$ is the equilibrium mass fraction, and θ is the relaxation time scale. The local void fraction is defined by Θ, P_{sat} is the saturation pressure, P_c is the critical pressure, θ_0 is the time scale coefficient usually taken as 3.84e-7 s.

While HRM was originally developed for one-dimensional analysis, it was extended to a multi-dimensional CFD model by Schmidt et al [192, 194, 195]. The introduction of variable rate phase change required a method for connecting predicted phase change with conservation of mass and momentum. This can be done using the chain rule shown below. This allows for the pressure to respond to compressibility, density change from phase change, and density change from turbulent mixing with the non-condensable gas.

$$\frac{D\rho}{Dt} = \left.\frac{\partial\rho}{\partial p}\right|_{x,h,y} \frac{D\rho}{Dt} + \left.\frac{\partial\rho}{\partial x}\right|_{p,h,y} \frac{Dx}{Dt} + \left.\frac{\partial\rho}{\partial h}\right|_{p,x,y} \frac{Dh}{Dt} + \left.\frac{\partial\rho}{\partial y}\right|_{p,x,h} \frac{Dy}{Dt} \tag{3.10}$$

The HRM model in diffuse interface Eulerian solvers has been used and validated in a wide variety of applications [146, 141, 92, 175].

3.3.3 Numerical Schemes

The choice of turbulence models used in internal nozzle flow simulations has shown to have a noticeable effect on the final solution [141, 92]. For the current simulations with the HRMFoam phase change model, the $k-\omega$ Shear Stress Transport (SST) turbulence model was employed [41, 137]. The spatial discretization was 2nd order Gauss linear and temporal discretization was 1st order Euler. To access thermodynamic properties HRMFoam uses a look up table from REFPROP [116]. The CFD simulations were a two-phase flow problem, the fuel in a liquid as well as vapor (gas) phase, and the ambient nitrogen in the gas phase. All species were considered compressible and miscible. The initial sac was partially filled with gas in order to simulate the real nozzle initial condition [141, 131], which consists of some residual gas. The boundary conditions include Pressure Inlet for the injector inlet and Pressure Outlet for the cylinder outlet. The walls have a *No-Slip* boundary conditions in order to produce boundary layer effects [190]. The injection pressure was 150 MPa and ambient gas pressure was 0.1 MPa for the Spray C simulations. For the Spray M1 simulations, the injection pressure was set to 100 MPa and ambient gas pressure was 2 MPa . The fuel inlet and ambient gas temperatures were both 25 degree Celsius (298 K), in accordance to the experimental conditions [refer Table 3.1, 3.2]. Note that these conditions differ from those used to experimentally measure mass flow by Payri et al. [161], so no direct validation of mass flow predictions are possible. The meshes used in the simulation are shown in Figure 3.6, and they include a mix of polyhedral and hexahedral (near walls) cells.

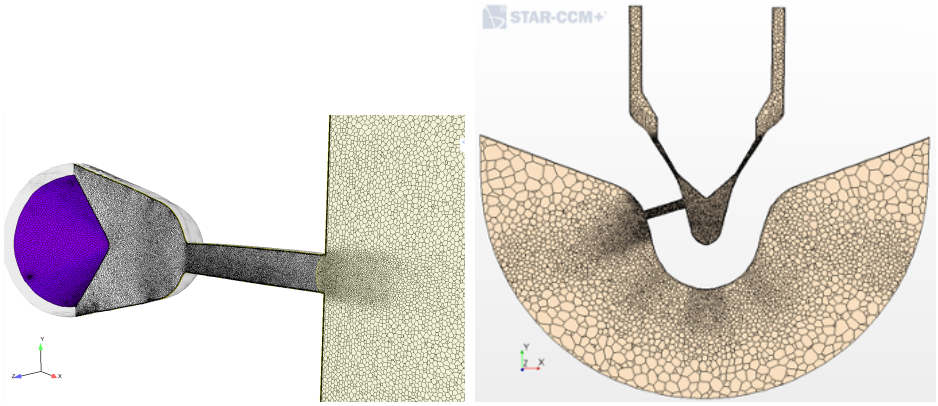

Figure 3.6. The internal and near nozzle regions are of interest in this study, therefore more mesh points have been used in these regions. To resolve boundary effects, adequate resolution of mesh points have been provided close to the walls. The nozzle regions are well resolved with both meshes resolving features up to $2\mu m$, inline with the resolution of the X-ray measurements. More details can be found in Table 3.3. [L-R] a clip of the Spray C mesh layout on the left, and a clip of the Spray M1 layout on the right. Due to the five-hole alignment, using a clip plane only one hole can be visualized at a given time.

Nozzle	Number of Mesh points	Minimum Grid size
Spray C	4.2 million	$2\mu m$
Spray M1	3.5 million	$2\mu m$

Table 3.3. Mesh Description for CFD simulation

3.4 Results and Discussions

3.4.1 Spray C

The results from the experimental studies [225, 226, 224] are compared to the CFD predictions in Figure 3.7 and 3.8. Figure 3.7 shows a direct comparison between experimental results from the methodology discussed earlier and liquid volume fraction contours from CFD data. Several slices perpendicular to the nozzle flow axis are shown for illustration. Both the imaging experiments and the simulations predict flow-separation and cavitation extending to the nozzle exit. The CFD simulations do predict the flow separation and resulting morphology of the liquid-gas interface fairly well. However, the imaging results show a thinner liquid-gas interface as opposed

to a diffuse interface predicted by the simulations. This can be attributed to the choice of the phase change model in HRM, which is a diffuse interface single-fluid Eulerian phase model. Certainly, the phase contrast effect and its retrieval using low-pass filtering may reduce the spatial resolution of the imaging results but it is expected that the interface will appear more diffuse than in reality due to these effects [225]. Finally, the images show the liquid-gas interface is wrinkled as it propagates downstream. This effect is not predicted by the CFD results.

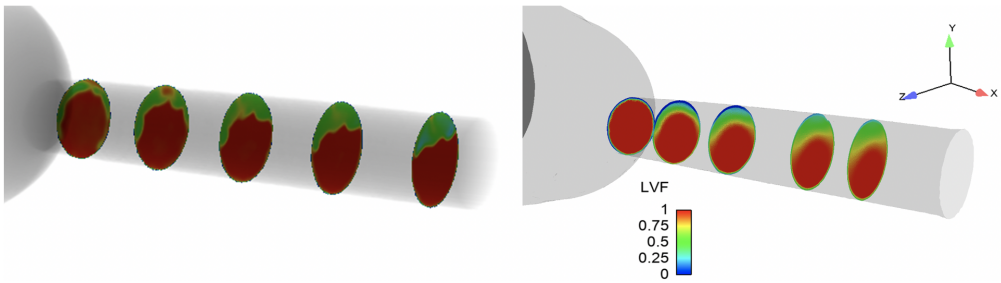

Figure 3.7. Comparison between intensity map from the X-ray imaging experiments (on the left) with the liquid volume fraction contours from CFD simulations (on the right). Both the X-ray and CFD simulations are scaled on the same range,

Figure 3.8 shows an orthogonal cut plane along XY plane that reveals the strong flow separation from the inlet corner along +Y. The results also indicate that there is significant blockage of the flow leading to lowered mass flow rate. The CFD simulation liquid volume fraction contours indicate that while the majority of the flow cross-section is in liquid phase, the transition to gas phase occurs over a diffuse liquid-gas interface. Only a thin film near the wall contains pure vapor, which then disappears downstream. Finally, due to the divergent geometry of the nozzle hole, there is thin flow-separation layer along the axis of the nozzle. This result is validated against the experimental observation using X-ray radiography (from [200]) in Figure 3.8 (left image panel). The propensity to cavitate in both the experiments and CFD are similar in nature, arising from the sharp inlet corner. The regions in the X-ray image corresponding to lighter shades indicate gas phase, and the darker shades is the liquid

fuel. It appears there is evidence of asymmetric cavitation in the nozzle, arising as a result of the sharp inlet corner. The lower end of the nozzle shows evidence of the presence of the gas phase due to cavitation. When compared to the CFD, the trends are strikingly similar, with a tendency to cavitate more from one side and a thin layer of gas along the lower edge of the nozzle.

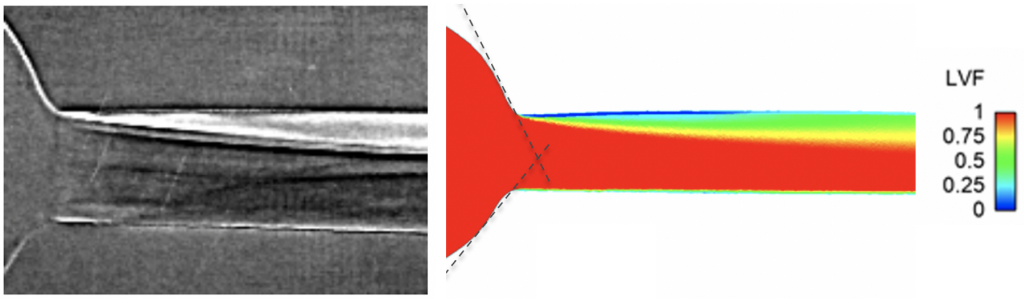

Figure 3.8. Comparison between internal flow radiography from the X-ray imaging experiments (on the left) with the liquid volume fraction contours from CFD simulations (on the right) show similar trends in the cavitation layer growth and the asymmetric nature of cavitation. Both the model and experiments predict a thin layer of cavitation on the smooth inlet corner (lower side). Experimental results from [200].

Through imaging data alone, it is difficult to distinguish between flow separation and cavitation as a pressure field or fuel vapor fraction could not be measured through the image intensity. But a further investigation of CFD results can reveal the extent of cavitation. Figure 3.9 shows a localized region just downstream of the sharp inlet corner that experiences pressures below the ambient gas pressure. This is the source of fuel vapor formation and explains the prediction from the model that aside from flow-separation there is significant cavitation that originates at the sharp inlet corner. Some studies [127, 75] have reported evidence of gas ingestion at the nozzle exit region for Spray C. That observation is verified in Figure 3.9 by showing the ambient gas volume fraction contour close to the exit. It appears there is a propensity for the ambient gas to enter the nozzle creating conditions conducive for a hydraulic flip -

a phenomenon particularly damaging to the performance of a fuel injector and the subsequent atomization process [212].

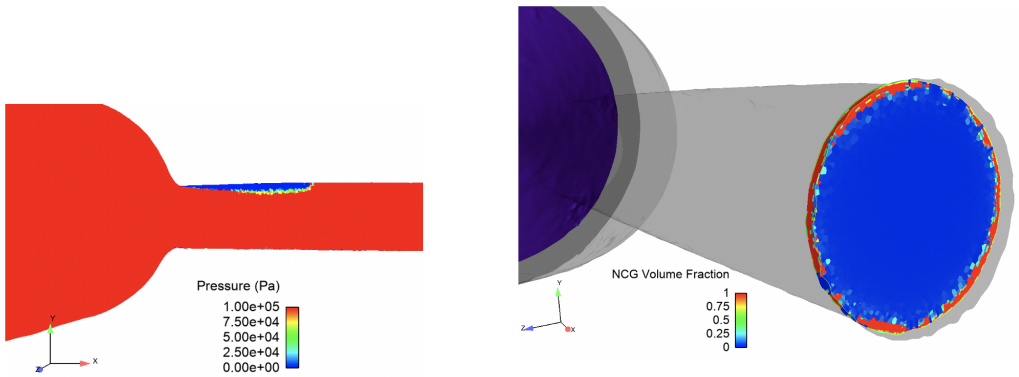

Figure 3.9. [L-R]: The pressure map inside the nozzle indicates a low pressure region corresponding to the sharp corner radius which causes cavitation within the nozzle. On the right, an ambient gas phase volume fraction contour, referred here as Non-Condensible Gas (NCG), shows evidence of hydraulic flip occurring at the nozzle exit, confirming predictions from other studies [127, 75].

While the validation of the model predictions to the internal flow experimental measurements are useful, the effect of the internal flow can be seen prominently in the near-field development of the spray. The X-ray radiography fuel density measurements at 0.1 mm, from the nozzle exit, are compared to predictions from the CFD, shown in Figure 3.10. Due to the asymmetric flow in the nozzle, the near-field radiography of the spray shows a region of fuel deficit corresponding to the cavitation zone. This behavior is confirmed from the CFD simulations, where the spray loses the circular shape and has a lower density region corresponding to the sharp inlet corner. Both the experiments and CFD results show regions of high densities opposite to the cavitation region. This shows the importance of accurately resolving the internal nozzle flow, even under quasi-steady conditions, and the subsequent effect of that in characterizing the spray behavior in the near-nozzle and far-field region.

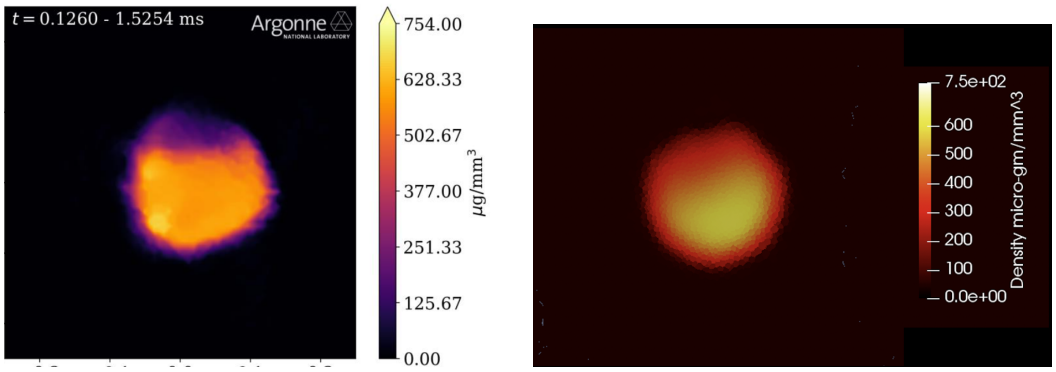

Figure 3.10. The near-nozzle spray at 0.1 mm from the nozzle exit shows the similarity in predictions between the CFD and the experimental observations. The region of fuel mass deficit corresponds to the cavitation layer within the nozzle. This further shows the importance of internal flow modeling in order to impose appropriate boundary conditions for external spray models. Experimental data from [200]

3.4.2 Spray M1

Since X-ray measurements of multi-hole injector geometries are challenging, due to interactions with the emerging neighborhood plumes, the experimental studies have been conducted using optical microscopy using high-speed cameras. More details about the setup and the results are provided in the corresponding publication [245, 131].

Figure 3.11 shows the time sequence images of the internal flow for the flow development in assessing a hole cavitation. The experimental setup is focussed on the sharp inlet corner for one of the holes [refer Figure 3.3]. Although difficult to portray in still images, the propensity for cavitation at the inlet of the hole during the higher needle-lift, steady period may also be determined from the image sequence shown in Figure 3.11. It shows that the inlet to the hole of interest becomes brighter at 105 μs and strengthens in intensity by 158 μs. The random flickering of light/dark structures at the hole inlet, indicate cavitation. The flickering is especially persistent at the upper side of the hole of interest, extending into the hole itself. The views at the left of the nozzle and other operating conditions indicate some cavitation layers

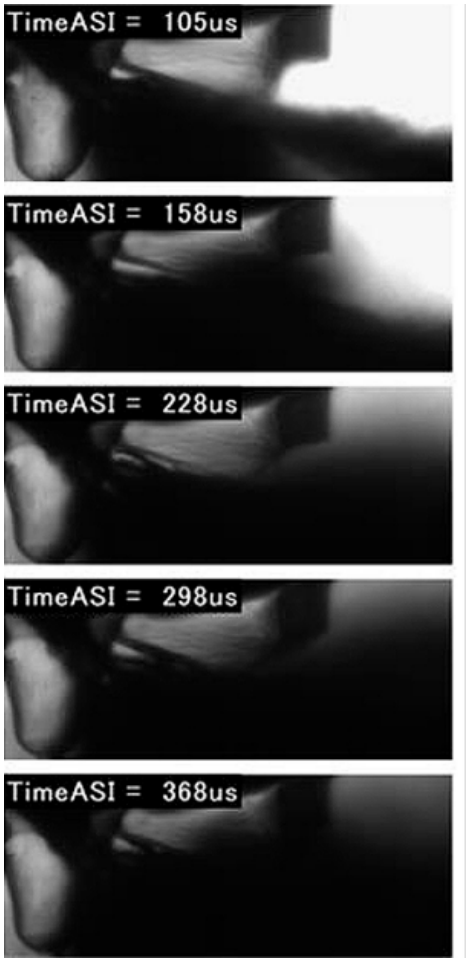

Figure 3.11. Time sequence images after start of injection (ASI) for assessing a hole cavitation. Fuel injection pressure 50 MPa, discharge pressure 2 MPa. Adopted from [245].

at the bottom of hole as well [245]. However, the major conclusion for the M1 nozzle during the steady period is that frequent cavitation is found at the hole inlet despite significant rounding by hydro-erosion. This is an important result as the cavitation occurs at the upper corner of the converging nozzle injector.

The X-ray CT characterized geometry for M1 nozzles are used and CFD simulations computed for similar operating conditions. The quasi-steady period, time-averaged CFD results indicate evidence of in-nozzle cavitation as well as near the upper corner of the inlet hole as seen from the experiments [245]. However the ex-

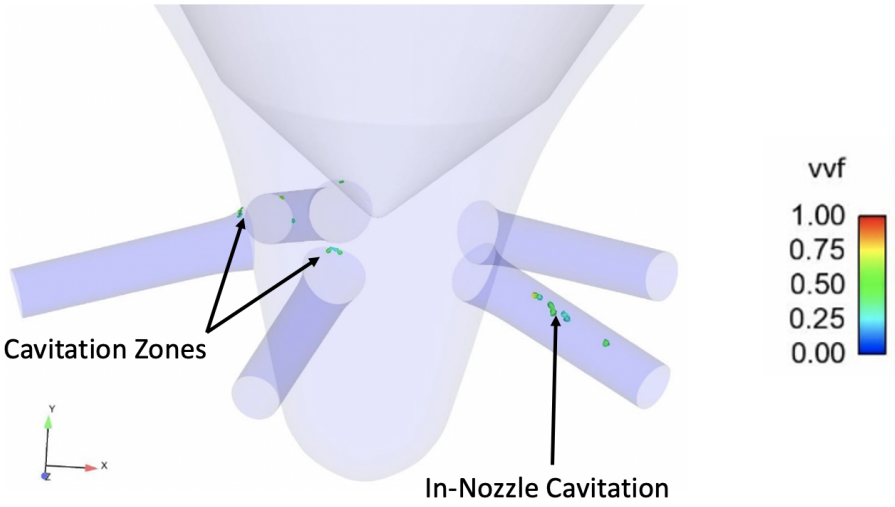

Figure 3.12. The iso-surface of fuel vapor fraction indicates in-nozzle cavitation and cavitation zones corresponding to the sharp inlet corners.

tent of cavitation zone region are found to be smaller in the CFD simulations, shown in Figure 3.12. This can be attributed to the mesh resolution, choice of turbulence models, and use of the Reynolds Averaged Navier Stokes (RANS) modeling framework among others. Nevertheless, this is a better result in comparison to the CFD simulations from [245] that did not predict any cavitation zone within the nozzle. In order to investigate the cause of cavitation further, we plot the pressure gradients within the nozzle. These pressure gradient maps are an indication of the flow characteristics. Figure 3.13 colors these pressure gradients into favorable pressure gradient regions (shown in blue) and adverse pressure gradient regions (shown in red). Comparing with Figure 3.12, there seems to be a correlation between the adverse pressure gradient regions and the propensity for the nozzle to cavitate.

Numerical studies of multi-hole injectors have shown to involve highly complicated, transient fluctuations in the mass flux of the injector [17, 218]. This is highly undesirable, as it could lead to inconsistent fuel delivery into the combustion chamber. This complex internal flow results in intermittent string cavitation phenomena when a strong vortex is injected and the resulting swirling spray contains a thermal

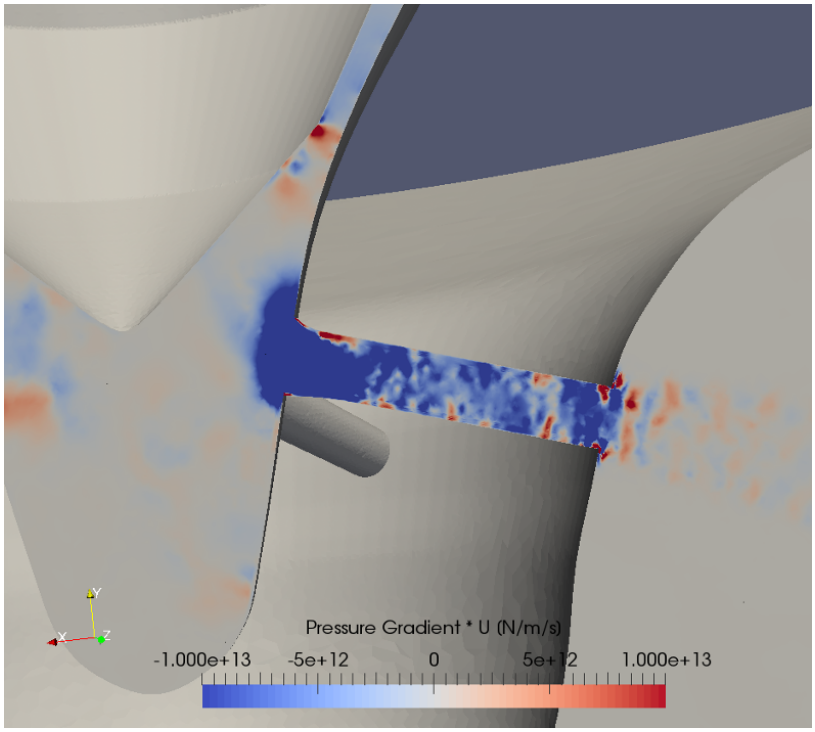

Figure 3.13. The plot of $U * \nabla.P$ shows region of adverse pressure gradient (in red) correpsonding to the cavitation zones from Figure 3.12.

non-equilibrium vapor core, which causes some nozzles to have a mass flow deficit [17]. In order to identify potential sources of vortex in the Spray M1 internal flow, the $\lambda 2$ based criterion is used [96]. The $\lambda 2$-criterion looks for a pressure minimum but removes the effects from unsteady straining and viscosity by discarding these terms. Taking the gradient of the Navier-Stokes equations results in

$$a_{ij} = -\frac{1}{\rho} p_{ij} + \nu u_{i,jkk} \tag{3.11}$$

where a_{ij} is the acceleration gradient and p_{ij} is symmetric. Decomposing the acceleration gradient into symmetric and anti-symmetric parts, we get the vorticity transport equation as the anti-symmetric part, and the symmetric part

$$\frac{DS_{ij}}{Dt} - \nu S_{ij,kk} + \Omega_{ik}\Omega_{kj} + S_{ik}S_{kj} = -\frac{1}{\rho} p_{ij} \tag{3.12}$$

The first two terms in the left hand side represent unsteady irrotational straining and viscous effects respectively. Therefore only $S^2 + \Omega^2$ is considered to determine if there is a local pressure minimum that entails a vortex. A vortex is defiened as 'a connected region with two negative eigen values of $S^2 + \Omega^2$' [96]. Since $S^2 + \Omega^2$ is symmetric, it has real eignevalues only, and by order the eigenvalues $\lambda 1 \leq \lambda 2 \leq \lambda 3$, the definition becomes equivalent to requiring that $\lambda 2 < 0$. Generally, these vortex cores are visualized as the isosurface for $\lambda 2$ that are slightly negative. Figure 3.14 plots the time-averaged negative $\lambda 2$ iso-surfaces. The vortex cores appear near the needle seat region, a constricted area of the sac as well as in significant portions of the nozzle inlet. This is because of the orientation of the nozzle holes with respect to the original flow direction, which requires the flow to turn thereby causing regions of re-circulation and low pressures resulting in a vortex formation. Figure 3.14 identifies a vortex core originating between two nozzle holes, marked by the letters A and B. This shows even for converging nozzles, and operating conditions (high ambient pressure, low injection pressure), the injector geometry has a leading order effect on the internal flow and the subsequent spray growth.

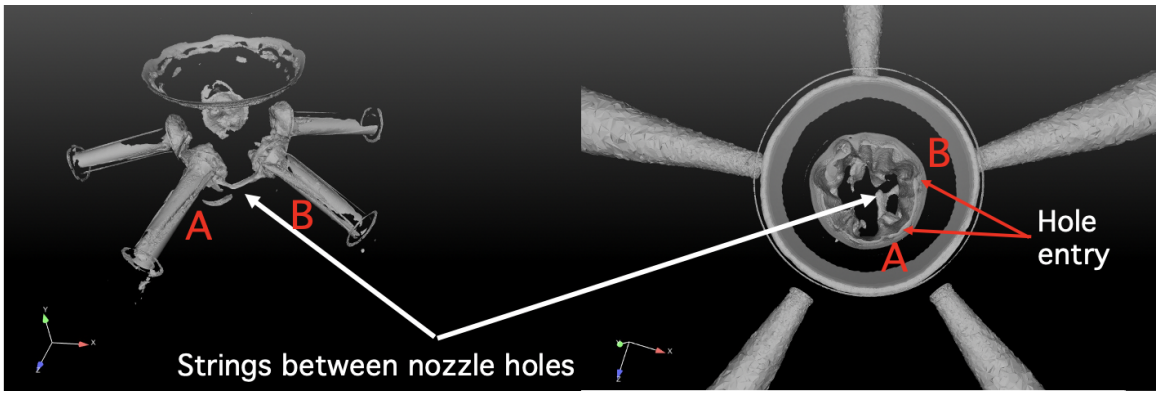

Figure 3.14. Time averaged isosurfaces for $\lambda 2$ indicates vortex cores between neighboring holes. Such vortex cores can cause mass fluctuations in the nozzles, deteriorating injector performance.

3.5 Inference

Comparisons between CFD simulations and experimental studies yielded important inferences:

- The Homogeneous Relaxation Model (HRM) is able to robustly predict the onset of cavitation within single-hole and multi-hole nozzles. While the extent of cavitation may differ between numerical and experimental observations, it can be attributed to the choice of numerical schemes, turbulence models as well as the modeling paradigm in RANS.

- The use of high-fidelity geometry inputs is a crucial step in improving model fidelity. This can be seen from the similarity in predictions to the experimental observations, in the internal flow as well as the near-field region.

- The onset of cavitation is a result of the pressure drop and adverse pressure gradients in the regions near sharp-inlet corners.

- Hole-to-hole vortex cores are observed for converging smooth nozzles, and is an important result in designing nozzles that suppress the propensity for this adverse phenomena.

- Finally, resolving the internal flow dynamics accurately remains a critical step in order to better inform boundary conditions, and tackle flow transients for the external spray modeling paradigm.

CHAPTER 4

MULTI-DIMENSIONAL MODELING OF A FULL INJECTION EVENT

4.1 Motivation

In the previous chapter we demonstrated the capabilities of the multi-dimensional Computational Fluid Dynamics (CFD) models in predicting phase change for converging and cylindrical nozzles. These investigations were limited to static mesh simulations of the quasi-steady phase of injection, arguably the longest and most important. However, in a realistic injection event the needle valve is never stationary and has been shown to have an effect on spray characteristics even during the steady phase [141, 245]. Another feature of these nozzles are multiple injection events, where the residual effect of previous injections effects the latter events. This phenomenon has been discussed widely in experimental literature [158, 57, 207], and it has shown to affect the combustion characteristics in the case of diesel engines [207]. For a truly predictive modeling tool, it is therefore important to be able to realistically model this transient phase of injection, including the dwell phase between injections, in which the fuel supply is shut off due to the closure of the needle valve.

The early and late transients during a fuel injection event have been shown to disproportionately contribute to the amount of emissions in a direct injection internal combustion engine [207]. The needle valve opening and closing event introduces many computational challenges. The changing topology due to the valve displacement necessitates adaptive meshing strategies, especially small mesh resolution when the gap between the valve and the seat is small resulting in expensive computations.

To model complete closure during the dwell phase, with a zero mesh thickness, adds additional complexities that previous modeling efforts have addressed using disconnected domains. Arienti et al., among others [14, 157], have used this idea to simulate a complete single injection cycle, starting and ending at a very low needle lift. Battistoni et al. [22, 21] modeled the the Start of Injection (SOI) and End of Injection (EOI) separately with the disconnected domain idea. To study SOI, Battistoni et al. [22, 21] initialized the sac with either liquid fuel or non condensible gas. For the case with non-condensible gas filled sac, an injection rate ramp up close to the available experimental data was observed. Other computational studies [155, 219], have shown evidence of cavitation during the needle opening and closing phase as well as temperature gradients observed during early transients due to mixing of hot fuel and cold ambient gas.

In order to alleviate this challenge in abrupt change of domain topology, a gradual needle closure numerical implementation has been used. The sealing force is applied in the needle seat region. While the full details of the implementation are discussed in Mohapatra et al. [145], the relevant numerics are briefly introduced in the following sections. In the original work, the needle sealing algorithm was implemented for multi-hole gasoline injectors, while in this work we study single-hole diesel nozzles. The big difference apart from injector configuration is the needle lift for these simulations. A typical diesel injector has much higher needle lifts (almost 3-5x) the lifts found in the gasoline injectors due to the fuel thermodynamic considerations. This introduces unique computational challenges in modeling transient behavior and therefore makes the single-hole diesel injector study, challenging. This gradual and easily-implemented model of sealing avoids spurious water-hammer effects observed in other numerical studies [155]. To help build intuition, synthetic needle lift data is plotted alongside the sealing algorithm function in Figure 4.1. As the needle lift stays low, say at a threshold around 10 μm, the sealing is activated (here shown by the value set to 1).

Whereas, in regions of needle lift greater than the threshold, the sealing algorithm is deactivated. This strategy helps in modeling the dwell phase between injections when the needle lift is near-zero, and therefore the sealing is activated to prevent flow through the seat region. The rate of sealing is determined by a time constant, τ introduce in the subsequent section.

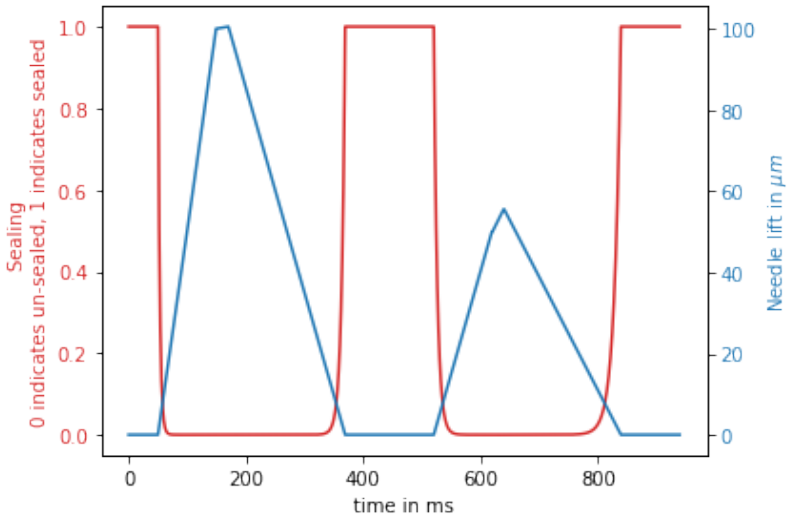

Figure 4.1. Synthetic multiple injection event and corresponding sealing activation

4.2 Nozzle

For this investigation, the Engine Combustion Network Spray D [3] injector has been studied (shown in Figure 4.2). This injector, as introduced in the previous chapter in the context of the M1 injector, is a converging nozzle single-hole injector. It has a nominal inlet diameter of 226 μm, an outlet diameter of 190 μm, with a hole length of 965 μm. The makes the orifice conical with a k-factor of about 1.5. The use of convergent nozzles is justified precisely because the convergence of the holes will discourage the onset of the cavitation phenomenon [161]. Numerous experimental and modeling results have shown the importance of injector geometry features such as inlet corner radius and hole conicity in dictating the onset of cavitation [193, 163, 163,

165, 131]. It is believed that a converging nozzle tends to suppress the tendency of a nozzle to cavitate, and this hypothesis has found wide-spread acceptance in the engine community. Previous research shows that geometry as well as initial sac conditions have a very important effect on the flow characteristics [141, 131]. To truly validate the numerical modeling results with experimental observations, it is important to provide a high- fidelity input in terms of nozzle geometry as well as initial conditions in sac and nozzle. We follow our approach first introduced in Chapter 3 (refer Figure 3.1) to establish the validations. In this chapter, while we mainly discuss results from the quasi-steady phase of injection [141], we shall introduce some dynamics related to the transient phase, in particular the early and the late phase of injection.

Figure 4.2. Output from x-ray CT scans showing the smooth inlet profile of the converging nozzle

4.2.1 Needle Lift

A multiple injection strategy was used in the experiments and the simulations, with 1 ms and 0.3 ms main and secondary injection duration, respectively, and a 0.3 ms dwell time between them. While the full needle lift profile is shown in Figure 4.3, we shall in the subsequent section briefly introduce the experimental methods used

to obtain these data. It is important to note that the first peak is twice as higher as the latter peak as the former is the main injection event.

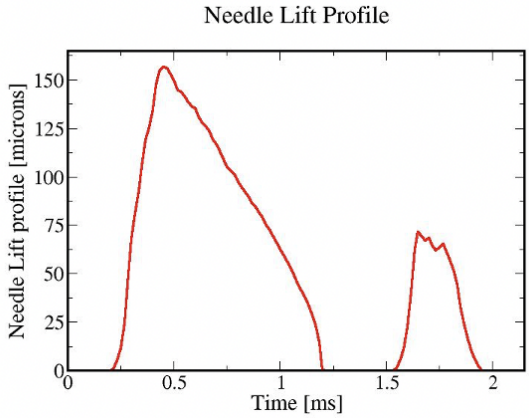

Figure 4.3. The multiple injection event characterized by the needle lift measured by x-ray diagnostics.

4.3 Experiments

In this section, the experimental methods, used to compare the model predictions, have been described in brief. These experimental studies have been conducted by research collaborators during the course of this study and are only presented here to aid the discussions. Full details can be found in the manuscript discussing the original experimental study [76, 103, 55, 200, 133, 131, 245, 130].

The experimental methods to characterize the nozzle geometry using x-ray Computed Tomography (CT) were introduced in Chapter 3 [refer to Section 3.2.1 for more details]. The experimental setup remains the same for characterizing the Spray D injector. Once the x-ray data were obtained, the resulting image stack from the X-ray scans was tomographically reconstructed to transform the nozzle projections into planar slices through the nozzle body. The software that was used to perform the reconstruction is TomoPy, an open-source Python code developed for synchrotron tomography [76]. The reconstruction algorithm is a direct Fourier-based method called

Gridrec. This method makes use of the Fourier slice theorem to build a 2D Fourier space from the transforms of parallel projections of the nozzle [103]. The spatial uncertainty associated with the full imaging and post-processing procedure was found to be 1.8 μm, determined from calibration of three steel pin gauges ranging in diameters between approximately 100 and 250 μm. The tomographic reconstruction method has been described in greater detail under a previous publication [133]. Figure 4.2 shows the output from the metal target nozzle characterization.

4.3.1 Needle Motion

A key step in characterizing the injectors was non-intrusively measuring the needle lift. The measurement serves as an essential input into the computational modeling as well as a fundamental parameter in the control of injector transients. The injector was horizontally mounted in a 0.5 L pressure chamber fitted with a pair of 12 x 30 mm x-ray transparent polyimide windows. Room-temperature fuel was pressurized to 100 MPa using a diesel common rail system, with the injector firing at 1 Hz into a nitrogen environment at 2 MPa absolute pressure. Unfiltered undulator white beam passed through a silicon monochromator crystal to absorb low-energy photons that would otherwise contribute to heat loading the nozzle. The filtered beam then passed through a mechanical chopper before intersecting the injector tip. As the beam impinged on the injector, x-ray photons were absorbed by the metal body, creating spatial contrast in the beam cross-section. Downstream of the spray chamber, the transmitted x-ray photons were directed onto a LYSO:Ce scintillator crystal, which converted the X-rays to visible light. The resulting image was magnified with a 5x objective lens and the 16-bit image was recorded using a high-speed camera operating at 80 kHz. In order to measure the full three-dimensional motion of the needle, images were recorded at two different viewing angles. To obtain the mean needle motion, 30

individual spray events were recorded and ensemble-averaged. The resulting needle lift is shown in Figure 4.3

4.3.2 Transparent Nozzle

To visualize the internal flow at high speeds, a transparent facsimile of the Spray D nozzle was constructed. This transparent nozzle provided the ability to assess to what degree the flow would cavitate and to also allow observation of spray growth as liquid mixes with charge-gas (nitrogen) in a pressure chamber. The challenges of the construction were to provide good optical access, sufficient mechanical strength for surviving high-pressure fuel injection, and a faithful adherence to the metal spray D geometry. The imaging setup has been previously introduced in Chapter 3. The full description of the experimental setup is described in detail in the literature[131, 130].

The imaging region for each camera included both the internal nozzle flow as well as the external spray, allowing the observations of how the injector flow influences spray development. However, because of the thickness of the acrylic nozzle and the small depth of field of the microscopy, the best focal plane to visualize internal flow is different than that of the external spray. As a compromise, one camera was adjusted for best focus on the internal flow, while the other camera was adjusted for best focus on the external spray. Features from the out-of-focus region remain observable and are clarified with the in-focus images from the orthogonal view. A metal fuel injector of the same variety as Spray D was ground flat just below the needle-seat sealing surface and replaced by an acrylic nozzle made as close as possible to the shape of the metal nozzle. The transparent tip was mated to a metal injector body and sealed in place using a combination of nozzle support and clamp. Detailed discussions for the transparent nozzle design and geometry can be found in a previous publication [141].

One concern over using the acrylic nozzles as a proxy for the metal nozzles is the effect of surface roughness. Surface roughness could trigger cavitation regardless of the gross geometry of the nozzle. Using a scanning electron microscope, the surface of the acrylic nozzle was examined. Figure 4.4 shows the features of approximately 2 μm in extent, which is similar to the surface roughness of the metal targets particularly when considering aging and deposits of an injector in service.

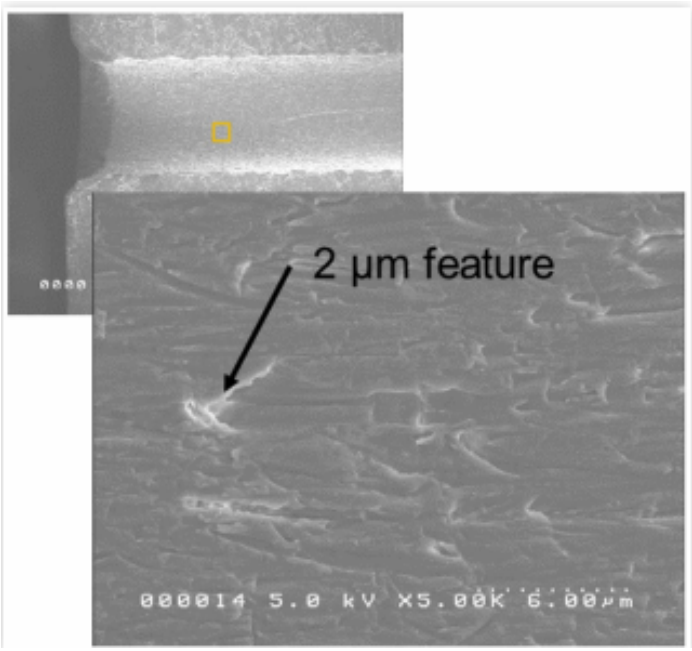

Figure 4.4. Scanning Electron Microscopy reveals the small scale features observed in the transparent nozzle. This is consistent with the x-ray CT scan resolution. Figure adopted from [141]. Credit to Pickett et al.

4.4 Modeling

Engineering-level CFD modeling was used because of its modest cost and the ability to perform parameter studies over multiple conditions. An Eulerian approach based on the Homogenous Relaxation Model (HRM) [195] introduced previously has been used for this study. The in-house solver is implemented using the foam-extend branch of the OpenFOAM CFD library [94]. While the governing equations intro-

duced in the previous chapter remains consistent, to model the numerical sealing phenomena some modifications have been introduced.

The discontinuous domain approach [14] briefly introduced in the previous section has a few shortcomings. Primary among them is the abruptness of topologically severing the domain would create massive disturbances in the flow, triggering powerful waves that could result in spurious water-hammer effects as observed by [155]. Instead, Mohapatra et al.[145] proposed to model what cannot be directly resolved in the simulation. When the needle is nearly closed against the seat, the drag forces produced by the walls begins to dominate the flow. The closure model employs a gradual increase in drag that halts the flow through the narrow gap. To avoid degenerate cells, their sealing model activates at low, but finite, needle valve lift. The threshold for this simulation is set to 10 μm in this simulation, whereas in the original work it was 5 μm. This can be attributed to the differences in the geometry, needle lift profile and the mesh motion strategy used to address the changing topology.

4.4.1 Numerical modeling of needle valve sealing

The governing equations for mass and momentum conservation are described in Eqn. (4.1) and (4.2) respectively. Both equations are written using barycentric velocity. All transported variables solved by the following formulations are Favre averaged, as the formulation accounts for turbulence and variable density.

$$\frac{\partial \rho}{\partial t} + \nabla . \phi = 0 \tag{4.1}$$

$$\frac{\partial \rho U}{\partial t} + \nabla . (\phi U) = -\nabla p + \nabla \bar{\bar{\tau}} + \bar{f} \tag{4.2}$$

In the above equations, ϕ represents the face valued mass flux. $\bar{\bar{\tau}}$ represents the stress tensor, which includes both viscous and turbulent stresses. The volumetric

source term, $\bar{f}$ in Equation 4.2 corresponds to an artificial drag force. This is a component of the needle-seat algorithm.

Mathematically, the sealing model can be expressed as:

$$\bar{f} = \rho U \frac{S_f}{S_d} \tag{4.3}$$

where the S_f and S_d terms are the sealing factor and drag constant respectively. The drag constant represents how vigorously the imposed force opposes the fluid motion. A very small value of 10^{-8} s is generally assigned to S_d. The general idea is that the value must be sufficiently high in order to bring the velocity of fluid in the sealing region down to near-zero values. In order to avoid an unrealistically violent closure or opening, the drag force is applied gradually modeled in Eqn 4.3 as S_f. Mathematically, it can be represented as:

$$S_f = S_\infty (1 - e^{-(\frac{t}{\tau})}) \tag{4.4}$$

S_∞ is 1 when needle seat sealing is applied, and 0 when it is not or when the needle is being opened. The τ term in this equation represents a time relaxation factor. Changing τ changes the rate at which sealing is applied or removed. The expression in Eqn. 4.4 creates an exponential decay of velocity dependent on the strength of the drag constant, S_d . The drag is turned on or off at a rate governed by τ. This gradual switching is controlled by S_f , which represents a relaxation towards the state indicated by S_∞. In the seat region, which is designated by a user input of a bounding box, the value of S_∞ is switched between zero and unity in order to represent closing or opening. For most of the domain, the value is always zero so that the artificial drag force is zero at all times. Full details on solver implementation are discussed in original work [145].

4.4.2 Geometry and Mesh Motion

The x-ray scans made of the metal Spray D target were used to generate a 3D domain for CFD computations. The changing needle lift necessitates resolving moving boundaries using adaptive meshing strategies. To accommodate this we use a boundary-fitted mesh with layering to follow the needle motion. The Eulerian HRMFoam solver had previously been modified to work with these dynamic mesh libraries, allowing for the modeling of transient needle motion during an injection event [16]. This was first demonstrated by Neroorkar et al. [152], and Baldwin [16]. In the former study, flow through a multi-hole diesel nozzle was modeled with a laminar, incompressible simulation using slip walls and upwinding convection schemes. Internal mesh motion was handled with layer addition and removal and without uniform layer stretching/squeezing. The minimum needle lift was set to 25 μm, and computed rate of injection profiles were compared to experimental data. The computational results predicted a significant influence of needle position on the internal flow structure, with low needle lift resulting in swirling flow within the nozzle sac. In addition, it was observed that hole-to-hole variation in flow rate was most significant at low needle lift positions. However, when used with the compressible implementation of the solver, artifacts of the mesh motion were produced in the flow field. Baldwin [16] parallelized the mesh motion library and added a flux correction step to HRMFoam solver to ensure that non-conservative fluxes do not adversely affect the prediction step of the SIMPLE algorithm. For modeling diesel injectors with high needle lift of upto 200 μm uniform layer stretching/squeezing might create degenerate cells at high and low lifts. Laplacian motion, while particularly suitable for low needle lift injectors such as found in the case of gasoline engines [145], are not suitable for high lift diesel injectors. Cases which require large deformations must address the problem of cell aspect ratios becoming large. This can be done through the addition or removal of cell layers. While this approach has the same structured grid limitation, if the case

starts with the lifting process and the addition of layers, the resulting mesh will inherently be structured and readily available for later removal. Layer addition/removal methods are commonly used for internal engine simulations where large deformations occur due to the motion of the piston. Baldwin [16] presented best practices for designing the meshing strategy which were followed in this study to a large extent. Granular level details on the implementation of the uniform mesh with layer addition and removal strategy are discussed in the original work [16].

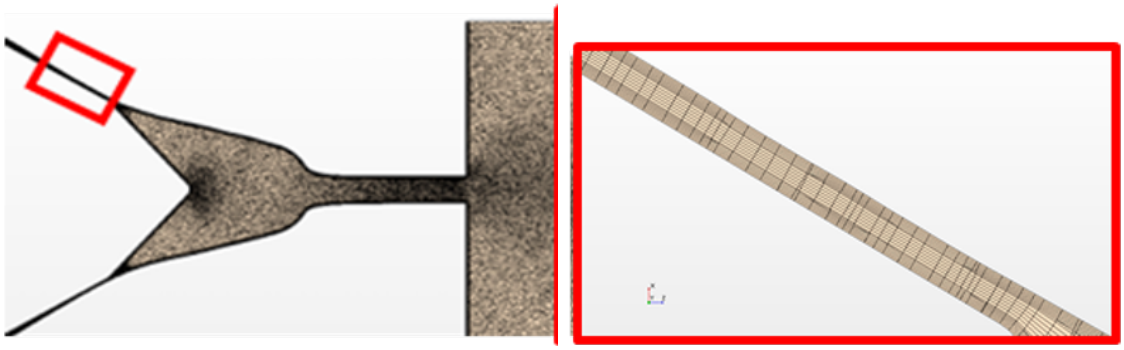

Figure 4.5. Mesh generated using a mix of hexahedral and polyhedral cells. The needle seat region is entirely composed of hexahedral cells to help with the uniform mesh layer addition/removal algorithm [16]

For the Spray D injector, the uniform mesh in the needle seat region was generated at low needle lifts as seen from Figure 4.5. Polyhedral cells were used in the other regions of the spray domain, while uniform boundary/wall cells were used to resolve the wall effects. The region of the needle seat is zoomed out to show the mesh resolution in the domain.

4.4.3 Model Details

The choice of turbulence models used in internal nozzle flow simulations has shown to have a noticeable effect on the final solution [141, 92]. For the current simulations with the HRMFoam phase change model, the $k-\omega$ Shear Stress Transport (SST) turbulence model was employed [41, 137]. In addition, a simulation with $k-\epsilon$ turbu-

Description	Details
Minimum Cell Size	2.5 μm
Total Cells	2.3 million
Turbulence Model	$k-\Omega$ SST, $k-epsilon$
Spatial discretization	2nd order Gauss linear
Temporal discretization	1st order Euler

Table 4.1. Details of numerical simulation

lence model was run to address effects due to the choice of turbulence models. The spatial discretization was second order Gauss linear and temporal discretization was first order Euler. To access thermodynamic properties HRMFoam uses a property table from REFPROP [116]. The CFD simulations were a two-phase flow problem, the fuel (n-dodecane) in a liquid as well as vapor (gas) phase, and the ambient nitrogen in the gas phase. All species were considered compressible and miscible.

Microscopic images indicated that after the end of injection, downstream charge gas (nitrogen) was entrained into the sac. Therefore, the initial condition for the next injection must include a mixture of both gas and liquid to model a realistic flow inside the injector. For the CFD models in the present work, an experimental image (Figure 4.6 right panel) was used to construct a qualitatively similar initial condition shown in Figure 4.6 left panel, below. Roughly 0.08 mm^3 or 12% of the combined nozzle and sac volume was initialized as gas. The fact that the computational initial condition has fewer, larger bubbles is deliberate, as experimentally it was found that the gas bubbles move and merge prior to the next injection, and the bubbles are randomly distributed from one injection to the next. Note that higher initial gas content in the sac (higher than 12%) is possible, as affected by expansion and compression cycles in an engine, as well as injector temperature. The boundary conditions include Pressure Inlet for the injector inlet and Pressure Outlet for the cylinder outlet. The walls have a *No-Slip* boundary conditions in order to produce boundary layer effects [190]. More details are provided in Table 4.1.

We studied the the injector dynamics mainly at two different inlet pressure conditions; with injection pressure set to 100 MPa for the baseline case, and a higher injection pressure simulation set to 150 MPa. For both these simulations, the ambient pressure was set to 2 MPa. The fuel temperature was initialized to 313 K, with the wall and gas temperature 300 K. In addition to the partial bubble initialization as shown in Figure 4.6), simulations were run without initial bubble distribution to compare effects of the former, in terms of injection delay and other flow parameters.

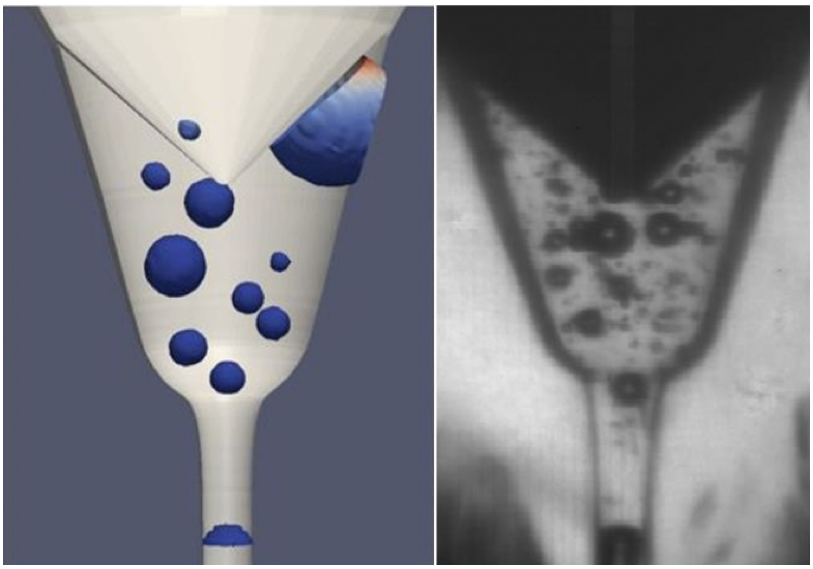

Figure 4.6. The initial condition for the computational model are based on observations from experiments showing presence of gas in the sac region. The experimental image, on the right, is courtesy Lyle Pickett (Sandia National Laboratories). Adopted from [141].

4.5 Results

4.5.1 Needle Motion

High-speed optical microscopy images and x-ray phase contrast images were processed for needle lift distance versus time. The needle lift profile for the first injection event, shown in Figure 4.7, indicates that the needle is never in a steady, seated position. Upon reaching a lift of about 170 μm, the needle immediately begins clos-

ing again. Hence, if a steady-state exists, it is despite the continuous motion of the needle. A moving needle at high needle lift does not constrain the flow, because the cross-sectional area of the orifice is smaller than the flow area of the seat region. Indeed, geometry analysis shows that when the needle lifts only 10 μm, the minimum flow area is the orifice rather than the seat passage. However, the moving needle does displace a significant amount of fuel in the sac and may have a persistent effect on the sac flow and potentially the spray. The needle lift is plotted in Figure 4.7 was used as an input for all the computational studies. In the simulations, the seat region remains sealed at needle lifts less than 10 μm, preventing communication between the flow upstream and downstream of the seat. The dashed line in Figure 4.7 indicates the region below which the sealing algorithms are employed. Because of the sealing algorithm, the injection in the CFD models begin at 0.235 ms and ends at nearly 1.2 ms. One issue with the sealing implementation below a lift height of 10 μm is that the computations may not restrict the flow appropriately, once flow is allowed, since the computational seat passage area is already greater than the hole area. Implications on the predicted ramp-up and ramp-down in rate of injection will be discussed. To better understand the quasi-steady state, we especially look at the internal flow dynamics in subsequent sections for three specific times, at 0.3ms, 0.5 ms and 1 ms that correspond to near peak rate of injection but different needle position.

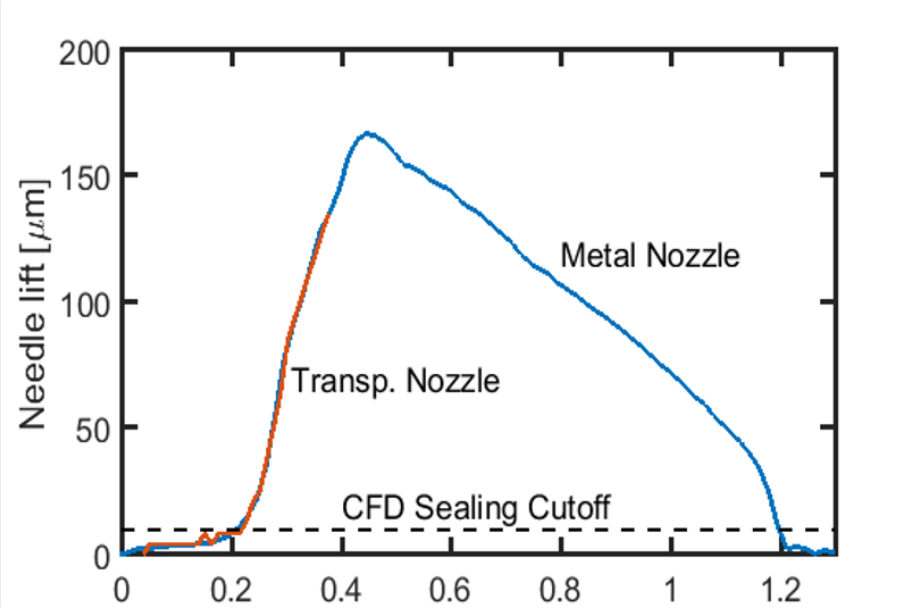

Figure 4.7. Needle lifts are measured. The overlapping behavior between metal/transparent nozzle measurements builds confidence. Adopted from [141]

4.5.2 Steady State

Figure 4.7 shows that the needle is never stationary. The nozzle discharge in Figure 4.8, does achieve a period of steady flow. This feature of the rate of injection (ROI) being remarkably constant, shortly after needle opening, can be attributed to the fact that as the needle moves away from the seat and creates a wider opening for the flow to pass, and as the flow opening at the needle seat grows to equal the nozzle exit area, the discharge is constant thereafter irrespective of the position of the needle itself. Thus, it shows that the needle motion itself has very little effect on the flow rate at the nozzle exit and is only a function of the opening at the seat. The rate shape model [58, 161] data for the baseline case, made publicly available by ECN shows remarkable consistency once the full value has been achieved. The computational predictions of mass flow from HRMFoam (for both turbulence models) agree very closely with the rate shape data during the quasi-steady part of injection. Though the CFD model has a faster ramp-up compared to the rate shape model,

the ROI remains nearly constant, once it reaches the peak value. By this point, all initial bubbles have been ejected, as evidenced by the uniformity of the mass flow rate. This level of agreement during quasi-steady state is relatively easy to achieve, because of the modest amount of phase change within the nozzle given its smooth features. However, there are some discrepancies at the start and end of injection. As a full disclosure, we obtain the rate shape model from [58, 161] and shorten it to our injection duration (first injection) of about 1ms. Previous experiments [161] on similar single-hole diesel injectors have shown that the ramp-up/ramp-down take about 0.2-0.3ms. In order to be consistent with the observed behavior, we retain the duration of the ramp-up/ramp-down phase of injection and shorten the quasi-steady state, leading to the arched behavior of the experimental curve in Figure 4.8. It also leads to difference between the CFD predictions and experiments in how quickly the ramping occurs. For more details about the experimental measurements please refer [58, 161]. To further explore the reason for the discrepancy between the CFD model ramp-up (transients) and the rate shape model, we plot the ratio of the needle flow passage area, around the needle seat, to the orifice area as a function of the needle lift in Figure 4.9. At the CFD sealing cut off value of 10 μm, the ratio of flow/orifice area is close to 1. This indicates that as the sealing is deactivated, the flow coming through the needle seat is equal to the orifice area and therefore the peak rate of injection value is reached almost instantly. This is contrary to the observations of the rate shape model, predominantly because they do not consider the possibility of sealing the needle seat while the needle lift is low (about 10 μm). This could also help explain the ramp-down shape disagreement between the numerical code and the rate shape model, where CFD predicts a much sharper ramp-down and an abrupt shut-off occurs. Some remnant flow might still exit the nozzle because of the pressure difference. For our simulation, the sealing algorithm [145] has a timescale term, τ that can be modified to determine the sensitivities of the effects on the abruptness

of the needle shut off. This is an important observation about the needle transient behavior and points to the key disagreements between the experimental model fit and observed CFD behavior.

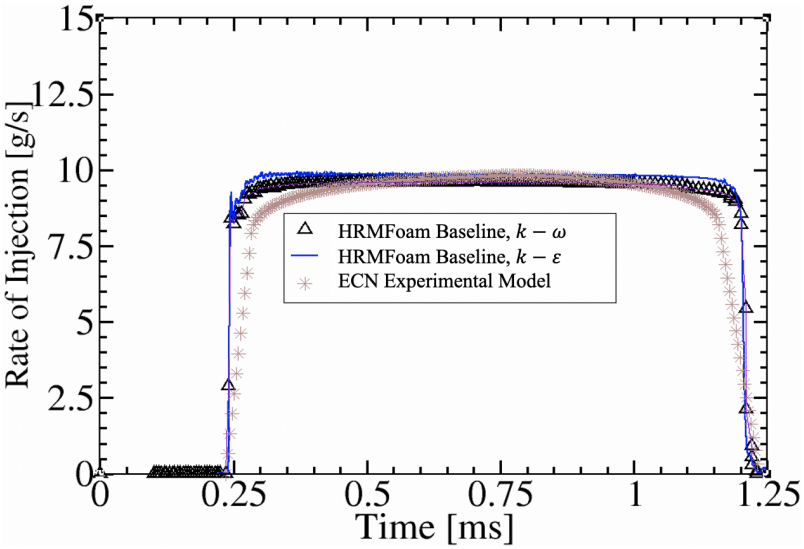

Figure 4.8. The fuel discharge or rate of injection, between computational models and experiments show similarities in prediction during the quasi-steady phase. However the transient phase of injection differ from experimental model [161]

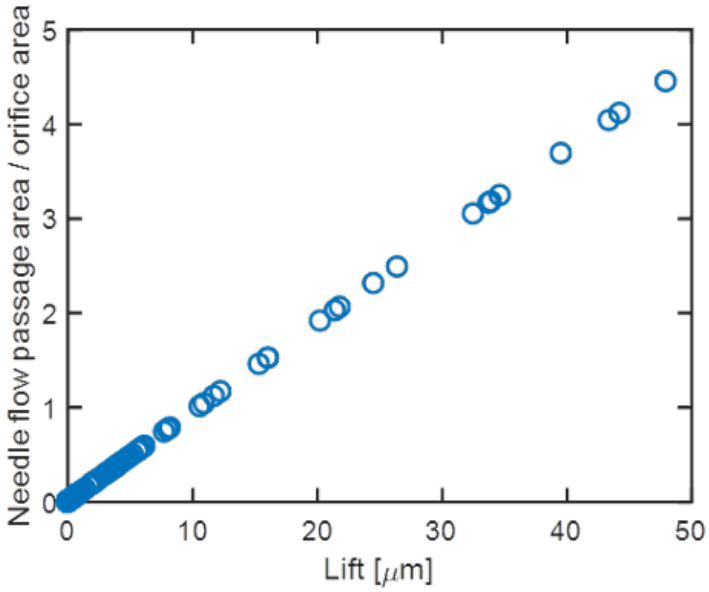

Figure 4.9. The effect of needle lift on the ratio of flow passage to orifice area shows an almost linear relationship. At the CFD sealing cut off range of 10 μm, the ratio is 1 which can help explain the faster ramp-up as seen in the CFD codes.

After the nozzle discharge, another important metric to characterize a nozzle performance is the momentum flux. For this characterization, we present results from a previously published research [141], wherein comparisons between slightly different implementations of HRM model were made. One is the HRMFoam model that has been introduced previously, and the second is CONVERGE-CFD, a commercial solver. This is done to provides an assessment of reproducibility of the computational results. Figure 4.10 shows the steady state momentum figure reproduced from [141]. The start and end of injection are marked by black vertical lines. Both CFD predictions show a decline in momentum after 1ms. This is surprising since the nozzle discharge (Figure 4.8) does not show this attribute. This means that there is a decline in the flow velocity. This can be manifested in a change in the shape of the profile of the velocity. At 0.5 ms, as shown in Figure 4.11, the flow is completely unrestricted by the needle and seat spacing. The velocity remains low as the flow passes this gap without significant acceleration. However, by 1 ms, the lift has diminished to the

point where a thin, annular jet has formed by the small gap. This jet accelerates to a velocity of approximately 75 m/s. Though this is still substantially lower than the velocity in the nozzle, the velocity still corresponds to a significant amount of dynamic pressure that is lost by the uncontrolled deceleration of the jet. The contrast between snapshots at t = 0.3ms (Figure 4.11 top panel) and t= 0.5 ms (Figure 4.11 middle panel) suggests that the annular jet's trajectory is not necessarily just a function of needle lift. Apparently, the jet can adhere to either the needle surface or the sac walls, depending on the history of the flow. This tendency of the jet to follow the wall resembles the Coanda effect and was observed at needle lifts of less than 20 microns. Once the flow approaches the injector nozzle, the flow experiences a strong favorable pressure gradient that will straighten the flow and largely erase the consequences of the jet. Hence there is little observable effect on the downstream jet other than the reduction in momentum shown in Figure 4.10 at about 1ms. Both CFD models appear to predict momentum close to the experimental validations [161], more in Table 4.2.

P_i (MPA)	P_j (MPA)	HRMFoam	CONVERGE	Experiments
100	2	4.71 N	4.55 N	4.59 N
150	2	7.1 N	6.9 N	7.1 N

Table 4.2. Predicted spray momentum compared to the experimental measurements [161]. P_i stands for injection pressure, and P_j for ambient/chamber pressure

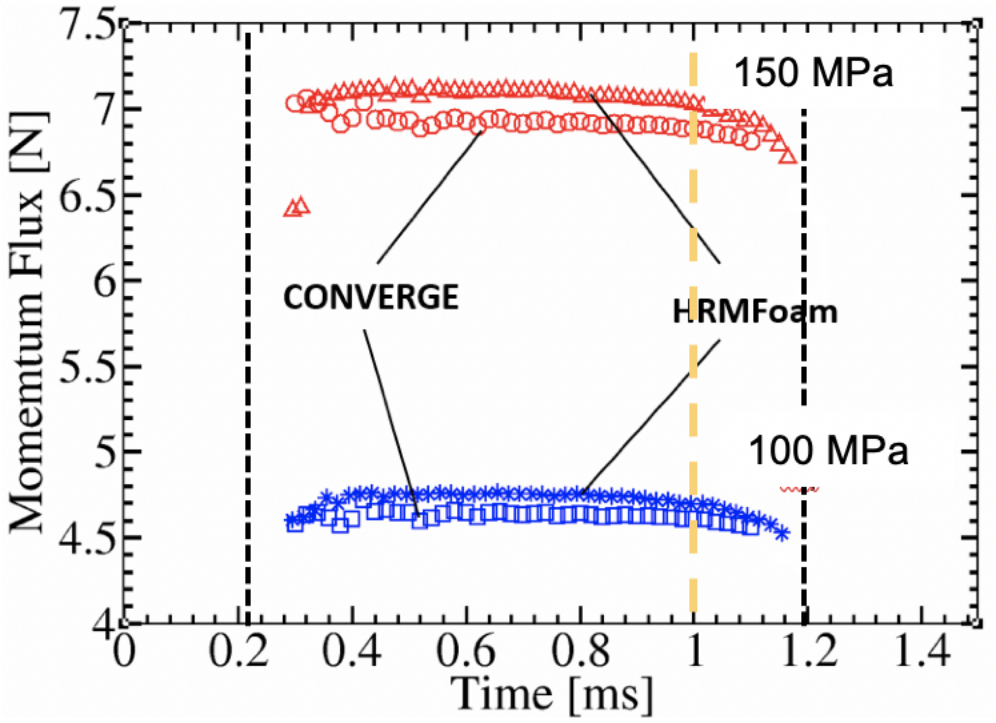

Figure 4.10. Momentum Flux comparisons between CFD models during the quasi-steady period show close agreement for two different injection pressures, however, diminishing momentum during the needle closure indicates loss of momentum head due to needle motion. Dashed vertical lines at 0.235ms and 1.2ms indicates start/end of injection respectively, in CFD models due to needle sealing.

To investigate this further, slices at transverse locations within the nozzle were taken and the vorticity field plotted in the middle panels of Figure 4.11 from the HRMFoam simulations for the baseline conditions, to understand the tendency of the fluid to rotate/ circulate because of the needle transients. During the early and late needle transients, especially when the lift is low at 0.3 ms and 1ms, higher vorticity is induced within the nozzle (Figure 4.11 top and bottom panel), compared to higher

needle lifts corresponding to 0.5 ms (Figure 4.11 middle panel) from the needle lift curve, when we see a quiet vorticity field. Therefore, the effect of nozzle displacement may have an impact on the spray development including momentum, although the ROI is nearly constant, even in a smooth, symmetrical nozzle like Spray D 134. This observation is consistent with a previous study, which explored the onset of string cavitation in a single, straight hole injector nozzle and attributed the same to a swirling upstream flow and the needle lift [172]. Although the current study has not explored the string cavitation aspect, Figure 4.11 shows the effect of the needle lift on producing a swirling flow inside the nozzle and therefore could answer the reduction in momentum as seen in Figure 4.10. Additionally, research on multi-hole injectors have shown prevalence of string cavitation during low needle lifts [54].

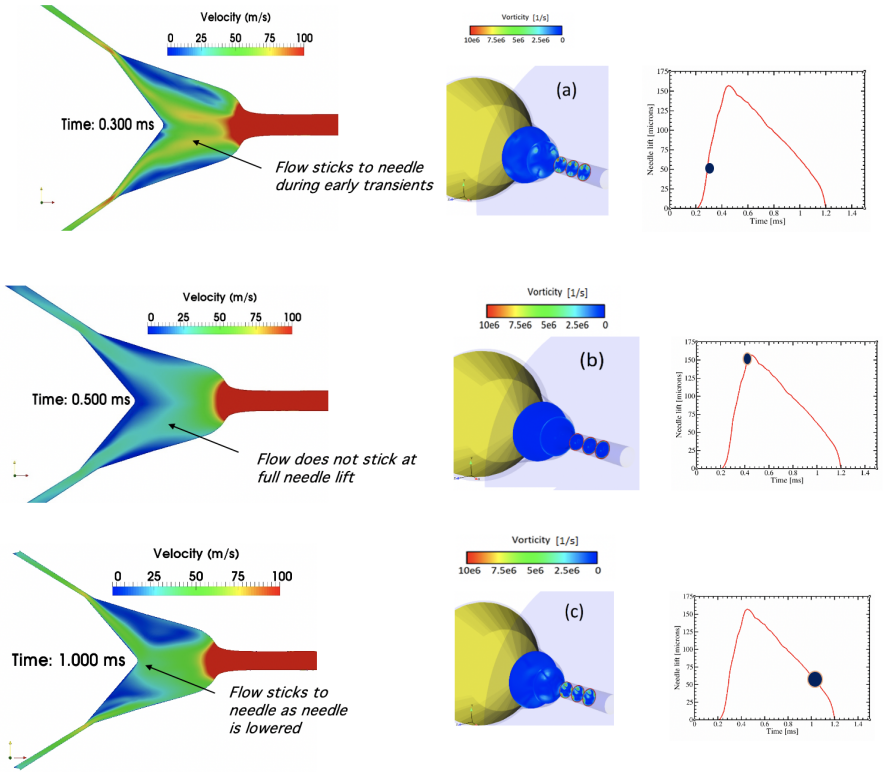

Figure 4.11. The effect of the transient needle motion is clearly evident on the internal flow during the quasi-steady state

Using $\lambda - 2$ criterion [90], we verify the previous observations in Figure 4.12. The vortex core iso-surfaces are colored by liquid volume fraction. Similar to observations from Figure 4.11, at low needle lifts the flow in the sac appears to be more chaotic with vortex cores stretching into the nozzle hole whereas at higher needle lifts, the vortex cores in the sac are limited to the upstream region.

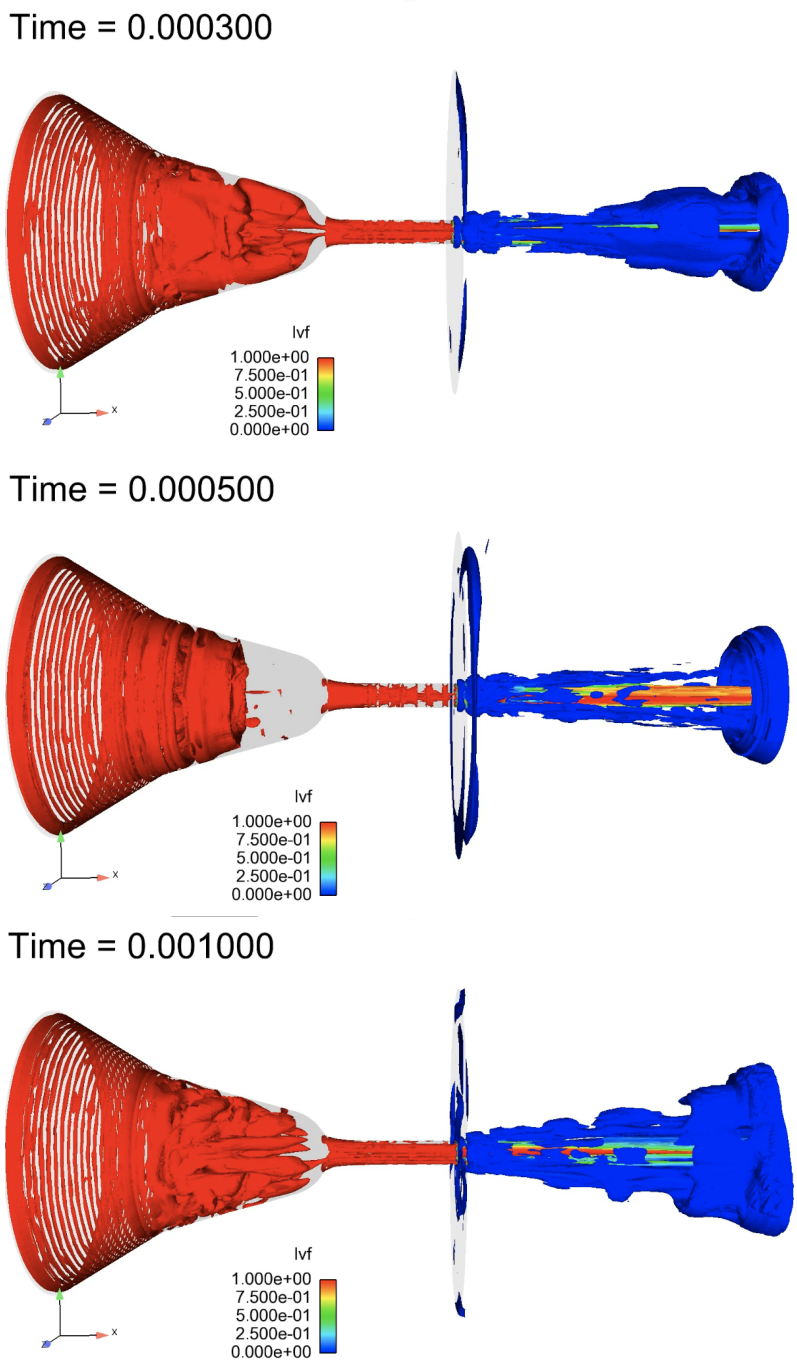

Figure 4.12. The $\lambda - 2$ vortex cores show similar unsteady internal nozzle flow behavior seen from Figure 4.11. Note all times are in seconds.

4.5.2.1 Two phase flow within the nozzle

A single microscopic image obtained during the steady period of injection is shown at the top of Figure 4.14. At this timing, the needle has lifted, and liquid is flowing through the sac and the hole. Because the refractive index of n-dodecane (n = 1.43) is close, but not exactly the same as the acrylic (n = 1.51), the interface between liquid and the acrylic produces contrast. Inside the nozzle, bubbles formed from cavitation, or the release or mixing with nitrogen (non-condensable) gas, will scatter light and appear dark when backlit. Outside of the nozzle, liquid droplets or any curved liquid surfaces will also scatter light and appear dark when backlit. This schematic is shown in Figure 4.13.

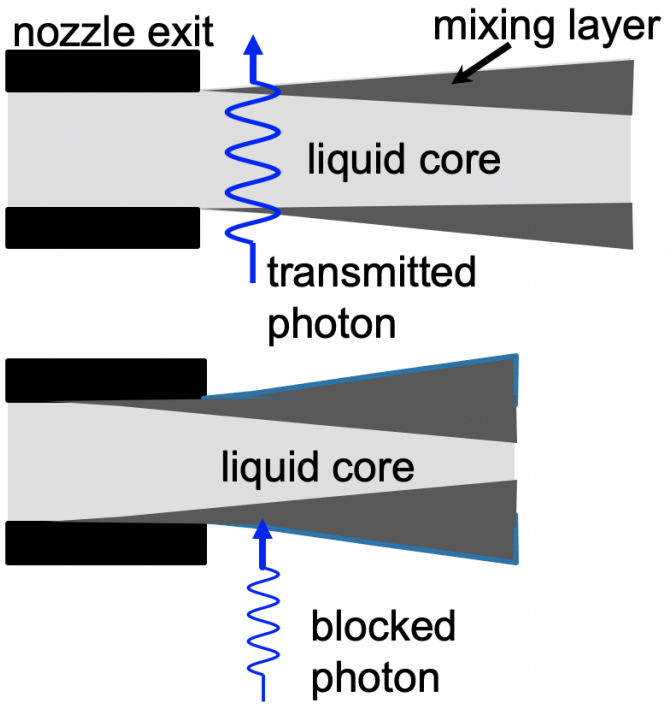

Figure 4.13. Schematic of the optical microscopy liquid core identification. Credit Manin et. al. [130].

Examination of the images, at different time instances, reveals that a fluctuating dark zone begins approximately 0.2 mm upstream of the nozzle exit. High-speed movies for this figure are available to view at [133, 131], which are more convincing.

A dark region shown at the inlet of the nozzle does not fluctuate, and we believe this to be caused by small micro-cracks induced by pressurization of the nozzle, rather than a flow feature. Though there are optical difficulties near the nozzle exit, fluctuations of these dark region are clearly observable. The transient and rapidly shifting nature associated with a fluid dynamic phenomenon was assessed by processing the images for the standard deviation over the steady period of injection as shown at the bottom of Figure 4.14. Ruling out the high standard deviation region that is associated with the moving needle, the most prominent fluctuations are found at the nozzle exit. Though the nozzles were designed to avoid cavitation by imposing a strongly converging shape, two-phase flow is detected in the transparent injectors. While cavitation induced by local divergence or wall roughness of the orifice is a likely cause for this optical effect, the rapid divergence near the nozzle exit could also cause some ingestion of non-condensable nitrogen gas if the jet separates from diverging walls. However, other acrylic nozzle samples were tested for consistency and even nozzles with almost zero divergence at the exit demonstrated evidence of two-phase flow. Moreover, movies showed transient fluctuations that were not necessarily connected to the downstream gas. These observations suggest that the two-phase flow in the nozzle is cavitation.

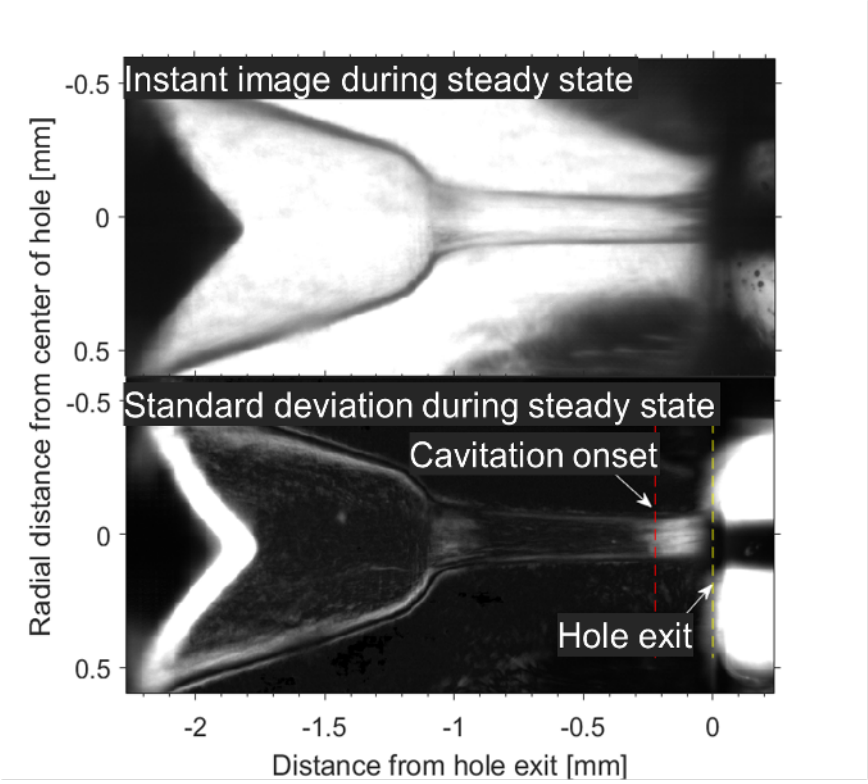

Figure 4.14. High speed microscopy showing the extent of cavitation inside the nozzle, beginning at about 0.2 mm upstream of the nozzle exit during quasi-steady phase of injection, identified by the standard deviation of the image set. Injection pressure = 100 MPa, discharge pressure = 2 MPa. Adopted from [141]. Experiment credits to Pickett et al.

Desiring to understand how the internal cavitation may affect spray development, we compare simultaneous images from the two orthogonal perspectives at the top and middle of Figure 4.15. The image from Figure 4.14 was taken with focus on the hole region to best measure the internal flow features, but it is obvious that the downstream spray is out of focus. By contrast, the middle picture in Figure 4.15 shows the spray in focus while the internal features are out of focus. The picture is also displayed with lower peak intensity (25% of the intensity outside of the jet), to highlight that the exiting jet has a small region of "transparency" at the nozzle exit. Pairing of the top and middle images at the same instant is useful in that the top definitively shows a region of cavitation internal to the nozzle while the middle shows

that this layer of cavitation, perhaps surprisingly, does not make the emerging spray completely opaque.

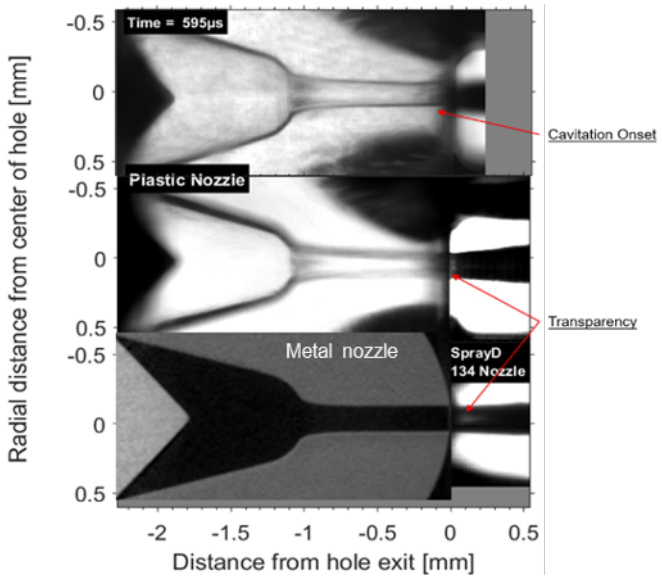

Figure 4.15. Dual line-of-sight images of acrylic-nozzle (top and middle) and an image of the spray emerging from metal-nozzle Spray D 134 with the same back lighting (bottom), during the steady period of injection. The metal nozzle is obviously opaque, but the metal nozzle tomography reconstruction on the plane of the hole (with the needle closed) is overlaid at the bottom as reference for the measured emerging spray. Injector and discharge conditions are identical to Figure 4.14. High-speed movies corresponding to this figure are available at [133, 131]. Adopted from [141].

Considering the processes of light scatter and refraction from liquid surfaces, we understand that a layer of droplets, bubbles, and liquid surfaces with curvature must be very small, and it must exist on both sides of the jet simultaneously, for there to be light transmission of significance through the spray [131]. The finding that some internal cavitation may exist, with the effect of promoting distortion to liquid surfaces, while simultaneously the emerging spray is at least partially transparent, is a new discovery. Analyzing results from other transparent nozzles with different propensity for cavitation, we find that the emerging spray does become opaque if the cavitation layer is found farther upstream within the nozzle but is in fact less opaque if the cavitation layer moves closer to the nozzle exit. With this relationship

between internal flow cavitation and the emerging spray appearance established, it is interesting to consider the transparency of the emerging spray for the metal nozzle target, as shown in the bottom of Figure 4.15(at the same image display range as the middle image, 0-25% of the intensity outside of the jet). An image during the steady period of injection shows that the metal Spray D 134 also produces a partially transparent liquid region, and the transparency is of similar magnitude as the acrylic nozzle. While the nozzles are clearly not identical from the standpoint of geometry, the emerging spray appears similar. This similarity supports the notion that the acrylic nozzle is a reasonable representative of the metal nozzle. More importantly, it suggests that internal cavitation is also occurring within the metal nozzle.

We caution that these results are specific to the geometry of nozzles for this study. While the nozzles are designed with inlet rounding and convergence, the reality is that the nozzle has a region of weak convergence, or possible local divergence. Other nozzles with stronger convergence throughout the hole may exhibit different characteristics, showing no cavitation within the nozzle. A very small amount of cavitation was predicted by HRMFoam near the exit, as shown in Figure 4.16, but not to the extent observed in the acrylic nozzle as shown in Figures 4.15. It is generally accepted that the geometry of the nozzle, especially the corner radius has a pronounced effect on the onset of cavitation within the nozzle [200]. Figure 4.16 shows no evidence of cavitation originating from the nozzle inlet corner, consistent with the microscopy studies on the transparent target in Figures 4.15 and other experimental studies [200]. This further demonstrates that the cavitation observed at the nozzle exit, predominantly from experiments and HRMFoam, is not because of any upstream source and is rather due to local effects near the nozzle exit. This discovery is significant because it identifies a fundamental source for spray breakup, growth of the initial mixing layer, nozzle flow coefficients, and the overall plume spreading angle.

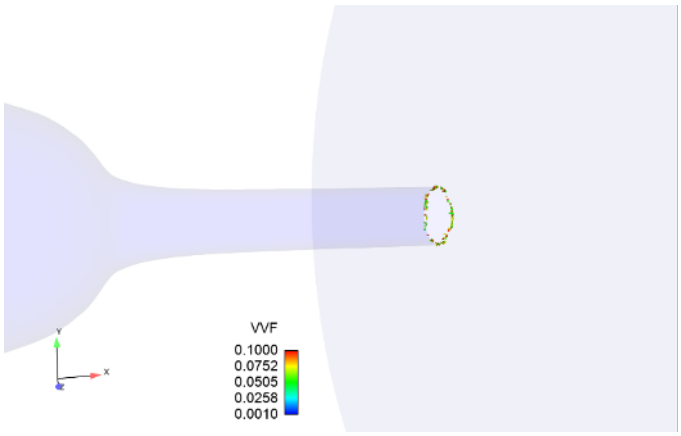

Figure 4.16. An iso-volume of vapor volume fraction (VVF) between 0.001-0.1 shows formation of vapor at the nozzle exit during peak ROI at t = 0.5 ms, however not to the extent observed in experimental images in Figure 4.14

4.5.2.2 Near Nozzle Comparisons

The internal nozzle flow comparisons yield important validation of model predictions. An other source of model validations are comparisons in the near-field region. This region is typically defined as within 10 mm downstream of the nozzle exit. While direct experimental observations in the dense core region is often times difficult, certain line-of-sight projection methods have proven to extract useful information. The projected mass density from numerical models are compared against the experimental radiography measurements conducted at Argonne National Laboratory [133]. The variable used for comparison is the projected mass density of the fuel, which is calculated by a line-of-sight integration along the X-ray beam [104, 105]. A similar procedure is replicated with the data from simulations to enable fair comparisons against experiments at 2mm downstream of the nozzle exit. To further validate the HRMFoam predictions, comparisons between engineering models and high fidelity sharp interface predictions such as in CLSVOF (for full model details refer [13, 12]) are made. The experimental observations show an asymmetry in the projected mass data with a slight off center profile. While no computational model is able to ac-

curately capture this asymmetry or eccentricity, CLSVOF matches the spray width successfully. The reduced order RANS based HRM models are zero centered, and HRMFoam performs better compared to CONVERGE in following the spray profile. However it appears both models are unable to resolve the eccentricity. This can be attributed to missing some phenomenological information due to the simplified model assumptions. However, these results are encouraging as the reduced order models are computational cheaper (compared to CLSVOF) and provide a general approximation of the spray behavior successfully.

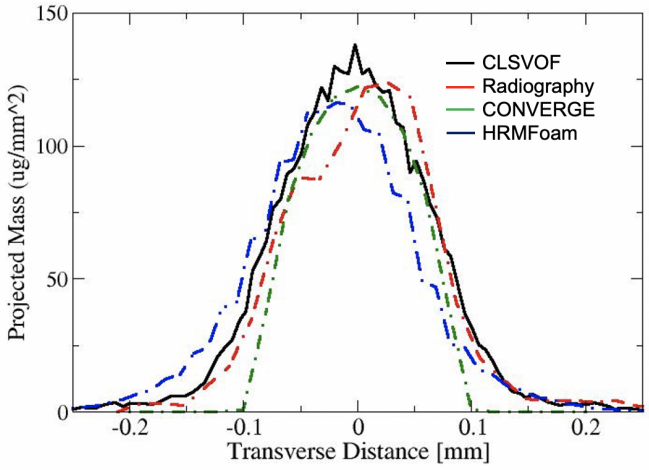

Figure 4.17. Measurements at 2mm downstream of the nozzle exit, for 100MPa injection pressure case shows differences between model predictions and experimental observation. CLSVOF simulations were performed by M. Arienti (Sandia National Laboratories) and CONVERGE-CFD simulations were performed by P. Srivastava

The interfacial area density Σ can be used to identify edges of a spray. These plots are expected to produce a bimodal distribution of data in the near field region, due to the nature of spray intact core. Figure 4.18 shows a schematic by comparing to available experimental data (not relevant to Spray D) [104].

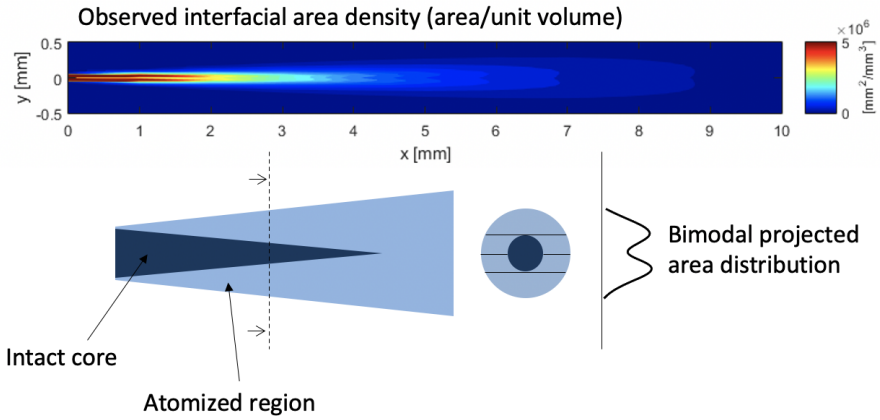

Figure 4.18. Interface area calculation schematic. Credit M. Arienti (Sandia National Laboratories)

While experimental measurement techniques such as ultra small angle x-ray scattering or USAXS, can be used to measure the Σ, in this work comparisons between HRMFoam predictions and high-fidelity CLSVOF models for interfacial area density are made. The CLSVOF simulation data is resolved at 10 x 10 μm resolution. It reveals that the bimodal structure of Σ is essentially captured by the HRMFoam simulation as seen in Figure 4.19. There is therefore an opportunity to re-calibrate the Σ model parameters either by fitting to data or by using data-driven/machine learning techniques to parameterize the model.

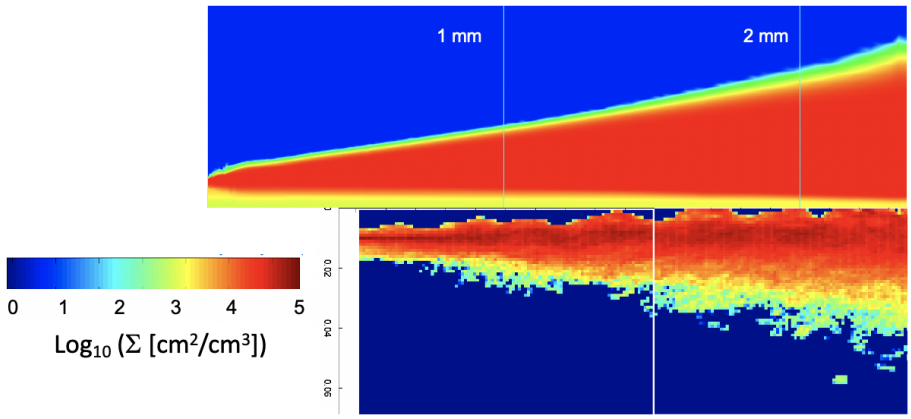

Figure 4.19. The prediction for Σ shows the presence of the bimodal structure in the spray for HRMFoam (top panel) in the near nozzle region. Due to the reduced order nature of the model, the edges are smeared compared to the high fidelity LES/CLSVOF simulation (bottom). Both these data are scaled similarly. [LES/CLSVOF simulation credit M. Arienti (Sandia National Laboratories)]

In the end, fuel density at three different locations downstream of the nozzle exit are made in Figure 4.20. For HRMFoam predictions, the projection of the nozzle exit are shown in the form of a dotted circle. The axial locations of comparison are at 0.1mm, 1mm, and 2mm downstream of the exit. Closer to the nozzle (top panel), the predictions between HRMFoam and experimental observations are comparable, however while the radiography data shows an onset of asymmetry the model predictions do not exhibit this behavior. At 1mm downstream (center panel), the radiography profile shows a severe asymmetry, which HRMFoam fails to capture. The overall profile of a thinner intact core is reproduced by HRMFoam predictions to some degree. At the farthest downstream location of 2mm (lowest panel) while the radiography data shows evidence of jet disintegration, the HRMFoam still predicts an almost intact core. A few reasons for the failure in capturing the eccentricity from Figure 4.17 and assymmetry in Figure 4.20 can be attributed to the coarser mesh resolution in the downstream region since most of the meshing was focused in the internal nozzle and close to the nozzle exit region. The computational domain for HRMFoam extended

until 2.2 mm downstream of the nozzle exit and therefore there may have been issues related to the pressure-outlet boundary effects close to the probe location. The asymmetry in the spray morphology was most likely caused by localized cavitation in the nozzle exit region. While HRMFoam is able to detect cavitation to some extent, there appears to be missing phenomena that needs to be better resolved.

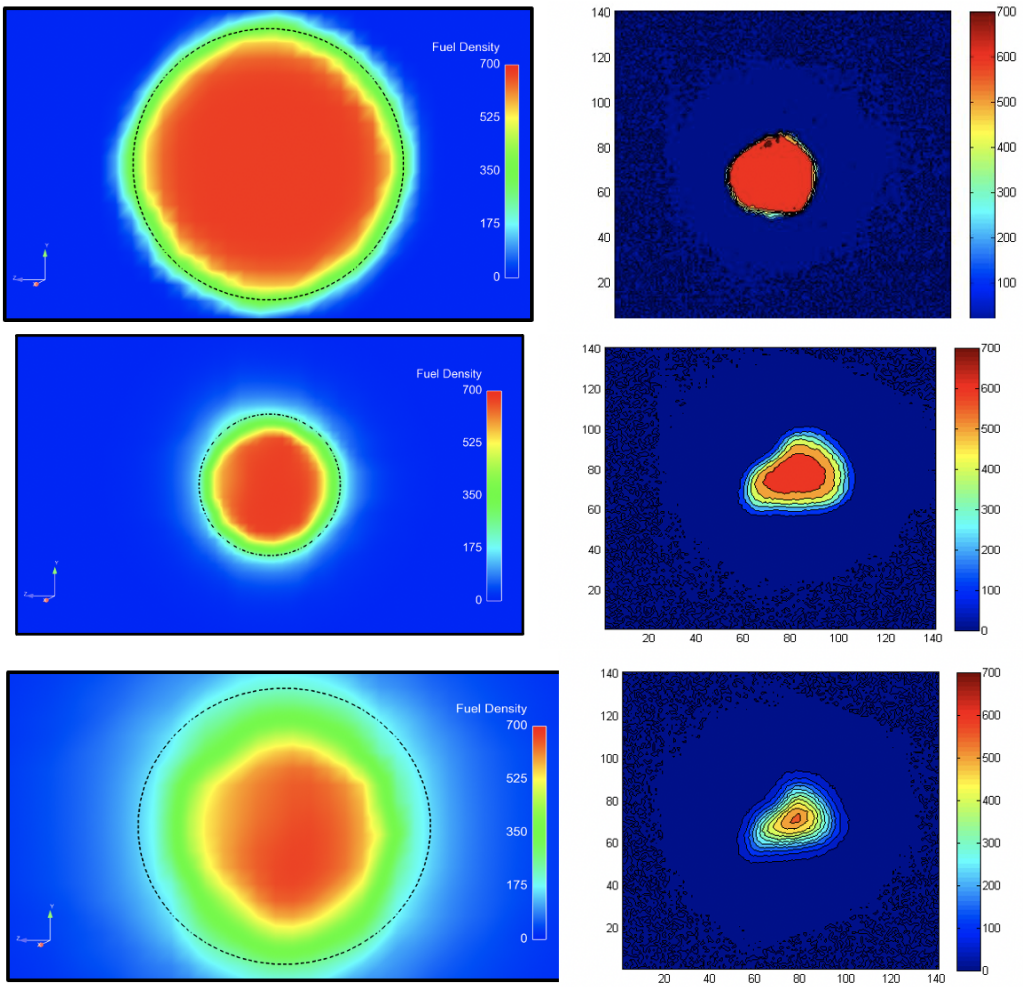

Figure 4.20. Comparisons between HRMFoam and radiography measurements reveal key differences in the predictions including spray morphology.

4.5.3 Transient phase of injection

4.5.3.1 Effect of initial conditions on early transients

We study the effect of initial conditions, namely a partially filled gas initialization and no gas initialization, by plotting the volume fraction of the phases within the sac and nozzle in Figure 4.21 from the HRMFoam simulations. The start of flow into the sac in the CFD model due to sealing is at 0.235ms. It demonstrates that there is little evidence of the initial condition bubbles impacting discharge beyond 0.24ms as the volume fraction curves from the 12% gas and 0% gas simulations overlap after 5 μs, indicating the similarity in the sac and hole conditions. This is further shown in the midplane clip colored by the liquid volume fraction in Figure 4.22. The persistence of the initial condition, inside the nozzle, is considerably less than the time required for the mass flow rate to achieve its maximum value. The volume fraction for each phase within the sac and nozzle region is as per equation 4.5.

$$Vf_{phase} = \frac{V_{phase}}{V_{sac+nozzle}} \tag{4.5}$$

where Vf is the fraction of volume, and V is the actual volume.

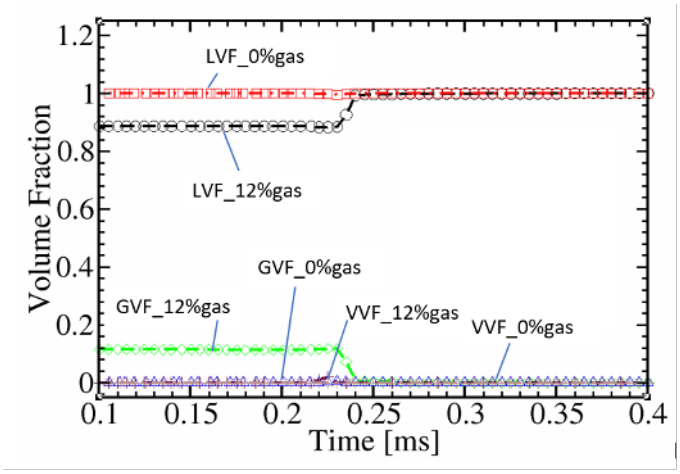

Figure 4.21. Volume fraction for all three phases - liquid, vapor and gas, during the early start of injection transients showing no effect of initial gas in sac after initial bubbles are compressed/dissolved in the liquid. LVF, VVF and GVF indicates liquid, vapor, gas volume fraction respectively. 12% gas and 0% gas indicate with gas bubbles and a full liquid initialization

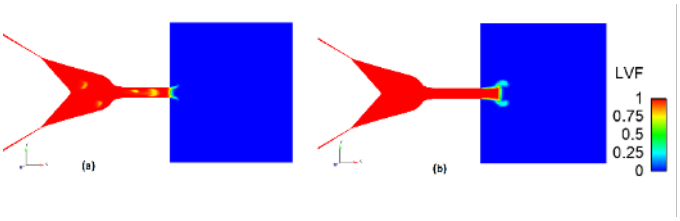

Figure 4.22. Effect on rate of injection due to presence of gas in sac and nozzle during early start of injection transients. The effect lasts for a very short time after which the 0% gas and 12 % gas rate of injection curves overlap.

Gas bubbles trapped in the sac at the end of injection, and held at pressure, dissolve into the liquid fuel at the start of next injection, as seen in Figure 4.22. With gas present within the sac, the ramping rate-of-injection profile and timing is modified compared to a sac full of liquid [245, 141, 131] shown in Figure 4.22. Manin et al. [130, 131] concluded from their experimental study that presence of gas strongly affects the injection rate, and this effect lasts for at least a few hundred microseconds. The prominent effect of the presence of bubble initialization is on the injection delay of the spray. Figure 4.22 further elucidates this by showing LVF snapshots at 5 μs

after start of injection. Figure 4.22, depicts the gas bubble initialization case shows a lagging spray compared to the no gas initialization simulation. This is remarkably consistent with the observation of [130], which concluded the time required for the flow to clear the nozzle of gas depends on the quantity of gas thereby causing an injection delay. Also, the presence of gas causes slight differences in the spray growth and development, consistent with the experimental observations on the same nozzle [130].

We demonstrate these phenomena by simulating three different conditions. Two of these conditions, with no residual gas and 12% residual gas has been simulated and results validated in Figure 4.8. The third case involves simulating an initial condition with all gas sac. As seen from the nozzle discharge plot in Figure 4.23, while there are obvious injection delays in the case of minimal residual (12%) gas condition, the ROI curve collapses on the no-gas curve within few microseconds as the gas dissolves in the liquid fuel. On the other hand, the simulation with all gas takes a a few hundred microseconds to reach 80% of the peak value. The flow coefficients (from Figure 4.24) indicates an interesting dynamic. For a converging, almost non-cavitating nozzle, the coefficient of area C_a is expected to be near 1. That is indeed the case for the no gas, and 12% gas simulation. However, for the 100% gas simulation, it appears that dissolution of gas in the liquid fuel is a much slower process and therefore the C_a has a much slower ramp up.

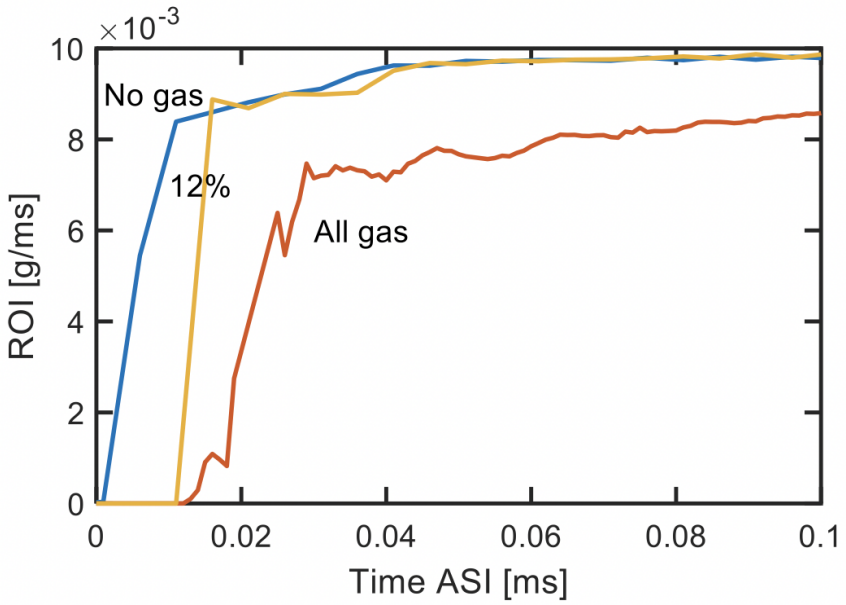

Figure 4.23. Effect of gas in sac on injection characteristics show as the amount of gas at SOI increases, the injection delays are much longer. This is consistent with experimental observation [131].

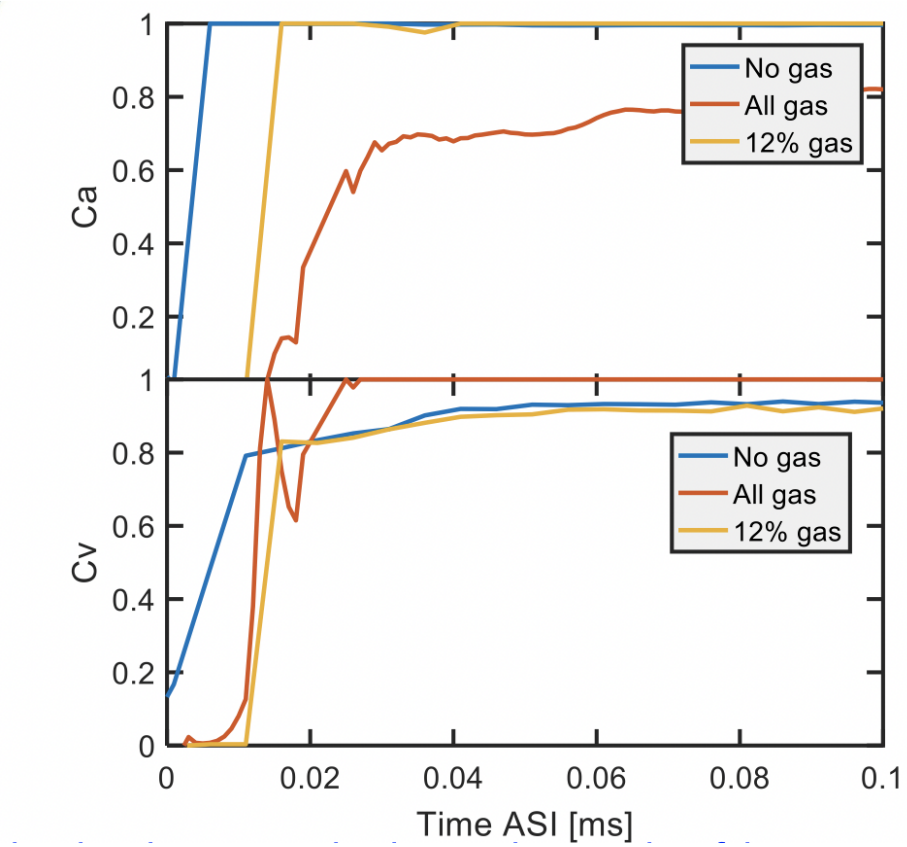

Figure 4.24. Effect of gas in sac on flow coefficients, C_a and C_v

4.5.3.2 End Of Injection and post-injection

For the Spray D simulation, the end of injection occurs at about 1.2 ms after start of injection. The end of injection, and the subsequent dwell phase are modeled as the needle closure event using the sealing algorithm as discussed previously. The optical microscopy experiments show bulk cavitation after the end of injection (middle and right panel in Figure 4.25). At the end of the injection, there is evidence of bulk cavitation inside the sac, which dissipates slightly and gas ingestion in the nozzle hole is seen (right panel). In the experiments the gas phase cannot be identified in terms of vapor phase fuel or charge gas. These set of observations are validated in the HRMFoam simulations in Figure 4.26 and 4.27. Almost right after the end of injection, evidence of bulk cavitation is seen in the simulation - left panel in Figure

4.26, and Figure 4.27. However, during the dwell phase the gas in the sac equilibriates with the outside pressure followed by the charge gas ingestion, shown in the right panel in Figure 4.26. At this time, the bulk cavitation or remnants thereof occurring with the sac from experiments is not replicated in the simulations fully. The charge gas reaching the sac is a significant discovery, as it will affect the subsequent injection. There are several challenges in calculating the gas volume in the sac.

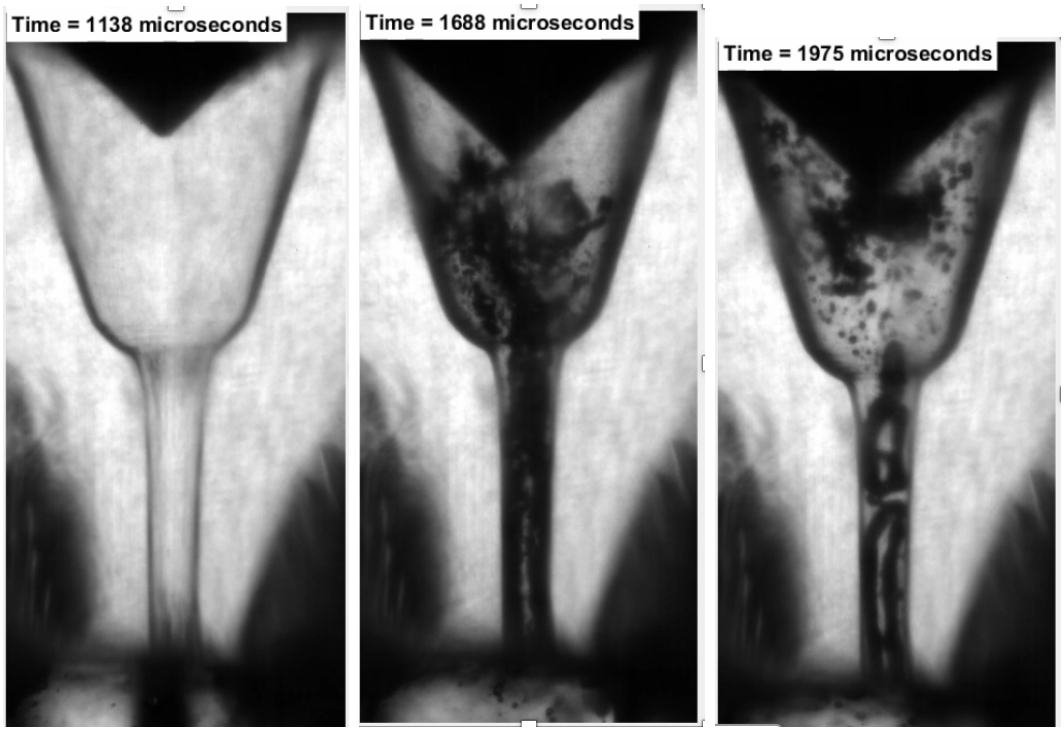

Figure 4.25. Experimental observations from end of injection show presence of gas in the sac. Unfortunately these results cannot distinguish between charge-gas or bulk cavitation. Adopted from [141].

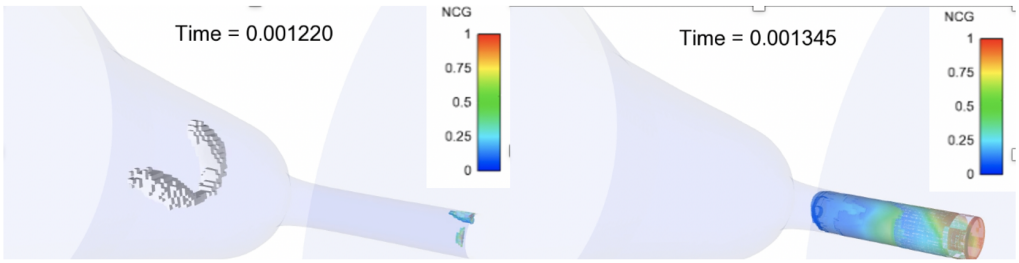

Figure 4.26. After the EOI, bulk cavitation is seen to occur within the sac and charge-gas ingestion occurs in the nozzle hole

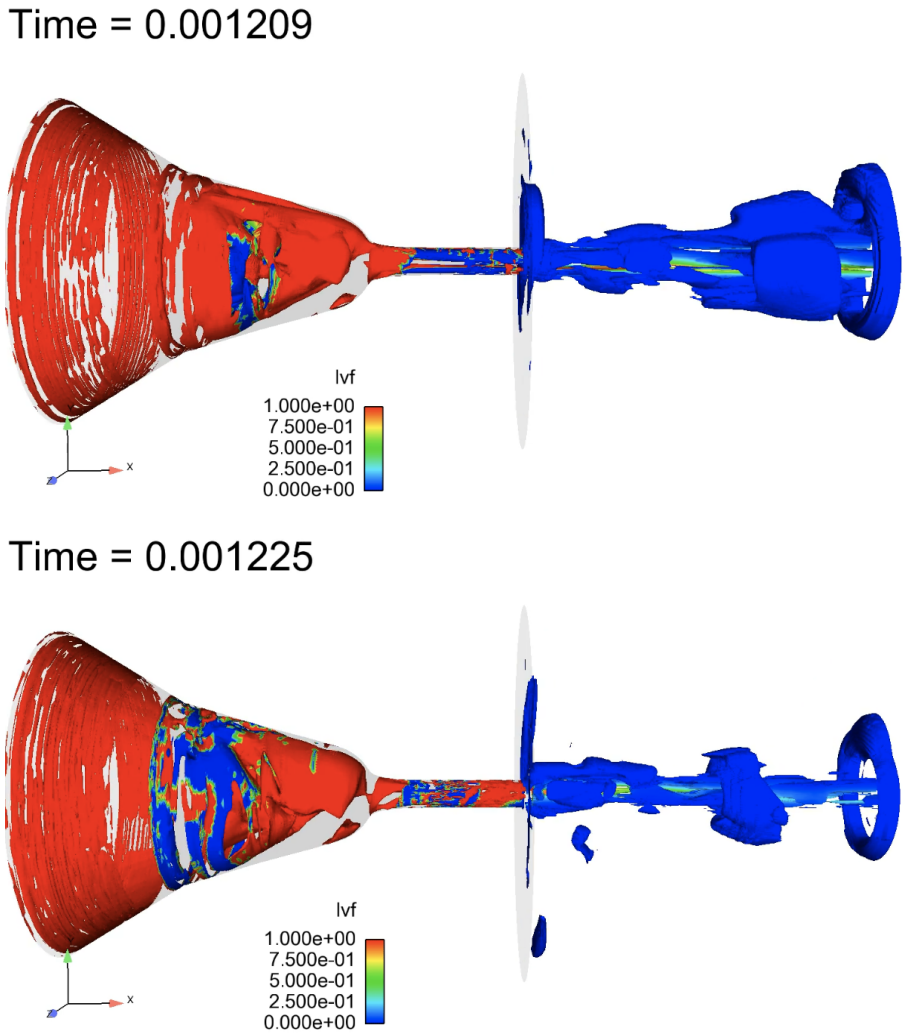

Figure 4.27. The $\lambda - 2$ criterion, colored by LVF, shows the presence of gas in the sac shortly after the end of injection. This phenomena is extremely transient and short-lived.

While qualitatively the HRMFoam simulations appear to be consistent with experimental observations, limited quantitative comparisons have been made to validate the model predictions. It is non-trivial to identify and characterize quantitatively the gas in the nozzle from experimental observations. Primary challenges include the transient nature of the process as well as limitations in the experimental setup in identifying the different phases of the gas. Gas inside the sac appears as dark pixels in images (due to non-matching refractivity index). Gas tends to form spherical bubbles, and therefore the volume can be easily be calculated from the radius calculated using edge detection algorithms. However, is to be noted that these results tend to be approximate because of the dissolved gas that is hard to measure, and that the needle is on the line-of-sight, as well as some cases of overlapping bubbles in the line of sight. The mechanism can be described as the injector closes, the flow of high pressure fuel supplying the sac with liquid is decreased and then cut off, causing a rapid pressure drop in the nozzle. The sudden pressure drop initiates bulk cavitation. Once the pressure equilibrates, the cavitation collapses and the fuel in the sack moves into these zones, which pulls fuel from the hole into the sac, which subsequently ingests gas (nitrogen) from the spray chamber into the nozzle tip. Data show significant dependencies upon the needle closing rate, injection pressure, hole shape (cavitating or smooth), and back pressure. Further the residual amounts of gas in the sac rapidly expand as soon as the sac equilibrates with the outside pressure. As the liquid fuel exits the orifice, the pressure inside the sac decreases below the outside pressure. Fuel velocity decreases, the liquid jet becomes laminar and dribble begins. As the cross-section of the orifice is not fully occupied by the liquid phase, gas begins to replace the volume left by the liquid.

Figure 4.28 reports the experimental observation and HRMFoam predictions. The experiments have been conducted at an injection pressure of 100 MPa and the ambient pressure of 2 MPa. To demonstrate the primary effect of the injection pressure,

we compare the predictions from the baseline to the high injection pressure simulation and note the relatively minor effect of the injection pressure on the prediction. While the CFD does predict gas inside the nozzle, the quantitative comparison shows the magnitude in difference between the estimated and the actual observation. Particularly this disagreement can be attributed to the needle closing rate within the simulation, that affects the traveling rarefaction pressure wave in the nozzle. The lowering of the sac pressure below the critical pressure is the primary driver for the bulk cavitation phenomena.

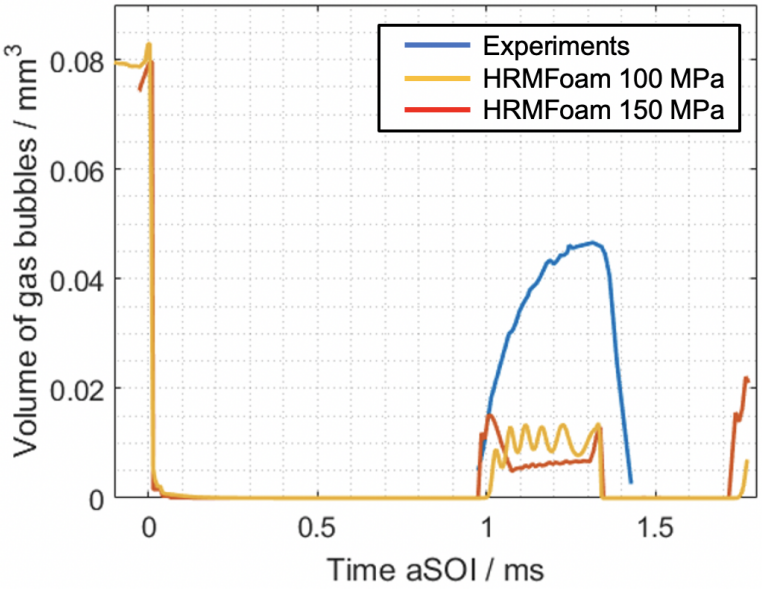

Figure 4.28. Quantitative comparisons show HRMFoam simulations underpredict cavitation in the sac/nozzle region compared to the experiments (injection pressure, 100MPa). Experiments were conducted by Lyle Pickett.

In Figure 4.29, the effect of the injection and ambient pressure on the presence of gas at the end of injection is shown. It is noted that the ambient conditions affects the gas ingestion more than the injection pressure. This result makes sense since with a lower ambient pressure, at needle closure the pressure drop would be higher and therefore cause larger amounts of bulk cavitation.

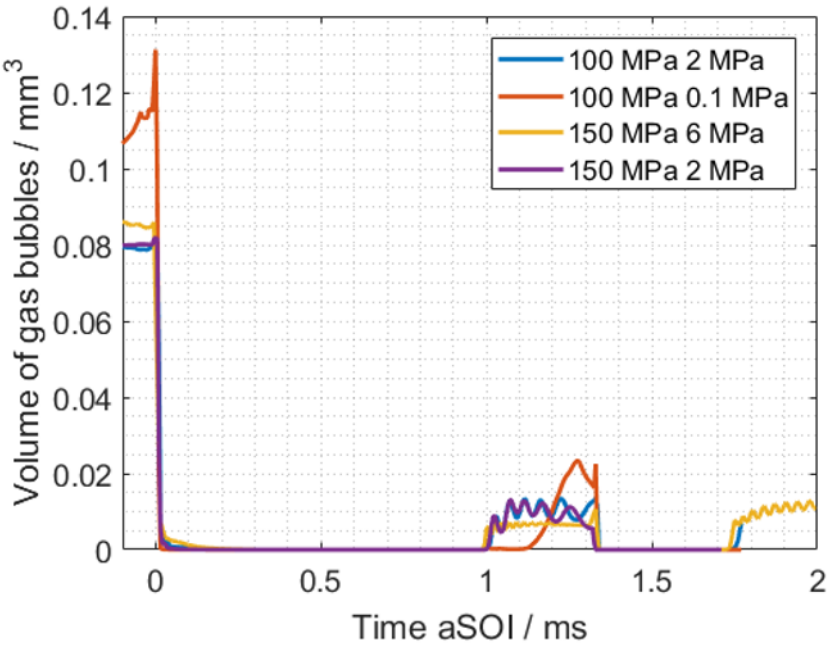

Figure 4.29. Prediction of vapor in the sac/nozzle show ambient pressure has a higher affect than injection pressure. Experiments were conducted by Lyle Pickett.

The presence of unburnt fuel at the nozzle exit after the end of injection is one of the major causes of soot formation in an engine. Therefore during the dwell phase between injections, it is important to understand the factors controlling the phenomena of so called fuel residual. In Figure 4.30 we show the region of investigation for fuel residual calculations. Figure 4.31 shows the fuel residual measurements from HRMFoam simulation. The amount of fuel is independent of injection pressure, as is expected. For lower ambient pressures, the fuel residual is significantly larger than the higher ambient conditions due to larger pressure difference.

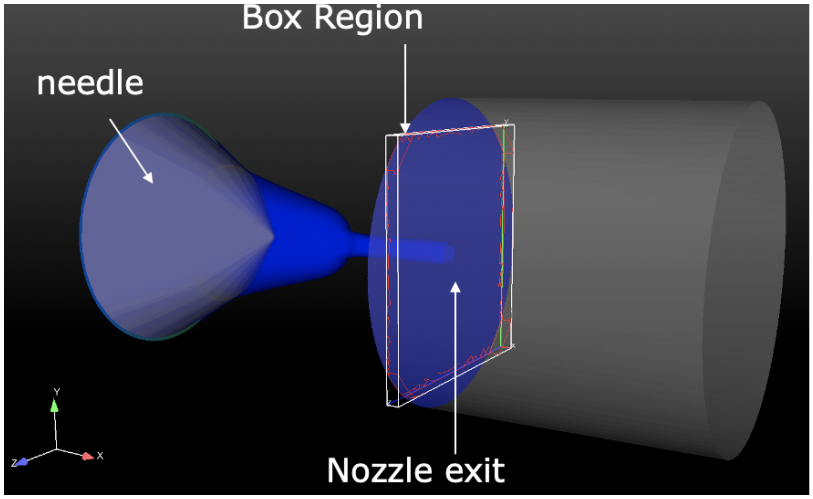

Figure 4.30. Schematic of the fuel residual measurement. The box region, close to the nozzle exit is the region the measurements are made.

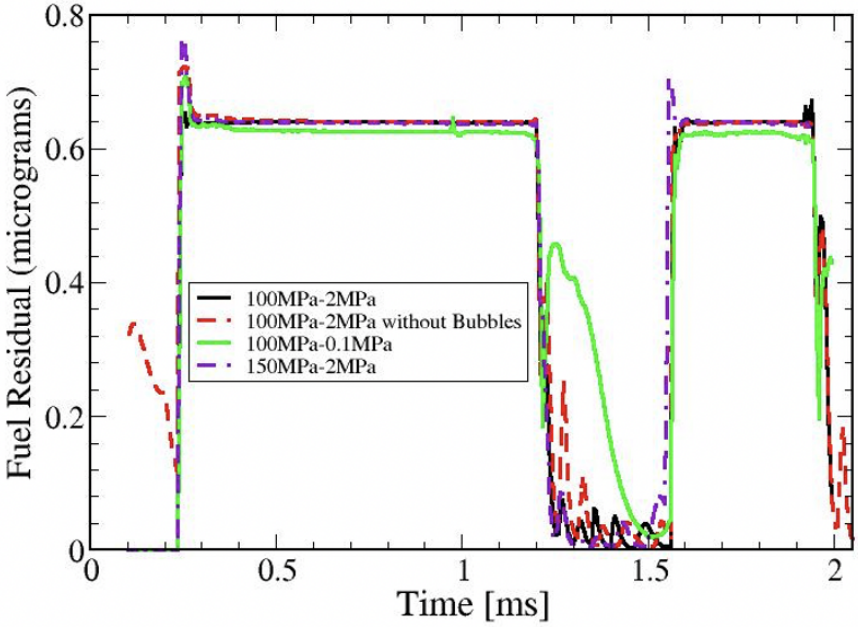

Figure 4.31. The unburnt fuel residual is impacted by the ambient conditions more than the injection pressure, and therefore is an important discovery for engine designers.

4.6 Inference

In summary, the numerical modeling of the multi-injection event involved overcoming various challenges around the needle closure event using adaptive meshing strategies for complicated geometries as well as phase-change thermodynamics. In the end, we draw the following conclusions in the study of a converging, symmetrical nozzle.

- Though conventionally believed that converging nozzles do not induce cavitation, in the experiments [141] (metal target, transparent nozzle) and simulations showed presence of cavitation near the nozzle exit. This source of cavitation does not come from the upstream of the nozzle but is rather local due to the slight divergence near the exit and the small features that were obtained from the high- fidelity experimental characterizations. Therefore, resolving local features is important to fully characterize the internal flow within the nozzle.

- The effect of needle lift on the flow characteristics is also studied using CFD and new discoveries made including the changing internal flow behavior during the quasi- steady state and the presence of an annular jet causing loss of pressure head and momentum towards the beginning of the end of injection. The location of the needle also does not seem to affect the flow characteristics like rate of injection if the needle opening area is greater than the nozzle flow area but influences the flow momentum and flow coefficients as well as the presence of vorticity in the nozzle, when the lift is low during quasi-steady state.

- With gas initialized in the sac and holes (upto 12% and 100%), computations show that gas is likely compressed/ dissolved in liquid in about a few μs, depending on the amount of gas in the sac, and the rate of injection is affected. Thus the gas has a an effect on the delay in the injection and initial spray growth.

- The comparisons between the CFD and the educated injection rate shape model [161] is good for the quasi- steady state, although it seems that the rate shape model does not include effects of the initial sac conditions, which would modify the ROI curve depending on the initial state of the sac and hole. Additionally, it appears that the ramp-up/ramp-down in CFD models are faster compared to the experimental model fit, and this discrepancy needs to be addressed.

- In the near nozzle, the performance of HRMFoam is comparable to the high fidelity models and experimental radiography measurements. While the HRM-Foam is unable to predict the asymmetry in the spray profile, the spray width is accurately modeled.

- Comparing HRMFoam to high fidelity DNS-CLSVOF simulation, yields that the Σ predictions from HRM are comparable (at least qualitatively) to the DNS predictions.

- The End of Injection dynamics show bulk cavitation in both the model predictions and experimental observations, and the mechanisms for them have been identified. However, the HRM model predictions are lower than the experiments which suggests further modeling improvements are needed.

CHAPTER 5

THE EULERIAN LAGRANGIAN MIXING ORIENTED (ELMO) MODEL

5.1 Introduction

Spray atomization models has a leading order effect on the prediction error for transient fuel combustion. In order for a numerical model to be truly predictive for combustion modeling, the atomization model has to be able to predict the correct quantity of fuel at the right time (in terms of the combustion cycle) and the appropriate location downstream of the spray injector. Current approaches in Lagrangian-Eulerian modeling, where the fluid is treated as Eulerian in the dense region and switches to the Lagrangian formulation in the dilute region depend heavily on sub-models for primary and secondary atomization. These include the Kelvin-Helmholtz Rayleigh Taylor (KH-RT) models among others. In a typical Lagrangian blob model Kelvin-Helmholtz breakup, there are many adjustable constants that need to be tuned and has shown to have an effect on model predictions. For example, B_0, B_1, C_t and C_{RT} are the KH droplet size constant, KH breakup time constant, RT droplet size and RT breakup time constant respectively. Alongside this, recent insights by experiments conducted by Siebers et al. [205, 151, 204] contradict the basis of past spray models such as KH-RT. Siebers [205, 151, 204] concluded '*the processes of atomization and the ensuing interphase transport of mass and energy at droplet surfaces are not limiting steps with respect to fuel vaporization in Direct Injection (DI) diesel sprays but rather they are limited by the turbulent mixing between phases*'.

Siebers [205] proposed a stand-alone steady jet one-dimensional simplified model for predicting liquid length for non-vaporizing conditions for such high We sprays.

Other studies, including Pastor et al. [159, 160], Desantes et al. [49, 48] and Musculus et al. [150] extended the original formulation using the Eulerian control volume approach to add physics of vaporization, model reacting flows, and the end of injection transients, respectively with great success. These models assumed a radial profile of the spray contrary to the assumptions in the original work [49, 159]. This is consistent from experimental and high-fidelity simulation observations for sprays. The typical profile for a spray during the course of its evolution is shown in Figure 5.1. The spray starts off as a slug flow near the nozzle exit, and as it progresses downstream of the nozzle exit, acquires a radial profile due to aerodynamic interactions. Here α is the shape factor defined in [150].

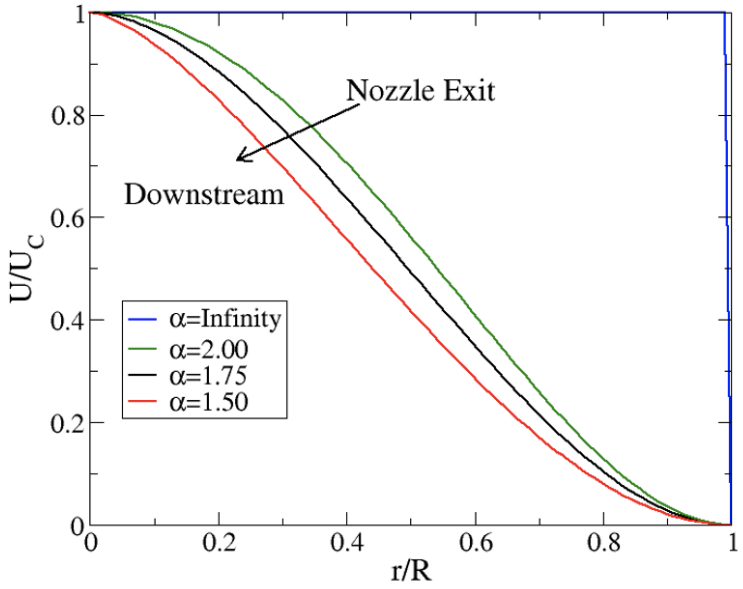

Figure 5.1. The spray profile behavior from nozzle exit to downstream. Image credit L. Pickett (Sandia National Laboratories)

However, these models were stipulated to be used only as stand-alone models and not be coupled with a multi-dimensional CFD solver. The current approach in the form of the Eulerian Lagrangian Mixing Oriented (ELMO) model, aims to

bridge that gap. The ELMO model inputs are informed by the internal flow and experimental observations including discharge rate, and spray angle. The near-nozzle spray evolution is based on the more accurate mixing-limited hypothesis inspired by previous efforts in this area [205, 49, 159, 150]. The ELMO model is coupled to a CFD solver and in the dilute region downstream of the nozzle, this model transitions to standard Lagrangian Eulerian and its droplet centric formulation. A schematic is shown in Figure 5.2.

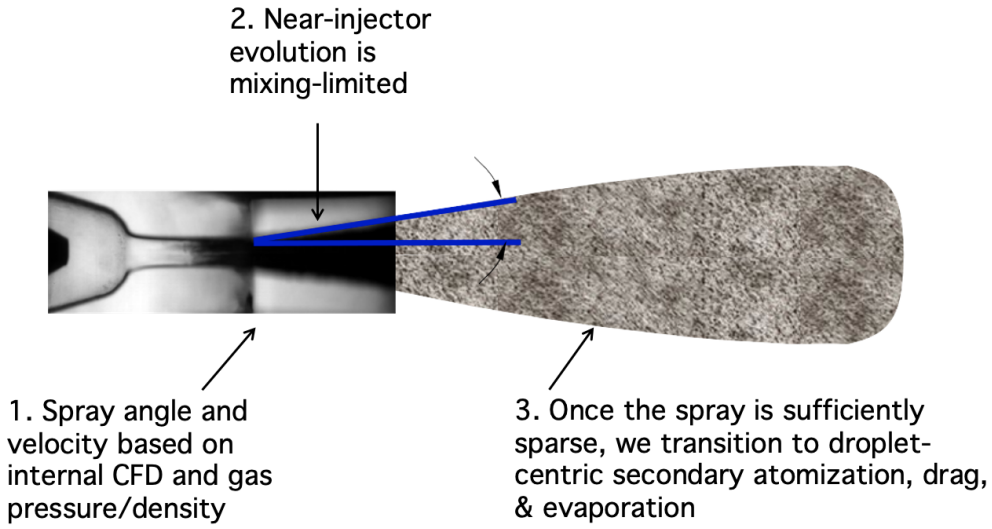

Figure 5.2. The ELMO model ideation is shown in this figure. In the near nozzle region, the ELMO operates under the mixing limited assumption. When the spray becomes sufficiently dilute, the ELMO transitions to standard LE and its droplet centric approach. Image credit D.P. Schmidt (UMass)

The advantages of this approach includes:

- The model can benefit from sophisticated transient internal models, informing the inputs for the external spray modeling paradigm.
- The near nozzle behavior, is consistent with the modern mixing limited hypothesis, therefore expected to be more accurate.

- In the downstream region, the model employs conventional droplet models (collisions, drag) where they work best [156].

- Since spray angle is an input to the model, it can be more effectively modeled than in the current state-of-the-art. Siebers [205] shows the spray angle to be transient during high/low needle lift. Most of the current modeling approaches ignore this transient, yet important, behavior.

5.1.1 Discussions about the ELMO model

The Musculus-Kattke model [150] is used as the starting point for the numerical formulation for the ELMO model. While the previous implementation is in an Eulerian frame of reference, the ELMO model is in its corresponding Lagrangian frame of reference (numerics introduced in the next section). This essentially provides the capability to use the Lagrangian 'capsules' as self contained liquid fuel parcels that entrain ambient air during the translation process along the spray axis. The capsule construct is shown in Figure 5.3 and involves a moving Lagrangian parcel over an underlying Eulerian mesh. This capsule therefore has two phases, the liquid fuel and the gas, including ambient air as well as vaporized fuel due to vaporization caused by the entrainment of hot ambient air. The in-capsule vaporization is modeled on the Desantes [49] formulation based on equilibrium evaporative physics. Therefore, the ELMO model is built on assumptions of inertial and thermal equilibrium, limited only by the entrainment rate of surrounding ambient air.

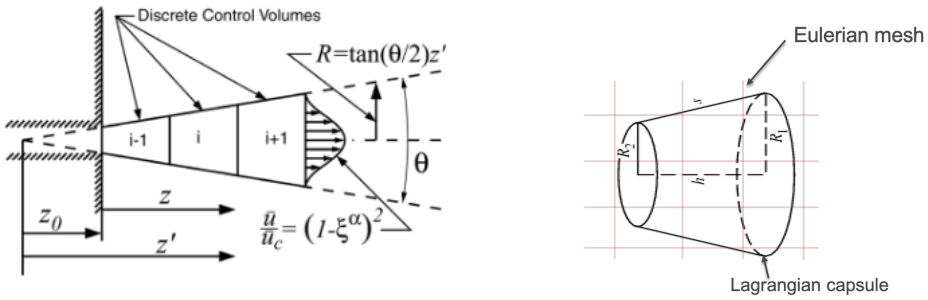

Figure 5.3. [Left]: The Musculus-Kattke model [150] on the left is implemented with the Eulerian control volume concept, whereas the ELMO model on the right panel is based on its equivalent Lagrangian formulation

The primary inputs for the ELMO model are the discharge rate and the spreading angle, both obtained from experiments. From Siebers [205], we know the spreading angle dictates the amount of entrainment and therefore vaporization.

5.2 The ELMO model numerics

In this section we will introduce the detailed discussion about the numerics for the formulation of the ELMO model, including some of its assumptions.

5.2.1 The capsule formulation

The Eulerian formulation in Musculus Kattke [150] while suitable for a standalone spray prediction model, is converted into its corresponding Lagrangian formulation. This is done in order to use the superior capabilities of the underlying CFD code CONVERGE [198, 5] in handling Lagrangian parcels. The ELMO model has a few assumptions:

- The mass of fuel in a capsule remains constant throughout the evolution of the capsule, upto the point of transition. It can be represented mathematically as:

$$\frac{dm_f}{dt} = 0 \tag{5.1}$$

where m_f is the fuel mass

- The momentum of the capsule remains constant. This translates to the capsule velocity slowing down in proportion to the mass of the ambient air entrained. Mathematically, at an individual capsule level this can be represented as:

$$\frac{dM}{dt} = 0 \tag{5.2}$$

where M is the momentum

- Each capsule starts off with its spray angle at birth. This spray angle dictates the shape, entrainment within the capsule, and remains constant throughout its life.

- Underlying all of the equations (introduced later in this section) are assumptions of equilibrium. Specifically, the phases all move at the same velocity and experience equal pressure. All phases are in thermal equilibrium as well. The capsule is in thermodynamic equilibrium, with an assumption of locally homogeneous flow. Therefore liquid and gas are at the same temperature.

Applying the Reynolds Transport Theorem (RTT), we convert the Eulerian MK control volumes [150] into the Lagrangian ELMO capsules.

$$\frac{dm_f}{dt} = \frac{\partial}{\partial t}\int_V \rho_f \bar{X}_f dV + \int_{CS} \rho_f \bar{X}_f (\vec{u} - \vec{u_r}) dA \tag{5.3}$$

by design, the left hand side of the equation is set to zero. The volumetric integral time is also zero due to fuel mass conservation principles, which reduces the above equation to

$$\int_{FS} \bar{X}_f \vec{u}.dA - \int_{FS} \bar{X}_f \vec{u_r} \cdot dA = 0 \tag{5.4}$$

which further can be reduced to:

$$\vec{u_r} = \beta\vec{u} \tag{5.5}$$

where m_f is the mass of fuel, u is the axial velocity, u_r is the control surface velocity, β is the shape factor of fuel volume fraction, ρ_f is the density of fuel, $\bar{X}_f$ is the fuel volume fraction, CS is the control surface and FS is the front surface. The term β accounts for the shape of the fuel volume fraction and velocity profiles, ranging from $\beta = 1$ for a uniform profile to approximately $\beta = 2$ for fully developed jets and is based on the representation shown in Figure 5.1 for α. For details about the shape factor term, please refer the original source of this work [150].

Similarly applying RTT, the momentum formulation can be represented as:

$$\frac{dM}{dt} = \frac{\partial}{\partial t}\int_V \bar{\bar{\rho}}\vec{u}dV + \int_{CS} \bar{\bar{\rho}}\bar{\bar{u}}(\vec{u} - \vec{u_r}.n)dA \tag{5.6}$$

per the design of the capsule and in absence of pressure gradient, the equation above reduces (approximately) to

$$\frac{\partial}{\partial t}\bar{\bar{u}}m = 0 \tag{5.7}$$

where M is the momentum of the capsule, $\bar{\bar{\rho}}$ is the mixture density of the capsule, $\vec{u}$ is the axial velocity, $\bar{\bar{u}}$ is the cross sectionally averaged axial velocity, u_r is the control surface velocity and m is the total mass of capsule.

Since the fuel mass in a capsule remains constant it can either exist in a liquid or vapor (gas) phase. The growth in volume (and mass) of the capsule at successive time instants is accounted by the entrainment of the ambient air (schematic shown in Figure 5.4). Therefore the mass and volume conservation equations have the following functional form:

$$m = m_l + m_v + m_a \tag{5.8}$$

$$V = V_l + V_v + V_a \tag{5.9}$$

where the subscripts l, v and a are liquid fuel, vapor and non-condensible gas respectively. The presence of the non-condensible gas in the capsule necessitates calculating phase interactions with the underlying Eulerian CFD control volume. This approach will be detailed in the coming sections.

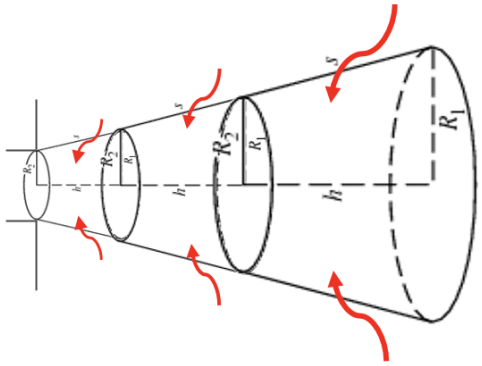

Figure 5.4. The schematic of the ambient air entrainment in the capsule

The increase in capsule mass due to the entrainment of air, Δm_{air}, can be represented as

$$m_{capsule}^{t+1} = m_{capsule}^{t} + \Delta m_{air} \tag{5.10}$$

In the present implementation, the effect due to entrained momentum is ignored, similar to the original work [150]. This therefore leads to a constant momentum for capsule, with decelerating capsule velocity. Like the implementation of Musculus and Kattke, the consideration of a radial velocity and density profile is included here, which has implications for momentum. This radial profile is manifested by the term

β such that the velocity is higher along the spray axis, where most of the mass is concentrated. As the capsule advances in time, the position is updated by a simple explicit Euler time integration given by:

$$z^{t+1} = z^t + \beta \bar{u} \Delta t \tag{5.11}$$

where z is the axial location of the capsule, and $\bar{u}$ is the capsule velocity at the location z. However it is to be noted that in the current implementation, the β value is set to 1. While the capsule volume increases radially, its axial extent diminishes as the control surfaces of the capsules are advected at different velocities based on their location along the spray axis. This same reduction in velocity represents a deceleration of the center of mass of each capsule. The derivation of the front side (fs) and back side (bs) location of each capsule begins with the assumptions that these capsule boundaries lie halfway between the centroid z of a capsule at its current position and next position for the front side or halfway between the current centroid location and previous location for the back side. Expressing this mathematically:

$$fs^t = \frac{z^t + z^{t+1}}{2} \tag{5.12}$$

$$bs^t = \frac{z^t + z^{t-1}}{2} \tag{5.13}$$

In order to make the train of capsules continuous, we constrain the back surface of the leading capsule to match with the front surface of the trailing capsule.

Once the location of the front and back sides are updated, the geometry of the capsule is fully determined and the kinematics of the advection process is complete. Based on the original model [150], as the axial extent of the capsules are known, the other features such as velocity, volume can be easily calculated.

5.2.2 Thermodynamic formulation

To model the process of vaporization, we follow the analysis first proposed by Desantes et al.[49]. Desantes et al. extended a one-dimensional nonevaporative spray model to take into account the effects of evaporation under direct-injection diesel engine conditions. Their approach is based on the assumption that evaporation is limited by the mixing process between fuel and ambient gas, and makes use of non-ideal gas state relationships for the description of the whole spray. Their model has been shown to accurately predict the influence of fuel type and both ambient and injection conditions on liquid spray penetration. The hot entrained ambient air is the primary driver of the vaporization process. The local thermodynamic equilibrium in the capsule depend on the mass fractions. Mathematically, they are represented as below.

$$Y_f = Y_{f,l} + Y_{f,v} \tag{5.14}$$

$$Y_a = 1 - Y_f \tag{5.15}$$

where suffixes f, a, l, v correspond to fuel, air, liquid and vapor phase respectively. While in the original work, non-ideal state relationships were considered, in this implementation we assume all gases to be ideal. This greatly simplifies our calculations without appreciable differences between predictions (as we shall see later).

The local mixture density is calculated using an ideal mixture assumption

$$\rho = \frac{1}{\frac{Y_{f,l}}{\rho_{f,l}} + \frac{Y_{f,v}}{\rho_{f,v}} + \frac{Y_a}{\rho_a}} \tag{5.16}$$

where Y_i is the mass fraction of the mixture i component and ρ_i is the density for the pure component i at the mixture temperature T and total pressure P.

The local mixture enthalpy h is also calculated through an ideal mixture assumption as

$$h(T) = Y_{f,l}.h_{f,l}(T) + Y_{f,v}.h_{f,v}(T) + Y_a.h_a(T) \tag{5.17}$$

where Y_i is the mass fraction of the mixture i component and h_i is the enthalpy for the pure component i at the mixture temperature T.

Due to the liquid-vapour equilibrium, the composition and temperatures of phases are coupled i.e., for a given Y_f, the proportion of liquid or evaporated fuel is not known, nor is the mixture temperature. It can however be solved iteratively, using the formulation below

$$\frac{Y_{f,v}}{1 - Y_f} = \frac{h_a(T_{a,\text{inf}} - h_a(T) - \frac{Y_f}{1-Y_f}[h_{f,l}(T) - h_{f,l}(T_{f,0})]}{\Delta h_v(T)} \tag{5.18}$$

In this equation, at any given instant if Y_f is known the right hand side terms depend only on local temperature (T), and the left hand side on $Y_{f,v}$. This equation fully determines the thermodynamic equilibrium state of the capsule. The equilibrium equation can be written in terms of mass fractions and molecular weights as

$$\frac{Y_{f,v}}{1 - Y_f} = \frac{MW_f}{MW_a} \frac{1}{\frac{p_a}{p_v} - 1} \tag{5.19}$$

where MW_f is the molecular weight of the fuel, and MW_a is the molecular weight of the ambient air. This helps to estimate the $Y_{f,v}$ term and therefore solve 5.18 using a root finding method to determine the temperature.

The ideal gas assumption necessitates using corrections (Poynting) in calculating vapor pressures [49].

$$p_v = p_v^0 * exp[V_l \frac{p_a - p_v^0}{R_{gas}T}] \tag{5.20}$$

where the uncorrected vapor pressure is p_v^0, R_{gas} is the gas constant, p_a is the ambient pressure. This corrected vapor pressure is used in 5.19. A two-way coupling mechanism based on Monte Carlo (MC) integration (introduced in the next section)

is used to sample thermodynamic variables within the capsule geometry. This MC mechanism is also used to provide gas phase source term, as a result of vaporized fuel, to the Eulerian CFD mesh. A key feature of this implementation is the presence of gas in both the capsule (to account for effects of vaporization, and mass/momentum) and the Eulerian mesh. To simplify our calculations, we define the gas phase gain equal to the loss of the liquid fuel. This evaporated fuel now in the gas phase is used in the source term for the mass. For example, consider the mass source due to vaporization, where the evaporated vapor mass during one time step is represented Δm_{evap}. Here, Δm_{evap} is positive for mass leaving the liquid

$$S_m = -(m_l^{n+1} - m_l^n) = \Delta m_{evap} \tag{5.21}$$

Similarly the source term for the momentum gained by the gas phase is equal to the momentum lost by the liquid phase

$$S_M = m_l^n \bar{u}^n - m_l^{n+1} \bar{u}^{n+1} \tag{5.22}$$

In the current implementation the momentum source terms are only the axial components. While radial contributions might be relevant, it is not included in the current modeling paradigm. This can be extended by using the β shape factor term introduced earlier. However this will be a subject of future development.

The energy source term can be expressed as:

$$S_E = m_l^n \bar{h}^n - m_l^{n+1} \bar{h}^{n+1} \tag{5.23}$$

5.2.3 Coupling formulation

The source terms introduced in the previous section must be distributed to the underlying Eulerian mesh. Depending on the capsule spatial and radial extent, and the underlying mesh resolution each capsule may overlap with numerous gas phase

cells. A Monte Carlo (MC) integration method is used to distribute the source terms to each gas phase cells. This coupling using MC is a two-way coupling, as the gas phase properties are obtained for the capsule advection equations. This distribution process has two major steps, the creation of Monte-Carlo points and the assignment of weights for source terms.

5.2.3.1 Creating of Monte Carlo (MC) points

The creation of MC points is a further two step process. In the first step, points are randomly generated with an uniform distribution within a cylinder enclosing the axial extents of the cone frustum. The radii of the cylinder is equal to the radii of the front side of the capsule (the larger radii). In the next step, an acceptance/rejection algorithm is used to determine points falling within the capsule cone frustum geometry. This slightly complicated approach is necessary since it is mathematically challenging to uniformly distribute points at random in a conical frustum. Only points falling inside the frustum are accepted and used for the coupling.

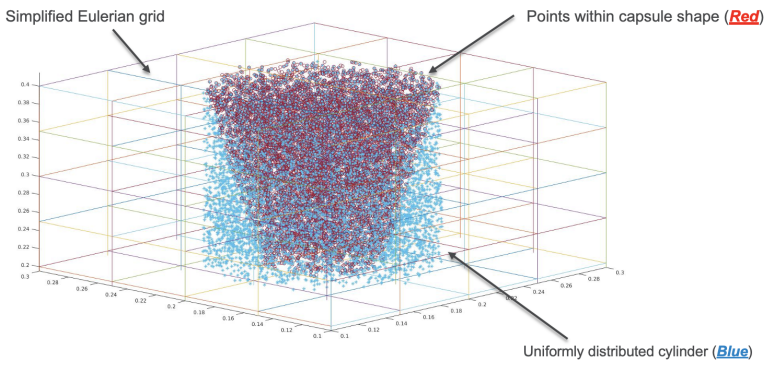

Figure 5.5. The acceptance/rejection algorithm shows the MC points creation

5.2.3.2 Redistribution of source terms

A single capsule may overlap with multiple mesh cells. This complicates the redistribution of source terms to the underlying mesh effectively. A simple strategy based on the same point finding algorithm used to determine in which cell a Lagrangian

parcel might reside is used. In this algorithm for a given capsule, once the points are created for each of these Monte Carlo points, a fraction of 1/N of the source term is deposited into the cell which contains the point. For example, in the Figure 5.6, cell # 8 has two randomly generated MC points coinciding with it. Therefore 2/N of the total mass/momentum/energy in the capsule is added to the gas phase source term. This stochastic approach obviates the need for calculating expensive geometric intersections. These interior points are used only for distributing source terms.

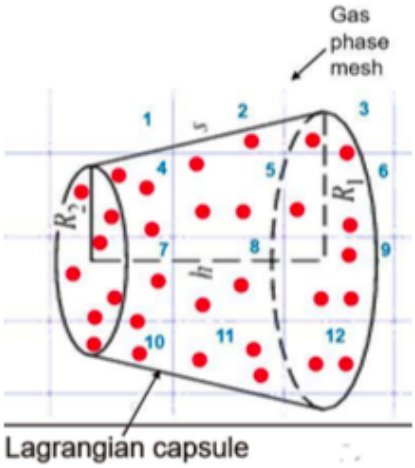

Figure 5.6. The schematic of a limited number of points distributed randomly within the capsule geometry, post acceptance/rejection step. The underlying Eulerian mesh computes the overlapping number of points within each cell and distributes the source terms.

5.2.3.3 Transition

As the capsule traverses along the axial direction, it ingests more ambient gas in the form of entrainment. This causes vaporization as the ambient gas temperatures are on average higher than the fuel temperature. This combined effect reduces the liquid fuel in the capsule, making the capsules diluted at a certain location downstream. At this point the Lagrangian capsule approach is not further needed, and the capsule is ready to be transitioned in Lagrangian parcels consisting of very small

droplets. This transition has several advantages including making use of existing frameworks to capture parcel impacts on cylinder walls. In the current formulation, the capsules only traverse in the direction of the initial spray, post transition parcels are not limited by this over-simplification. This will help in the model's ability to successfully capture the phenomena of spray bending, as is observed in the case of multi-injector nozzles.

The criterion for transition in this current implementation is based on two factors:

- The liquid volume fraction is below 0.005
- The vapor volume fraction of the total fuel volume fraction is above 99%

These thresholds are somewhat arbitrary and are a tunable parameter in the model. The 99% threshold means the liquid phase has all but converted into the vapor phase and that the spray is dilute enough to be converted into smaller droplets (or parcels). The LVF criterion of 0.005 correspond to six diameter distance between consecutive droplets arranged in a uniform lattice structure. Quan et al. [174] showed the impact of inter droplet spacing especially on the trailing droplet including its drag coefficient.

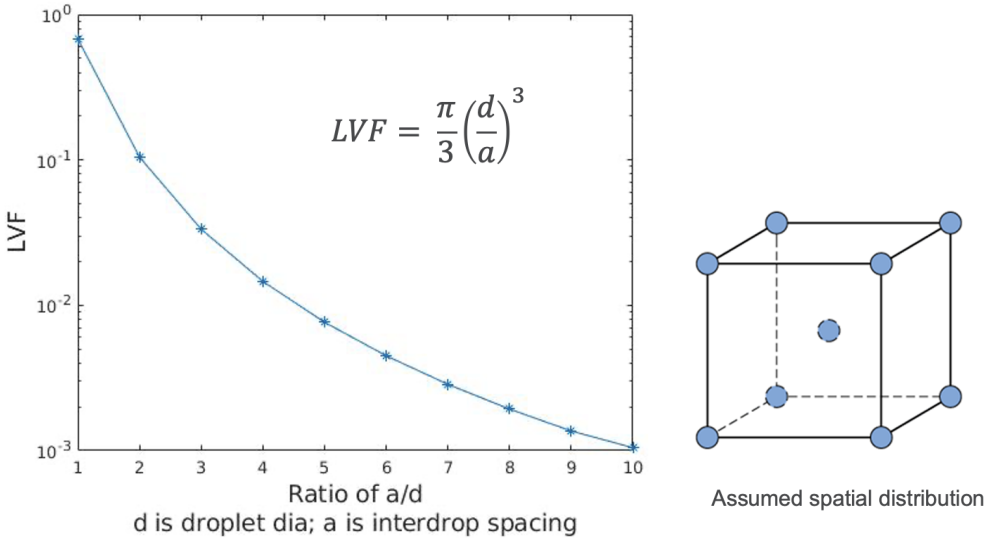

Figure 5.7. On the left shows the plot of LVF against ratio of droplet dia and interdroplet spacing based on an assumed spatial distribution

During transition, the remaining liquid mass in the capsule is proposed to be divided equally into a predetermined number of parcels at locations generated randomly within the capsule's geometrical extents. Since the capsule geometry is small, of the order of a few Eulerian mesh cells, the random distribution of parcels is not expected to make a huge difference. The velocity imparted to the newly generated parcels is randomly selected and the total magnitude is equal to that of the original capsule. The droplet size is based on a critical Weber number of 6. This is informed from experiments [168]. Once the transition to Lagrangian parcels is complete the parcels evolve with typical spray models appropriate for secondary atomization of dilute sprays.

5.2.3.4 Overall approach

The kinematic equations are solved first using the same time-step (δt) as the CFD solver. The discharge rate curve is interpolated based on the overall timestep of the solution. The discharge rate determines the mass of fuel in each capsule upon

initialization. The capsules are initialized at the nozzle exit, with an initial spray angle, and an axial length equal to the velocity (at birth) times the δt. The capsules are then updated using equations 5.11 - 5.13. The new positions determine the volume of the capsule as the one dimensional implementation is projected onto three dimensions using simplified assumptions of radial profiles, shown in (left panel) Figure 5.3. The entrained ambient air mass is calculated and the new velocity determined, by conserving the capsule momentum. The solution of the thermodynamic state using the relationships discussed between 5.14 - 5.20, requires iterative solution due to the coupling of the variables. Once the new thermodynamic state is found, the capsule calculations are concluded. At the end of each cycle of calculations, capsules are examined against the transition criterion and the ones meeting it are transitioned into smaller Lagrangian parcels.

5.3 Results

In this section, we will first present the results from the validation studies. Following which, we shall present performance of the ELMO model against experimental data and comparisons against the current state-of-the-art models such as Lagrangian Eulerian. It is to be noted that the ELMO model is implemented within the CONVERGE CFD solver and used in place of the primary atomization KH-RT model. In this section, results for validation against spray statistics measured for Spray A, Spray H, and Spray G will be presented. Lastly, we shall compare effect of some of the hyperparameters such as number of MC points per capsule and sensitivity to turbulence models in the last section.

5.3.1 Validation Studies

To validate the Lagrangian implementation of ELMO, we use the test case used originally in the Musculus Kattke paper [150]. In this test case, a square-wave injec-

T_a	900 K
ρ_f	670 kg/m3
P_{rail}	1200 bar
ρ_a	19.3 kg/m3
C_a	0.83
C_d	0.58
θ	17
d_0	0.2 mm

Table 5.1. Conditions for MK/ELMO verification. Source [150]

tion pulse with duration of 0.5 ms is simulated using both the 1-d transient model. The simulation conditions are listed in Table 5.1. The ambient gas conditions are similar to the Low Temperature Combustion (LTC) condition, with Exhaust Gas Recirculation (EGR)-diluted ambient gases at 900 K with a density of 19.3 kg/m3 and 12.7% O2. The orifice discharge and area-contraction coefficients (C_d and C_a respectively) and spreading angle (θ) were not measured, but are assumed using typical values from [151]. The spatial resolution in the axial direction is 0.1 mm, and the domain is 100 mm long.

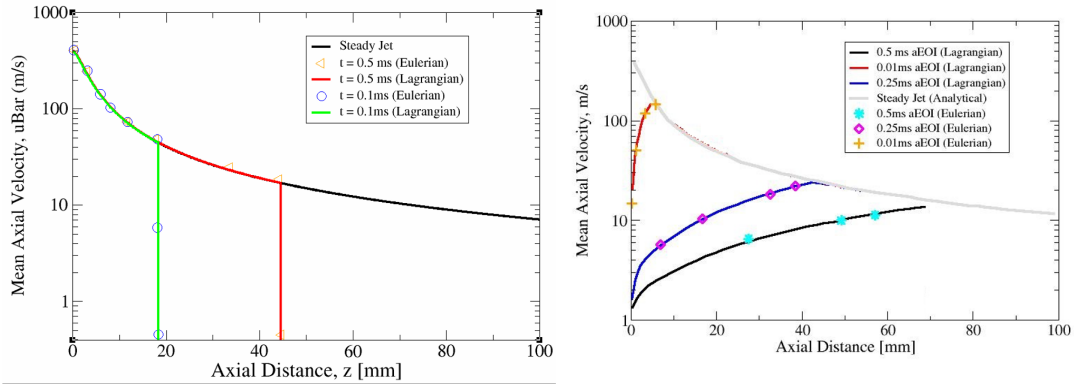

Figure 5.8. Cross-sectionally averaged velocity predictions of the one-dimensional jet model for the Low Temperature Conditions (LTC) condition. The plots on the left and right show the predictions before and after EOI at 0.5 ms, respectively. Each curve is labeled according to its time in the simulation, relative to the start of injection.

At each instant of the simulation shown in Figure 5.8 the velocity $\bar{\bar{u}}$ varies approximately inversely with the axial distance , z as is typical of steady jets. The sharp drop in $\bar{\bar{u}}$ at each instant in the simulation (the vertical line in the left panel of Figure 5.8 indicates the head of the jet, whose penetration in the numerical solution agrees with the analytical solution (black curve) of the Naber model [151]. Therefore, up to End of Injection (EOI), the numerical solution agrees with the analytical solution for steady jets, as it should. The prediction of $\bar{\bar{u}}$ is shown in the right panel. Again, the predictions at various instants in the simulation are indicated by the annotations on the plots, and the gray line represents the steady-jet behavior. Some aspects of the predictions after EOI are noteworthy.

Firstly, the model predictions show that after EOI, $\bar{\bar{u}}$ decrease very rapidly to zero near the injector. As the simulation proceeds in time, the values of $\bar{\bar{u}}$ remain anchored at zero near the injector. Farther downstream, the $\bar{\bar{u}}$ increases nearly linearly to the downstream steady jet curve (gray). The linear dependence is not revealed in the log scale plot. Secondly, the departure from steady-jet behavior moves downstream at roughly twice the rate of the initial jet penetration. In the early part of the simulation after EOI, the $\bar{\bar{u}}$ curves join the steady-jet behavior at some distance downstream. The downstream distance where $\bar{\bar{u}}$ and join the steady jet behavior moves at twice the initial jet penetration rate shown in the left plot. Therefore, the EOI disturbance (departure of $\bar{\bar{u}}$ from steady-jet behavior) reaches a given downstream position in roughly half the time required for the initial penetration of the head of the jet.

The ELMO (Lagrangian implementation) and the Musculus-Kattke [150] predictions are almost identical leading to confidence in the model formulation. It is to be noted that the ELMO results do not use vaporization formulation for this study in order to be consistent with the non-evaporative Musculus-Kattke model.

In order to further validate the steady jet behavior (based on [151]), the ELMO results are plotted against the Musculus Kattke model predictions in Figure 5.9. The

predictions are comparable building confidence in the Lagrangian implementation of the model. The conditions for this study is detailed in Table 5.1.

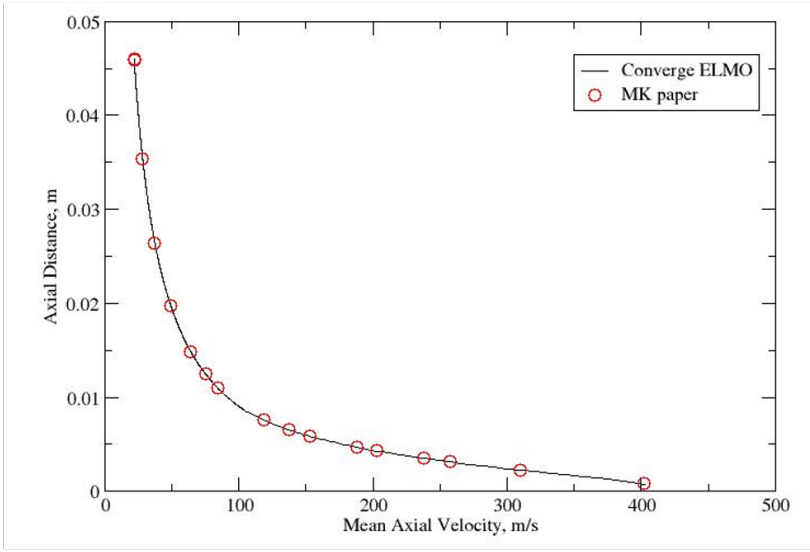

Figure 5.9. Verification of steady state behavior for the ELMO model

Using Eqs 5.14 - 5.20, the state relationships for the whole mass fraction range from 1 (pure fuel) to 0 (ambient gas) can be obtained. In order to demonstrate the implementation in the current study we compare the ELMO predictions with the data obtained from the Desantes work [49]. Figure 5.10 shows an example of such solutions for typical diesel conditions employing decane as fuel. Figure 5.10 shows that the main evaporation process takes place between $Y_f = 0.75$ and $Y_{f,evap} = 0.37$. The latter point is also a transition for the temperature evolution. The ELMO predictions are a close match (here in red) and the slight differences can be attributed to the ideal gas assumption instead of the real gas treatment used in the original study [49]. The ideal gas assumption greatly simplified the numerics of the computation.

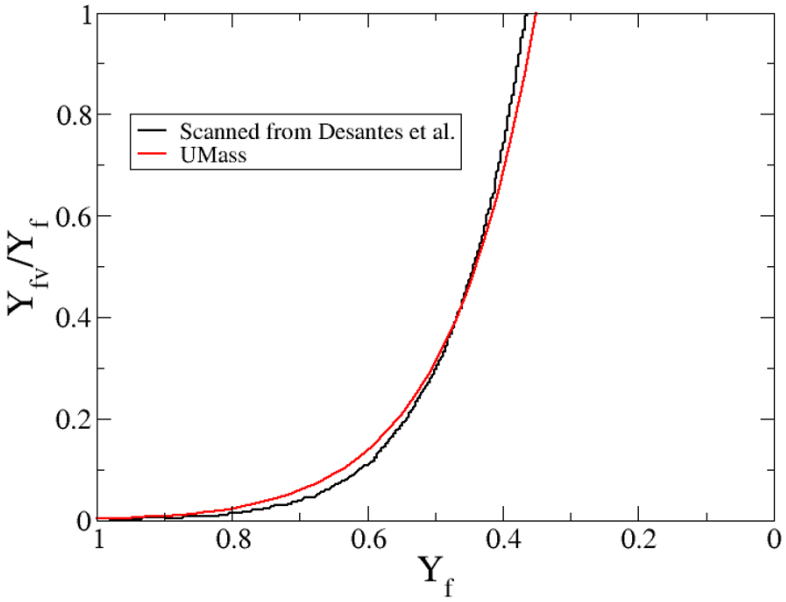

Figure 5.10. State relationships example results from [49]. Fuel: decane, P = 7 MPa, T_a = 885 K, $T_f, 0$ = 373 K. Left axis is vaporized fuel mass fraction

To further validate the ELMO temperature predictions, critical for modeling vaporization, we make comparisons to the adiabatic mixing curve. Here the Lagrangian Eulerian (LE) implementation in CONVERGE is also shown for comparisons. The conditions for this study are similar to the Spray A conditions discussed in Table 5.2. The data were obtained at 18 mm from the nozzle exit, that corresponds to the lift off length for Spray A. The ELMO transition occurs at around 4-5 mm from the tip. This comparison is important as it can portray the effect of the locally homogeneous flow assumptions within the capsules and the mixing limited physics that drives ELMO vaporization predictions. Here we see the ELMO model predictions follow the blue mixing curve up until 0.14 vapor fuel mass fraction and predict lower temperatures thereafter. More effect of this behavior will be evident in the reacting simulations. It is to be noted that the LE predictions are consistently almost always higher than the adiabatic mixing plot and predicts higher temperatures.

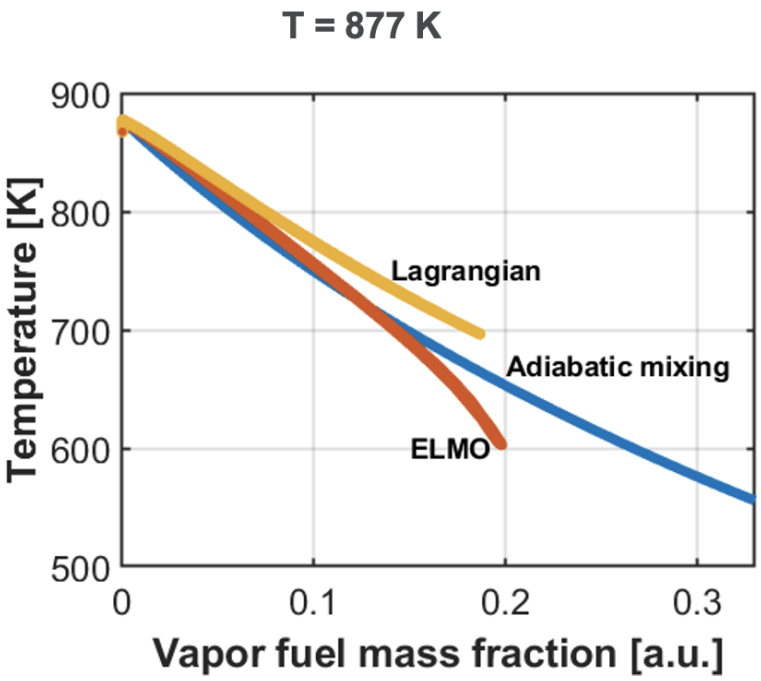

Figure 5.11. Validation of ELMO vaporization model against adiabatic mixing limits

5.3.2 Spray Studies

In this section, we will study three different injectors that operate in vastly different regimes. The objective is to study the performance of the ELMO model under different conditions. While we shall briefly summarize the hydraulic and thermodynamic conditions in this section, more details on the injectors and their geometries can be found on the ECN webpage [1].

5.3.2.1 Spray A

Spray A is a single-hole diesel fuel injector. The base simulation case is similar to the Spray A example case within CONVERGE [198, 5]. More hydraulic and thermodynamic details are presented in Table 5.2.

Number of Injectors	1
Injection Type	Single
Injection Duration	1.54 ms
Gas Temperature	877 K
Injection Temperature	373 K
Nozzle diameter	90 μm
Turbulence model	k-epsilon RNG
Spray Angle	21 deg
C_a	0.98

Table 5.2. Spray A test conditions

The vapor penetration comparison plot is shown in Figure 5.12. The experimental studies from Sandia [164] and CMT [162] show the effect of variability in conditions and measurements. The LE model predictions, while accurate in the early stages of the diverge from the experiments at around 0.5 after Start of Injection (aSOI). The ELMO predictions are shown for two different transition criteria, one for LVF based transition of 0.005 and the other for 0.015. For both these conditions, the ELMO predictions are much closer to the experimental predictions up until 1.2 ms aSOI that translates to about 45 mm downstream of the nozzle exit. This is in spite of much fewer tunable constants for the primary atomization process compared to standard LE models.

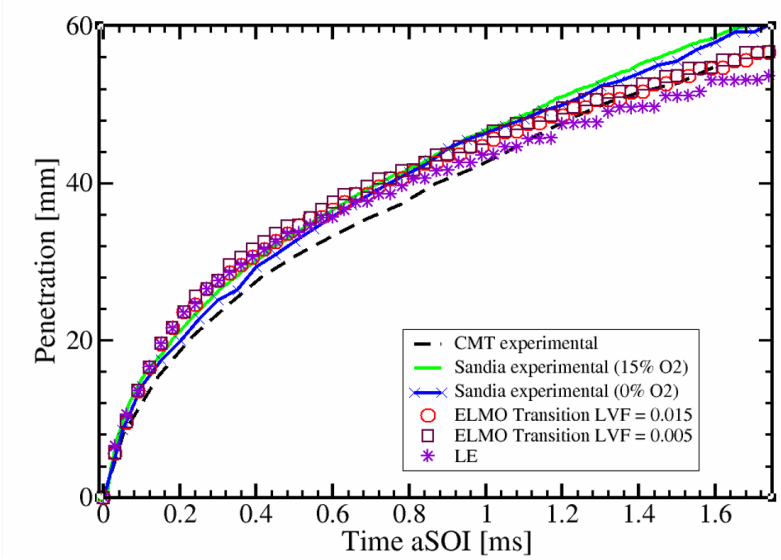

Figure 5.12. Vapor penetration plots for Spray A comparing experimental results with ELMO and LE predictions.

The mixture fraction is an important metric in estimating the amount of fuel at a location downstream of the nozzle exit. This is of particular interest from the perspective of fuel concentrations from a combustion perspective. We measure the mixture fraction at two downstream locations, 25 mm and 45 mm. The ELMO predictions are very comparable in the peak value and the profile of the spray to the experimental observations at around 25 mm downstream. This translates from the vapor penetration plots discussed previously. One important factor to note here is that for LVF transition criterion of 0.015 the predicted peak is much lower and comparable to the LE predictions. Whereas for the 0.005 transition criterion, the ELMO predictions match the experiments particularly well. This further reinforces the initial thought behind choosing these arbitrary transition criterion. Further downstream at 45 mm both the ELMO and LE predictions are slight off consistent to the observations made in the vaporization plot. The transition from ELMO capsules to LE parcels occurs at around 4-5 mm downstream of the nozzle exit, therefore at 45 mm location, the effect

of ELMO is not as evident. This explains the overlap between ELMO/LE predictions further downstream.

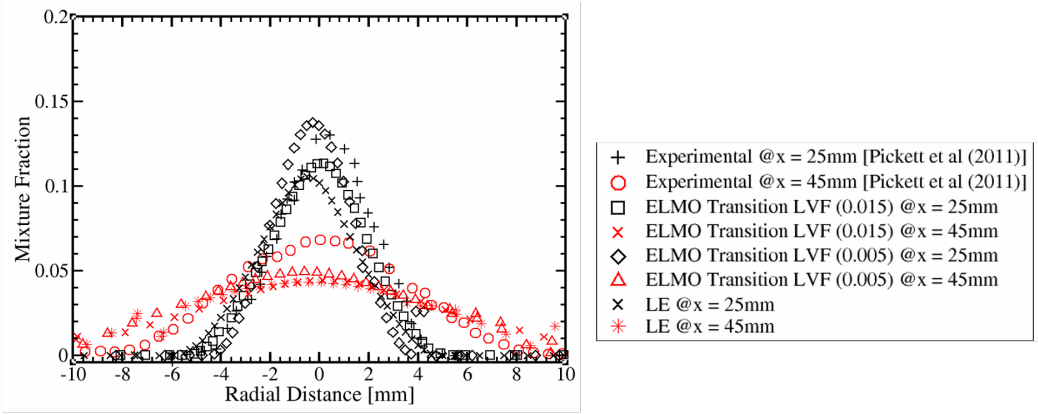

Figure 5.13. Mixture fraction comparisons at x = 25 mm and 45 mm.

The other relevant comparison for ELMO is the prediction of the gas phase velocity. Per the mathematical formulation of ELMO, using the MC points the mass, momentum and energy source terms are exchanged with the gas phase cells. To validate the effect of these source terms, based on assumptions of local homogeneity, comparisons for the gas phase velocity to the experimental observations from literature [62] are made. The conditions are for Spray A injector. Figure 5.14 shows the gas phase velocity predictions of ELMO track similarly to the experimental observations, and is inverse proportional to the axial location. The profile of the gas velocity is similar, while ELMO slightly overpredicts the magnitude. This can be attributed to the assumptions in the model about the mixing limited physics and locally homogeneous flow. Overall this is a satisfactory result.

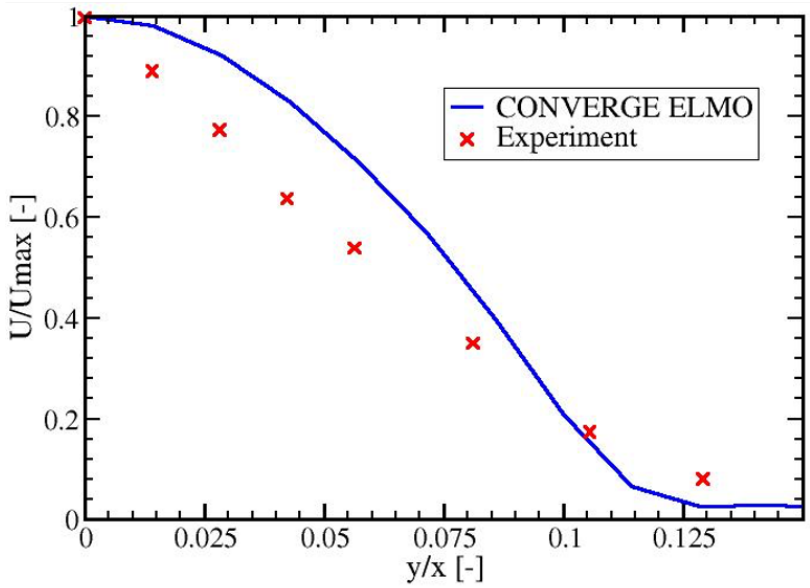

Figure 5.14. Gas phase velocity predictions against experiments shows comparisons [62]. The y/x axis is the transverse direction at 15 mm downstream of the nozzle exit.

5.3.2.2 Spray H

The next test case is the ECN Spray H [1]. The Spray H injector is a single axial hole injector with a sharp entrance, more likely to cavitate than Spray A. Indicative of this sharper nozzle entrance, the value of the coefficient of area is lower than Spray A. The injected fuel is n-heptane, which is more volatile than dodecane, making Spray H a substantially different target than Spray A. The conditions used in the simulation are given in Table 5.3. The rate of injection was described by an idealized profile, rather than assumed constant.

Number of Injectors	1
Injection Type	Single
Injection Duration	6.8 ms
Gas Temperature	1000 K
Injection Temperature	373 K
Nozzle diameter	100 μm
Turbulence model	k-epsilon RNG
Spray Angle	23 deg
C_a	0.86

Table 5.3. Spray H experimental conditions

Figure 5.15 shows the vapor penetration and mixture fraction results respectively. For vapor penetration, ELMO predictions are higher than LE yet comparable to the experimental measurements [164]. The mixture fraction predictions show ELMO outperform LE. However the profile of the spray in ELMO and LE both are wider than the experimental measurements. It is to be noted that because Spray H is more volatile and the gas temperature so high, the capsules vaporize quickly and the transition to parcels occurs only at very early times, near the start of injection. This essentially means that at the locations of measurements downstream of the nozzle, the ELMO capsules have transitioned to LE parcels.

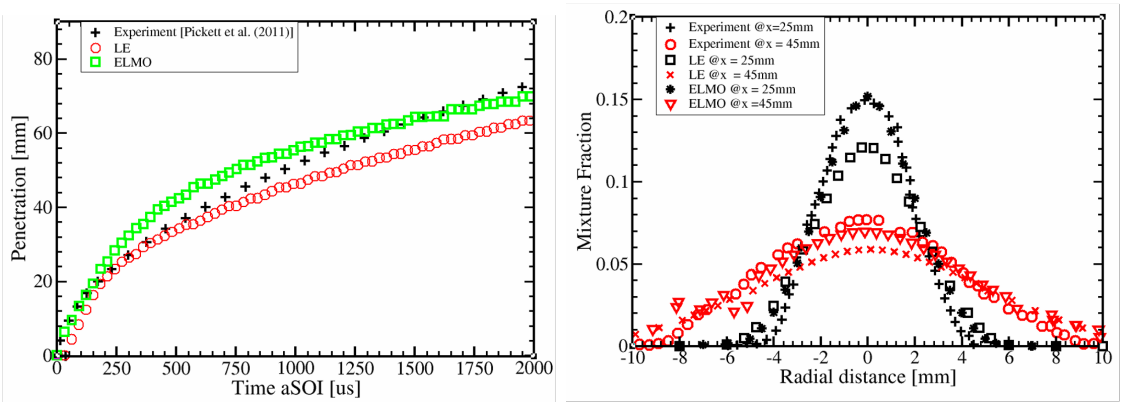

Figure 5.15. Vapor penetration comparisons (on the left) and mixing predictions at x= 25 mm and 45 mm show close predictions by ELMO to experimental results

5.3.2.3 Spray G

The final test is the Spray G injector. This is a multi-hole gasoline injector with 8 holes [1]. The fuel used in this injector study is iso-octane under non-flashing conditions [1]. While the mixing limited hypothesis by Siebers [205] was essentially built for diesel type conditions, this test case for a gasoline injector will be an empirical study in the model's prediction abilities. While being a multihole injector, the ELMO simulations have been carried out on a single sector mesh comprising of a single nozzle. The conditions for simulations are shown in Table 5.4. The vapor penetration predictions are shown in Figure 5.16. The spray angle, a major input to the ELMO model, used in this simulation are inspired by the work of Sphicas et al. [213] and Payri et al. [201]. This demonstrates the sensitivities of the inputs to the model performance. The vapor penetration plot shows that previously noted spray angles [213] of 34 degrees underpredicts the penetration, whereas the Payri [201] measurement of 25 degrees is more accurate in predicting the spray penetration. Surprisingly the ELMO predictions are closer to the experimental observations for lower angles than reported in the original work [213].

Number of Injectors	8
Injection Type	Single
Injection Duration	0.78 ms
Gas Temperature	573 K
Injection Temperature	363 K
Nozzle diameter	165 μm
Turbulence model	k-epsilon RNG
Spray Angle	25/34 deg
C_a	0.68

Table 5.4. Spray G conditions

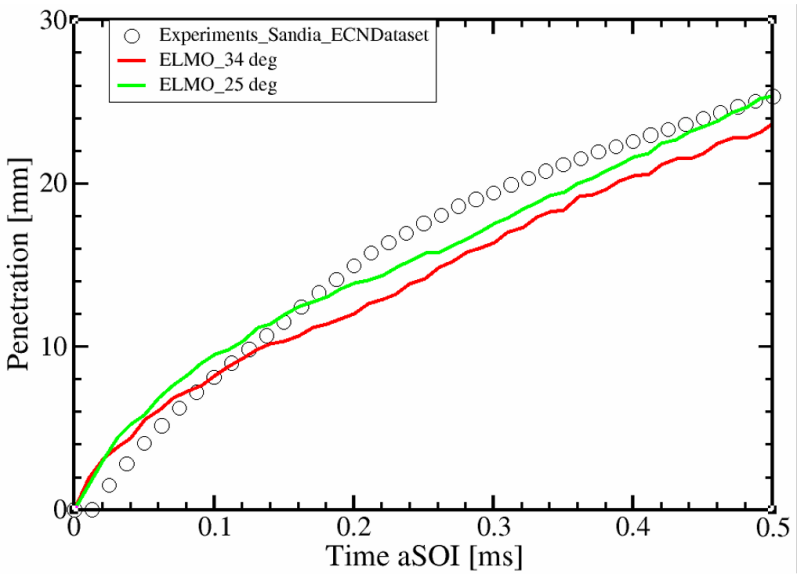

Figure 5.16. Vapor penetration comparisons with two spray angle inputs, show dependency on the initial setup on the model predictions.

5.3.3 Engine study

To demonstrate the effect of incorporating this current version of ELMO in a combustion simulation, a test engine case CAT 3400 is simulated. This is a diesel engine, based on the Caterpillar 3401 design. The simulation has a single injection event, between -9 CAD to 12 CAD, where CAD is the Crank Angle Degree. The Engine cycle is between -142 CAD to 70 CAD. The initial temperature is set of 355K, and species fraction set to O2 (23%) and N2 (76%). The combustion model used is a finite rate chemistry model. Figure 5.17 compares the in-cylinder pressure trace between the experimental observations, LE and ELMO. The experimental validations for the pressure trace are obtained from [199]. The ELMO predictions for pressure trace are about 5% lower than the experimental predictions, while maintaining the profile. This can be attributed to the lower temperature predictions from Figure 5.11 that may suppress combustion. This observation translates to Figure 5.18 where the mean temperature predictions by ELMO and LE (left panel) differ after the point of injection. This is expected to have an effect on the species prediction shown on the

right panel. Here we see that ELMO predicts species that indicate partial oxidation that can be contributed to the lower temperatures in the gas phase. This is an area of future improvement for the ELMO model

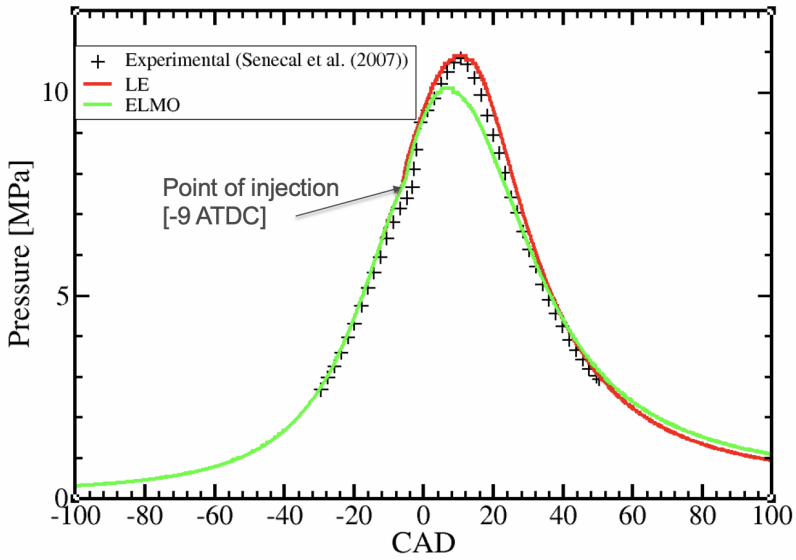

Figure 5.17. In-cylinder pressure trace shows ELMO following the profile however predicting 5% lower than experimental observations.

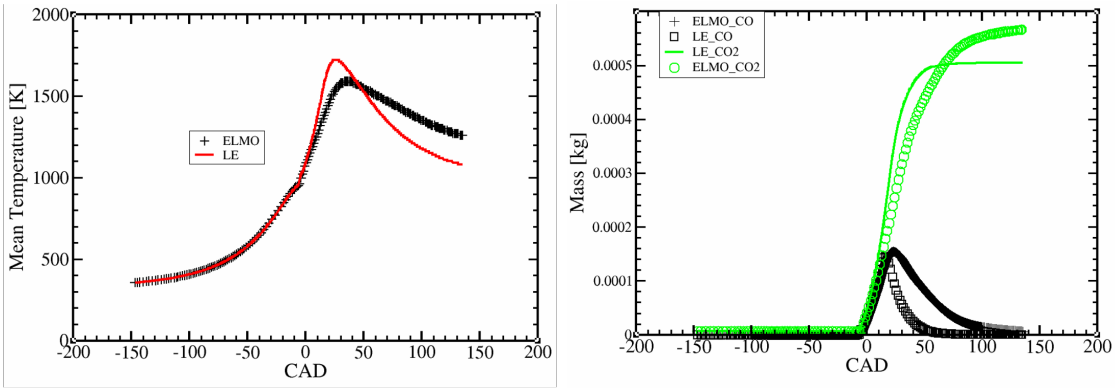

Figure 5.18. The mean temperature profile (on the left) and the species fractions on the right show the differences in ELMO predictions compared to standard LE. This is consistent to the lower than adiabatic mixing temperature predictions from ELMO shown earlier.

5.3.4 Parametric studies

The hyperparameters in the ELMO model such as the number of MC points, the sensitivity to the grid size, and the effect of the turbulence models have been studied. All the test cases are based on the Spray A injector introduced earlier. Figures 5.19 and 5.20 show the effect of different numbers of MC points on the mixture fraction and vapor penetration. There are no discernible effects of the number of points chosen and the model predictions. This can be attributed to the very small capsule sizes that even five hundred points are seemingly sufficient. This is encouraging as choosing fewer points makes the calculations faster and does not effect predictions.

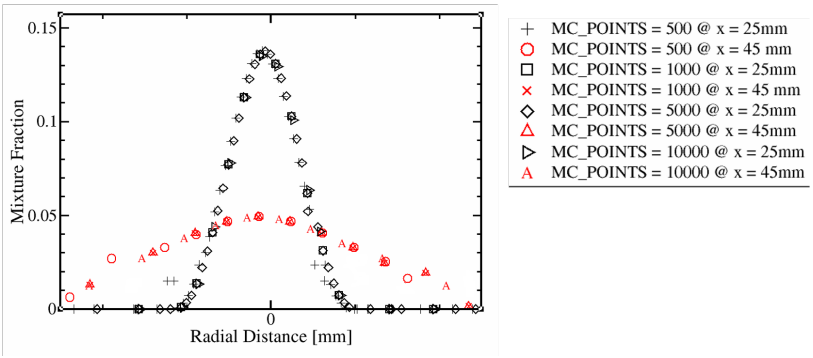

Figure 5.19. Effect of different MC points on mixing fraction

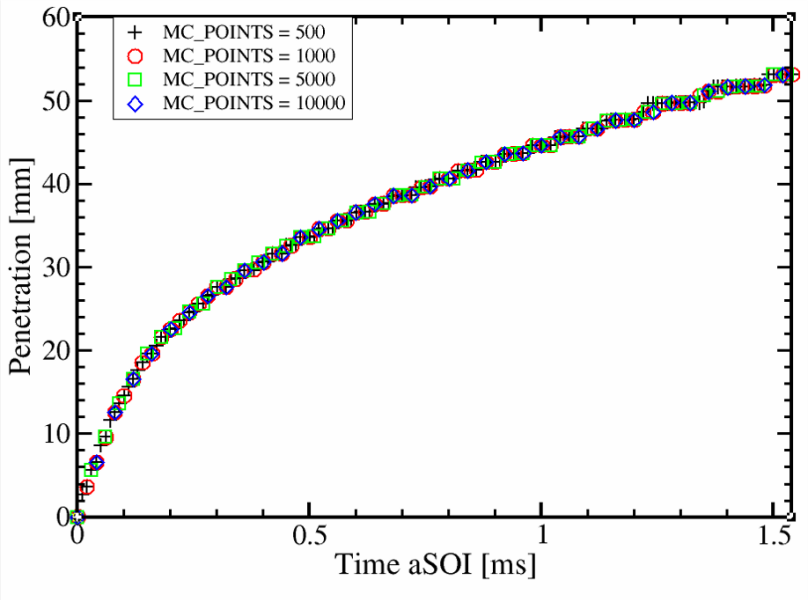

Figure 5.20. Effect of different MC points on spray penetration

Turbulence models have a major role to play in determining the mixing at locations downstream. It is therefore expected to play a role in the prediction quality. In this study we compare the $k-\epsilon$ and $k-\omega$ models. In Figure 5.21 we compare the effect of different turbulence models on the mixing profile for ELMO and LE simulations. As expected there are some minor differences but not much to convince that the solution is dependent on the choice of turbulence models. It is to be noted that for most of the hot ambient conditions, the ELMO capsules transition to Lagrangian parcels upstream of the locations where these measurements are made. Therefore it is not surprising that the ELMO predictions mirror the LE behavior in terms of turbulence model sensitivity.

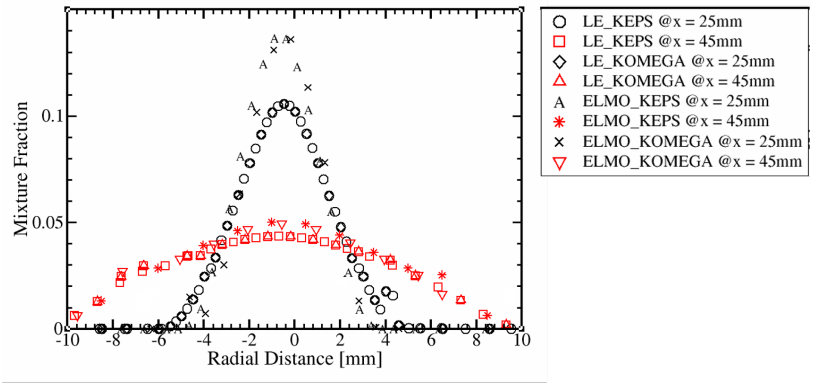

Figure 5.21. Effect of turbulence models on mixing predictions

The effect of the underlying mesh resolution on the spray is expected to play an effect on the prediction quality. This is especially true for the ELMO model where the two way exchange of source terms and mass/momentum takes place between the capsules and the mesh. In this study we find some effect of the grid size on the penetration and mixing shown in Figure 5.22 for Spray A conditions. This is consistent with our expectations.

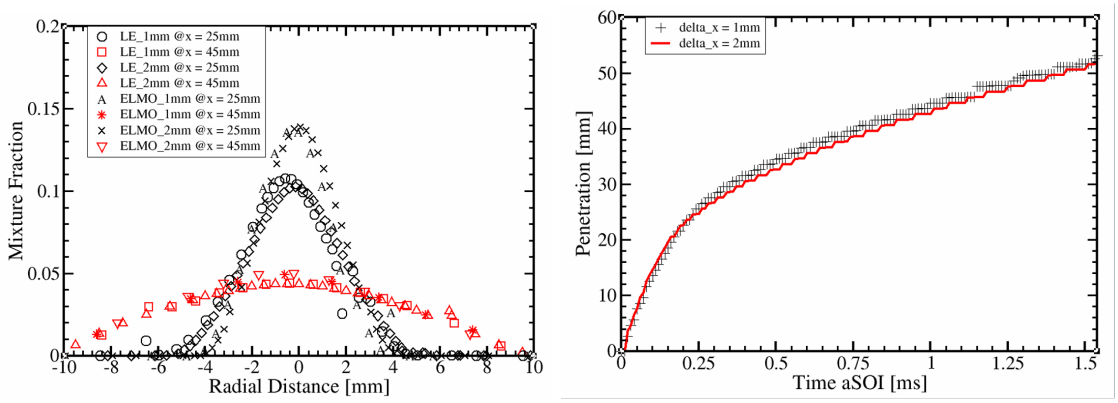

Figure 5.22. Effect of Eulerian grid mesh size on predictions

5.4 Summary

To summarize, the Eulerian Lagrangian Mixing Oriented (ELMO) model is a novel extension of the previous efforts in building physically consistent reduced order spray

models [205, 49, 150, 159] for the dense core region. The challenges in incorporating these quasi-1D models into a multidimensional CFD code was overcome by applying Monte Carlo integration and developing the concept of a two-phase capsule structure using a Lagrangian formulation. Some key takeaways and challenges from this study include:

- This new model, ELMO, is based on ideas of thermodynamic and inertial equilibrium in the dense spray core, resulting in an approach that targets mixing-limited conditions in the near-nozzle region, where conventional LE approaches fail [156].
- The ELMO model's performance for diesel sprays in predicting spray penetration, mixing and gas phase velocity is satisfactory and comparable to experimental observations [164].
- The comparable performance of ELMO to gasoline Spray G injector's spray penetration indicates that atleast in the near-nozzle region, the physics can be modeled as mixing limited.

CHAPTER 6

DATA-DRIVEN TURBULENCE CLOSURE

6.1 Motivation

In the previous chapters, we discussed modeling strategies ranging from engineering-level models (RANS/HRMFoam) to reduced order models such as ELMO in modeling different aspects of the internal flow, near-field and the far-field spray. In this chapter, we will explore the possibility in building a data-driven surrogate model for complex physical processes. While the long term objective is to establish a path towards physics-guided data-driven coarse-graining, in this chapter we shall focus on building a data-driven algebraic turbulence closure model.

In this chapter, we present an automatic data-driven machine learning (ML) approach for the development, evaluation and interpretation of deep neural networks (DNNs) for turbulence closures and demonstrate their usage in the context of cold-flow large-eddy simulation (LES) of the four-stroke Darmstadt engine using an open-source compressible multi-dimensional CFD solver OFICE, in a hybrid PDE-ML framework. Rather than explicitly using canonical formulations of closure terms, these DNNs robustly discover the functional relationships between the large-scale features of the resolved flow (cell Re, strain and rotation rate tensors etc.) obtained by solving the Navier Stokes to the small-scale unresolved terms. Experimentally validated high-fidelity LES solutions of the engine at different crank angles are utilized as the ground truth to train the DNN based closure models. Since optimizing these DNNs can be a laborious process for scientific datasets, and often require expertise, we propose a Bayesian optimization framework that automatically determines the best set of net-

work parameters, including the architecture and training hyperparameters - batch size, regularization etc. for optimum performance. We compare and contrast various networks for their effectiveness in an a-priori testing setting. Finally, the best 'learnt' network is integrated with the open-source CFD solver (OFICE), and solutions are obtained over several injection cycles. These experiments reveal that the DNN models temporally track resolved scalar variance with a good accuracy. Additionally, we interpret the artificial neural networks with sensitivity analysis to determine the relevant large-scale features for the learning process.

6.2 Introduction

Multi-dimensional Computational Fluid Dynamics (CFD) modeling of the in-cylinder flow dynamics is a fundamental tool for the design and development of the Internal Combustion Engines (ICE). At the center of the ICE is the process of transient combustion caused by oxidizing energy-dense fuel in a combustion chamber. One of the key limiting factors in the combustion process is the phenomena of fluid turbulence that affects mixing and the energy exchange. Therefore, this non-linear, non-local phenomenon has been at the center of modeling efforts in the CFD community [197, 186, 64]. Fluid turbulence is a multi-scale phenomenon and is an essential component of modeling engineering-relevant flows such as in ICE [see Figure 6.1 for the range of scales typical in an ICE simulation]. While solving the full Navier Stokes using Direct Numerical Simulation (DNS) results in the most accurate representation of the complicated, non-linear, non-local, multi-scale phenomenon, DNS is often computationally intractable. Engineering level solutions based on Reynolds Averaged Navier-Stokes (RANS) and Large Eddy Simulations (LES) alleviate this issue by resolving the larger integral length scales and modelling the smaller unresolved scales to reduce the simulation complexity. These models however suffer from the difficulty of turbulence closure. The linear eddy-viscosity model represents one of the most

popular methods for Reynolds stress closure for Smagorinsky-LES models [65, 208]. However, these approximate models are commonly phenomenological/heuristic in nature and thus require fitting to high fidelity DNS datasets for idealized flows [170]. With the incredible strides in the development of sophisticated Machine Learning (ML) algorithms made in the last decade it is only logical that these tools be widely adopted for use with scientific applications (refer to [109, 33, 56] for a review on ML for fluids applications). In particular, Artificial Neural Networks (ANN) have shown great performance for function approximations [246] and recently Deep Neural Network (DNN)-based approaches for fluid problems have gained wide attention [171, 44, 178, 123]. Specifically, the use of ML methods for developing data-driven Reynolds stress closures has shown a lot of promise for canonical as well as complex engineering flows [123, 134, 202, 242]. We extend this effort to emulate a data-driven closure term for compressible flows relevant to modeling advanced propulsion systems. In order to isolate the effects of turbulence, the scope of the current study is the cold-flow ICE simulation. Lastly, interrogating the trained network in an inference setting needs to be fast. This is because this inference engine will be used at each cell node and each timestep. A modestly slower engine would lead to massive latencies between the PDE solvers and the neural network inference. It is with this motivation that we investigate effects of network pruning and quantization within the scope of this study.

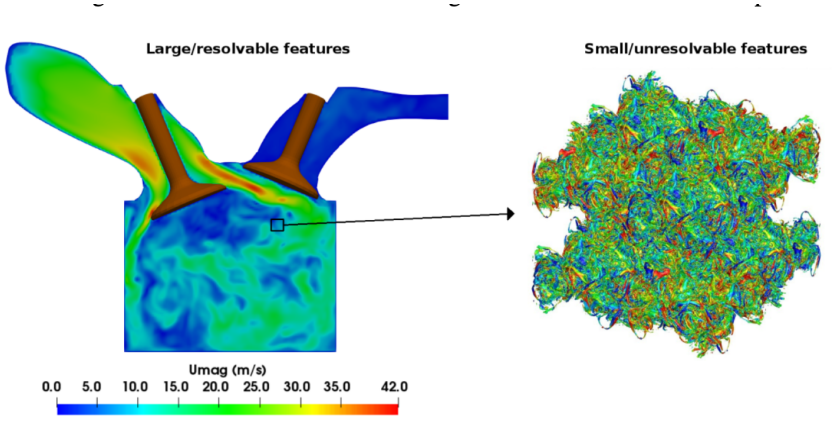

Figure 6.1. Mutiscale nature of turbulence in a typical Internal Combustion Engine (ICE) simulation. Figure adopted from [51].

6.3 Our Problem

In this study, we demonstrate the capability of DNNs to learn an algebraic LES closure, as a function of filtered variables. The algebraic closure approach are one of the most popular closure models used in the literature for related efforts [208, 93, 244]. The machine-learnt model is expected to relate the resolved flow parameters to the sub-filter (unresolved) scale [see Figure 6.1]. While Portwood et al. [171], discusses the limitations of the so called functional models such as the Smagorinsky type models, which are phenomenological and limited in the range of dynamics they are able to model, these models are comparatively numerically stable and therefore of great interest to the engine community [93, 51, 24, 25]. A major criticism of these data-driven turbulence models is that they act like a black-boxes. In order to alleviate this shortcoming, we have made efforts to incorporate physical constraints and statistical symmetries into a set of physics-informed inputs discussed in more detail in a later section.

When training a neural network, one of the the main difficulties is in setting the tunable hyper-parameters that directly dictate the performance of the data-driven

model. Manually tuning these hyper-parameters requires expertise and a-priori information about their sensitivities, which can be difficult to develop as more of these techniques are applied to large-scale scientific datasets. The traditional best practices suggestions based on applications such as computer vision, natural language processing among others for setting up a neural network architectures and hyperparameters translates poorly to scientific datasets. This is because scientific data are often high-dimensional, complex and/or sparse. Thereby the corresponding loss manifolds are non-convex and complicated and may have many local regions of optima (see Figure 6.2 for a schematic). Automated grid/random searches of the hyperparameter space, while is easy to implement, often take a very long and arduous path to the optimum result in such a setting. Bayesian Optimization (BayesOpt) based automatic Machine Learning (autoML) promises to be an important effort in this area where the optimization occurs without human intervention and often times yields robust results, in a limited set of evaluations. Along with tuning hyperparameters, searching for optimal neural network architecture or Neural Architecture Search (NAS), in terms of layer depth and width, has previously shown a performance boost [124, 98] . We extend this BayesOpt autoML approach for NAS tasks in this study.

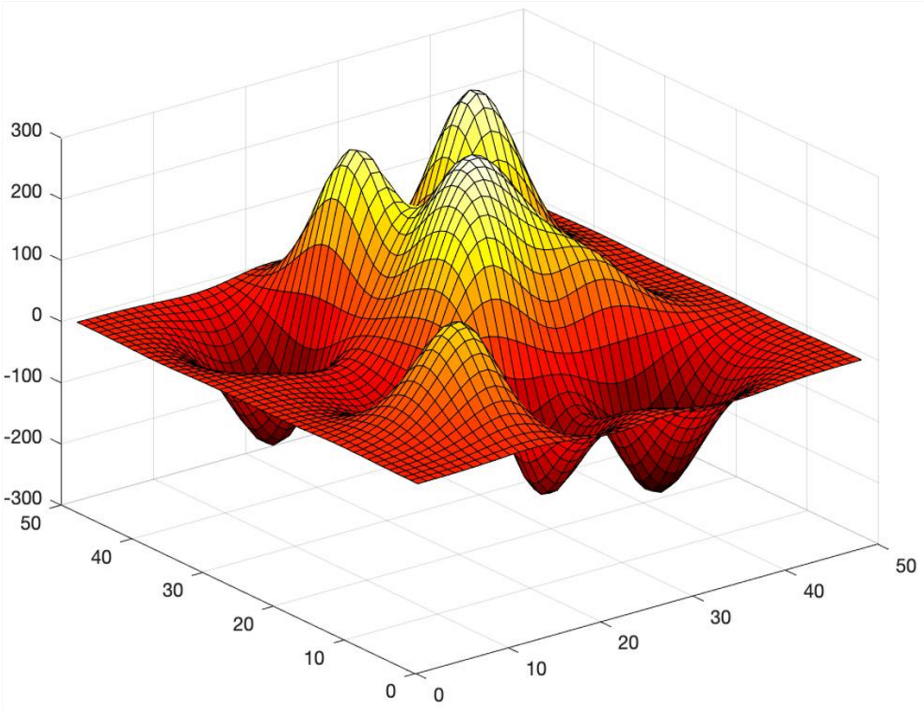

Figure 6.2. A typical non-convex loss manifold for high dimensional scientific dataset. This image is generated by translating and scaling Gaussian distributions, as commonly found in engineering and scientific datasets.

In the end, in order to demonstrate the effectiveness of the machine-learnt model, the best performing network, after the autoML step, is integrated into OpenFOAM [94, 95] and a-posteriori comparisons are made. The overall workflow is shown in Figure 2, including the intermediate step in using Bayesian Optimization for optimizing network hyperparameters. While the scope of this study is limited to deriving machine-learnt closure models, the long-term vision of this line of investigation is to build coarse-grained solutions.

While this exercise was limited to replicating and calibrating the algebraic closure models, this workflow can be extended for other closure schemes such as the one-equation and two-equation models. The final objective of this line of investigation is to build representative coarse-grained models. In a nutshell, coarsening the computational grid increases the grid-induced error. A set of coarse-grained models

will produce a surrogate model predicting the distribution of the CG-CFD local error and correcting the fluid-flow variables. Given sufficiently fine-mesh simulations, a surrogate model can be trained to predict the CG-CFD local errors as a function of the coarse-grid local flow features. This remains the subject of a subsequent study.

The overall workflow is presented in Figure 6.3. Following which numerical and machine learning methods are presented. The Results and Discussion section evaluate the performance of the data-driven model in a-priori and a-posteriori setting. Finally, the evolution of the network learning and parameter sensitivity studies are conducted to improve our understanding of these physics-informed data-driven methods.

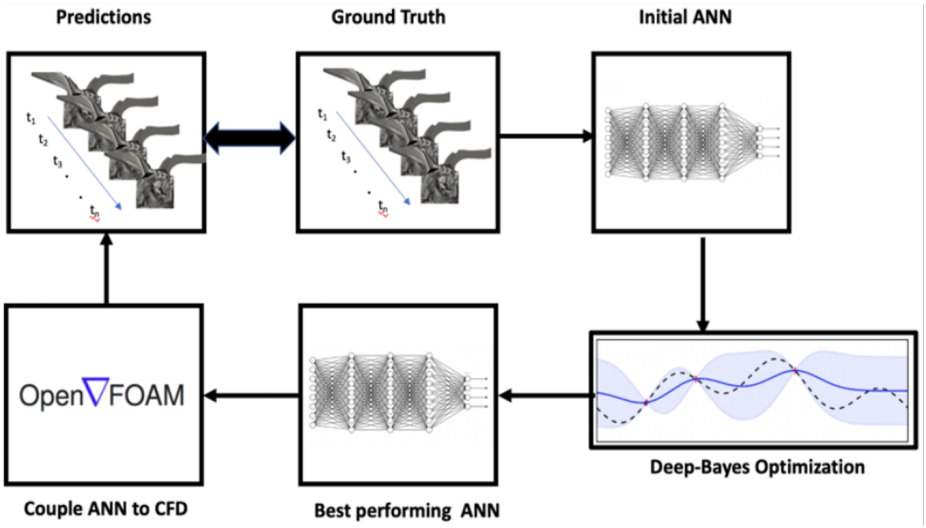

Figure 6.3. Overall workflow of the data-driven framework

6.4 Physics of the Problem

The LES filtered compressible governing equations for the mass conservation, momentum and energy and species read:

$$\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial (\bar{\rho} \tilde{u}_i)}{\partial x_i} = S_p \tag{6.1}$$

$$\frac{\partial(\bar{\rho}\tilde{u_i})}{\partial t}+\frac{\partial(\bar{\rho}\tilde{u_i}\tilde{u_j})}{\partial x_j}=\frac{\partial}{\partial x_j}[\bar{\rho}\bar{\nu}(\frac{\partial\tilde{u_i}}{\partial x_j}+\frac{\partial\tilde{u_j}}{\partial x_i})-\frac{2}{3}\bar{\rho}\bar{\nu}\frac{\partial\tilde{u_k}}{\partial x_k}\delta_{ij}-\rho\tau_{ij}^s]-\frac{\partial\bar{p_j}}{\partial x_i}+\bar{\rho}g_i+S_u \quad (6.2)$$

$$\frac{\partial(\bar{\rho}\tilde{h})}{\partial t}+\frac{\partial(\bar{p}\tilde{h}\tilde{u_i})}{\partial x_i}+\frac{\partial(\bar{\rho}\tilde{K})}{\partial t}+\frac{\partial(\bar{\rho}\tilde{K}\tilde{u_i})}{\partial x_i}=\frac{\partial}{\partial x_i}(\alpha_{eff}\frac{\partial\bar{h}}{\partial x_{j]}})+\frac{\partial\bar{p}}{\partial t}+S_h \quad (6.3)$$

$$\frac{\partial(\bar{\rho}\tilde{Y})}{\partial t}+\frac{\partial(\bar{\rho}\tilde{Y\tilde{u_j}})}{\partial x_j}=\frac{\partial}{\partial x_j}(\mu_{eff}\frac{\partial\tilde{h}}{\partial x_i})+S_{Yi} \quad (6.4)$$

where u is the velocity vector, ρ is the density, ν is the dynamic viscosity, Y_i is the mixture fraction, τ_{ij}^s the momentum sub-grid stress, h is the enthalpy, K is the kinetic energy, α_{eff} is the effective thermal diffusivity, and μ_{eff} is the effective dynamic viscosity. Also the S terms stand for the source terms added for each gas phase equation from the spray. Since in the current simulation we mainly deal with cold flow, in the absence of spray, the S terms are all set to zero. The tilde denotes filtered quantities, while bar sign indicates an averaged quantity. For data generation in this study, sub-grid stresses of the momentum equation τ_{ij}^s are modeled using the algebraic closure model proposed by Smagorinsky [208].

The sub-grid stresses, τ_{ij}^s , as shown in Eq 6.2 employ the Boussinesq hypothesis [214] and the assumption by Smagorinsky [208] that the smallest scales are isotropic. In this case, the Prandtl's mixing length hypothesis can be used to derive the subgrid viscosity in terms of one characteristic length and one velocity scale [234].

$$\tau_{ij}^s-\frac{1}{3}\tau_{kk}^s\delta_{ij}=-\mu_{sgs}S_{ij} \quad (6.5)$$

where S_{ij}, is the rate of strain tensor, calculated as shown in Equation 6, and the kinematic viscosity is presented in Equation 6.8:

$$S_{ij} = \frac{1}{2}(\frac{\partial \bar{u_j}}{\partial x_i} + \frac{\partial \bar{u_i}}{\partial x_j}) \tag{6.6}$$

$$|\bar{S}| = \sqrt{2 S_{ij} S_{ij}} \tag{6.7}$$

$$\nu_t = \frac{\mu_{sgs}}{\rho} = (C_s \Delta)^2 |\bar{S}| \tag{6.8}$$

where Δ is the filter width and C_s is the Smagorinsky constant. In summary, the above approximation for the kinematic viscosity assumes that changes in the resolved fields are slow, so that sub grid eddies can adjust themselves quickly to the rate-of-strain tensor. Thus, a closure based on a single constant is not universally true and the constant value may have to be adjusted. In this work the original Smagorinsky constant value of 0.168, as in Pope [169], was chosen based on good results achieved in previous ICE simulation works [93, 51]. Since some form of data-fitting is needed to optimize the sub grid scale model parameters even for this simplified approach, one can envisage a purely data-driven method to optimally approximate this changing constant based on large scale resolved terms, motivating our approach. More details about the LES implementation and extensive validation against experimental data are reported in [51].

6.4.1 Data-driven modeling scope

In this study, we propose to derive a functional relationship between the large-scale resolved flow features and the sub-grid scale unresolved terms in the case of cold flow simulation for an internal combustion engine. The specific quantity of interest is the subgrid scale kinematic viscosity introduced in Equation 6.8. The large-scale resolved inputs are non-dimensionalized appropriately, inspired from related literature [242]. Functionally, this relationship, in a non-dimensionalized form, can be expressed as follows:

$$\nu_t = f(R_c, S, \Omega, K, Y) \tag{6.9}$$

where R_c is the cell Reynolds number, measured as $R_c \equiv \frac{\rho u \delta_c}{\mu}$ where δ_c is cube root of the cell volume, S is the shear-rate tensor, Ω is the rotation-rate tensor, K is the gradient of the kinetic energy, and Y is the non-dimensional measure of the wall distance. The normalization factors for each term are discussed in Table 6.1.

Quantity	**Raw Input**	**Normalization Factor**
Kinematic Viscosity	ν_t	$\delta_c\sqrt{k}$
Strain-rate tensor	S	$\frac{\sqrt{k}}{\delta_c}$
Rotation-rate tensor	Ω	$\frac{\sqrt{k}}{\delta_c}$
TKE gradient	K	$\frac{k}{\delta_c}$
Wall distance	Y	δ_c

Table 6.1. Normalization of Input features

To effectively model the different regimes of the cold flow simulation, a representative sample of an entire combustion cycle is used as the dataset. The details are in Table 6.2. This dataset, agglomerated from five different snapshots, contains 18.47 million cells. The differences in the cell sizes at different Crank Angle Degrees (CAD) can be attributed to the changing topology of the cylinder necessitating the changes in the mesh structure and total count. The mesh type is composed of unstructured hexahedral grids and uses the Laplacian mesh motion library to model the changing volume. More details about the OFICE solver, based on coldEngineFOAM, including extensive validation against experimental datasets are described in full detail in the literature [51].

The dataset described in Table 6.2 is used for training as well as testing the machine learning model and divided into a train/test/validation split in the ratios 0.7, 0.1 and 0.2 respectively. The details of the machine learning models are discussed in the next section.

Angle	90 CAD	-90 CAD	180 CAD	5 CAD	265 CAD
Number of cells	3.71 mil	3.71 mil	4.43 mil	2.84 mil	3.8 mil

Table 6.2. The dataset used for training and testing

6.5 Machine Learning

Neural Networks are organized as sequential layers of neurons and they have shown excellent performance in functional mapping between a set of inputs/outputs, across various disciplines. Differential programming, based on principles of automatic differentiation, is a paradigm in machine learning where parameters of the neural network are trained by gradient-based optimizations [23, 120]. This is indeed useful for scientific ML tasks and helps to improve the parameterization of the approximations of the neural network [88]. In this work, fully connected ANNs are used to find the functional mapping between the target to the non-dimensional input features [more in Table 6.1 and Equation 6.8]. While the neural network implementation is fairly straightforward and done using MATLAB's Deep Learning Toolbox, the a-priori estimation of the best network hyper-parameters, and the architecture itself are non-trivial due to the multi-scale, non-local, non-linear nature of the fluid turbulence. To discover the best performing settings for our dataset, an Automatic Machine Learning (Auto ML) strategy is used and, while there are many available optimization methods [82], the Bayesian Optimization (BayesOpt) approach, which has shown good performance for other data-driven tasks ([209, 63]) is used. Our overall approach is shown in Figure 6.3 and the overview of the approach described in detail in the literature [[78, 140]]. For the purposes of this chapter we limit ourselves to identifying the best performing ANN, given by the BayesOpt algorithm.

6.5.1 Automatic hyper-parameter tuning using Bayesian Optimization

For a simpler setup, conducting a randomized or grid search for the best hyperparameters would be useful. However, these methods are computationally intractable for a large-scale problem such as this. Instead, we explore an automatic hyper-parameter tuning setup within the Bayesian Optimization framework, in which the learning algorithm's generalization performance is modeled as a sample from a Gaussian Process (GP) [179, 240]. The posterior distribution induced by the GP leads to efficient use of information gathered by previous experiments, enabling optimal choices for what parameters to try next. To pick the hyper-parameters of the next experiment, one can optimize the acquisition functions (introduced in the next section) such as expected improvement (EI) [144] over the current best result or the Gaussian process upper confidence bound (UCB) [215]. EI and UCB have been shown to be efficient in the number of function evaluations required to find the global optimum of many multimodal black-box functions [144, 215]. While there are many hyper-parameters that have an effect on the ANN performance and accuracy, we limit the set of parameters to the ones in Table 3, due to their leading order effect on the network performance. The autoML workflow is shown in Figure 3. Figure 6.4 provides the BayesOpt workflow and Figure 6.5 shows the components of the Neural Architecture Search.

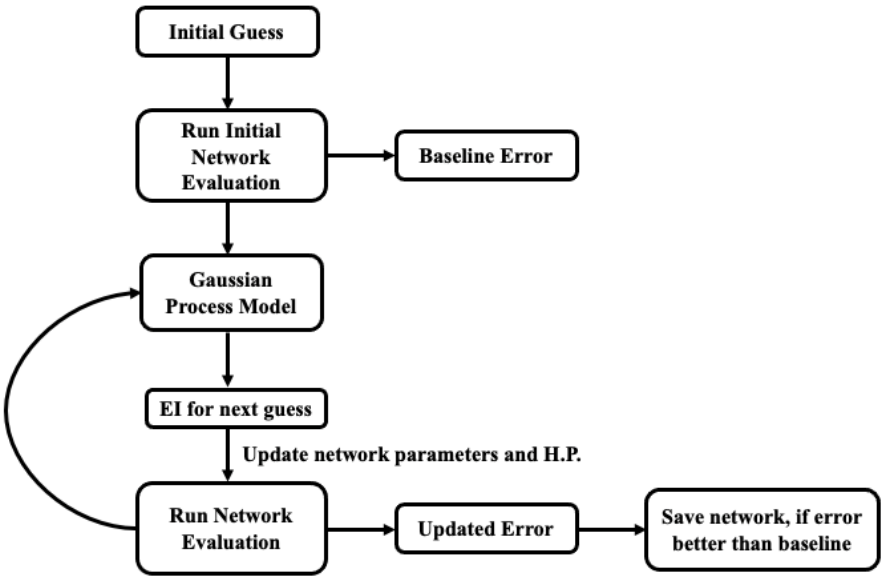

Figure 6.4. Bayesian optimization workflow. The BayesOpt is modeled as a Gaussian prior with EI used to update parameters. Within each BayesOpt evaluation, the training is run and best results against validation data stored.

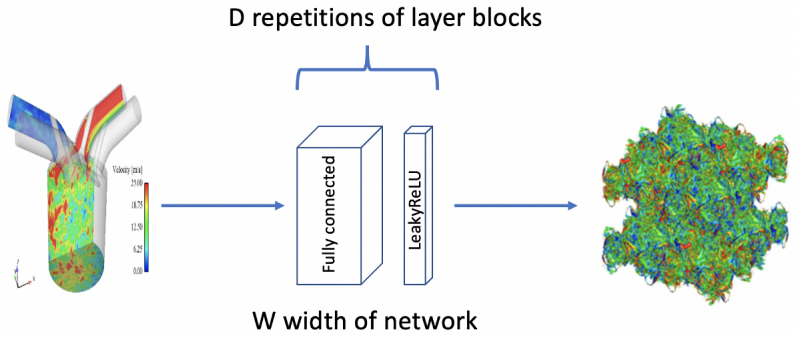

Figure 6.5. The Neural Architecture Search schematic shows the repeating blocks of weight and activation. Both the network depth and width are optimized using the BayesOpt method.

6.5.2 Acquistion function for Bayesian optimization

The acquisition functions evaluate the effectiveness of a point, x, based on the posteriori distribution function, Q [209]. The Expected Improvement function can be used to maximize the Expected Improvement (EI) over the current best. Functionally it is represented as:

$$EI(x, Q) = E_Q[max(0, \mu_Q(X_{best} - f(X))] \tag{6.10}$$

where $\mu_Q(X_{best}$ is the lowest value of the posteriori mean, and X_{best} is the location of the lowest posteriori mean.

6.5.2.1 Pruning

Accelerating the inference engine is critical to improving the cost-accuracy tradeoff. The function call to the inference engine at runtime involves querying the functional map $\mathbf{F^{-1}}$ involving repeated calculations on the trained weights and biases. Since this inference engine is used at every interior cell in the mesh during runtime, ideally a smaller matrix would lead to larger savings in compute and therefore make the cost-accuracy tradeoff more favorable, a desired goal for this exercise. We first limit this by only exploring smaller networks (less than 8 layers, 50 neurons each) within the Bayesian Optimization step. However for some hyperparameter explorations these networks tend to be much larger in size and explore the full potential permitted within the optimization step.

In this (and a related) study, we explore the Synaptic Gradient Flow [223] approach in improving inference cost over the baseline trained models using techniques of pruning. While we will refer to these techniques briefly in this section, for a full review of the work please refer the original article [142].

Typically for an engineering level fluid simulation, the number of meshpoints and timesteps are on the order of a few million. This essentially means the number of network inference function calls are of the same order. For a deep neural network that amounts to $O*T*N$ operation calls, where O is the number of matrix operations, T is the number of overall simulation time-steps and N is the total number of node points. Even if the well-trained surrogate model alleviates the need to numerically solve for PDEs for specific tasks, these repetitive costs accumulate and become expensive for

reasonably sized problem. This also leads to a latency between different components of the numerical solver that incorporates the inference calls, which slows down overall computation.

In weight-based or iterative pruning methods [38], the neurons with the lowest scores are pruned after training. However, during a neural network training for a deep network, each layer learns a different feature of the dataset. That translates to some layers with low scores or weights but relevant information that should not be lost for generalization purposes. Using the iterative pruning method while is faster, can lead to a situation where there are not enough plausible connections between two layers and causes a *layer collapse* [223]. While these pruning algorithms can indeed compress neural networks at test time, there is no reduction in the cost of training. On the other hand the Synaptic Gradient Flow (SynFlow) approach [223] uses synaptic saliency as a metric for pruning, that leads to savings during the training time itself and is defined as:

$$S(\theta) = \frac{\partial \mathrm{R}}{\partial \theta} \odot \theta; S^{in} = \frac{\partial \mathrm{R}}{\partial \theta^{in}} \odot \theta^{in}; S^{out} = \frac{\partial \mathrm{R}}{\partial \theta^{out}} \odot \theta^{out} \tag{6.11}$$

where R is a scalar loss function of the output y of a feed-forward network parameterized by θ. The term $\odot$ is the Hadamard product between two matrices. When R is the training loss L, the resulting synaptic saliency metric is equivalent (modulo sign) to - $\frac{\partial \mathrm{L}}{\partial \theta} \odot \theta$, the score metric used in skeletonization [114]. The resulting score metric is also closely related to $|\frac{\partial \mathrm{R}}{\partial \theta} \odot \theta|$. In the SynFlow approach, [223], the neuron-wise conservation of synaptic saliency is maintained such $S^{in} = S^{out}$.

The overall end-to-end network training approach can therefore be summarized as the following workflow image shown in Figure 6.6.

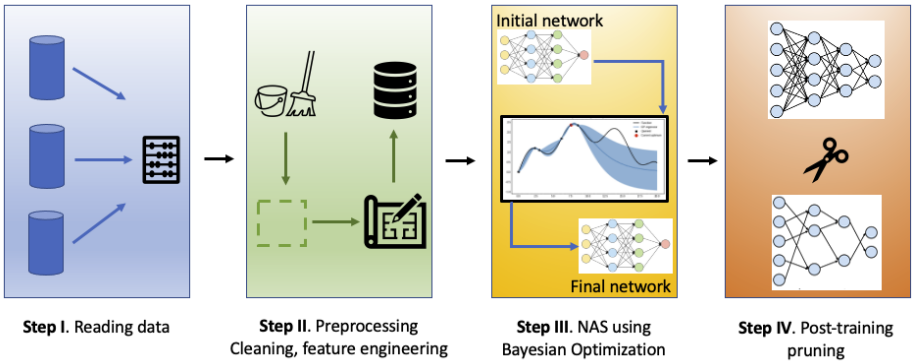

Figure 6.6. The overall workflow used in training the surrogate model. The last step of pruning, is optional and only used if extra savings would be needed

6.5.3 Solvers and Initialization of learnable parameters

The functional mapping, using a neural network, can be viewed as a function of the learnable parameters such as weights and biases. In fact, the training process of the neural network is often stochastic - very often two separate runs do not yield the same result. This can be attributed to the choice of optimizer, the precision of the learning process, as well as the complex, non-convex loss manifold and the initialization of the weights. This last step has a determination on the performance of the neural network during the optimization process itself [132, 68, 81]. In this work we explore the effect of choosing different initialization, from well-known ones including Glorot [68], He [81] and narrow-normal. In addition to the initialization of the weights the performance and evolution of the training process itself are studied by using different solvers such as ADAM [107], RMSProp [228] and SGDM [221], within the autoML framework.

6.5.4 End-to-end workflow for coupling to CFD

The best performing network, following the auto ML step is integrated into the OFICE solver [51] built using OpenFOAM [95], employing steps described in Figure 6.7. The network is first converted into a C++ source code using the MATLAB

Coder. The C++ source code is then integrated into OFICE as a special turbulence class.

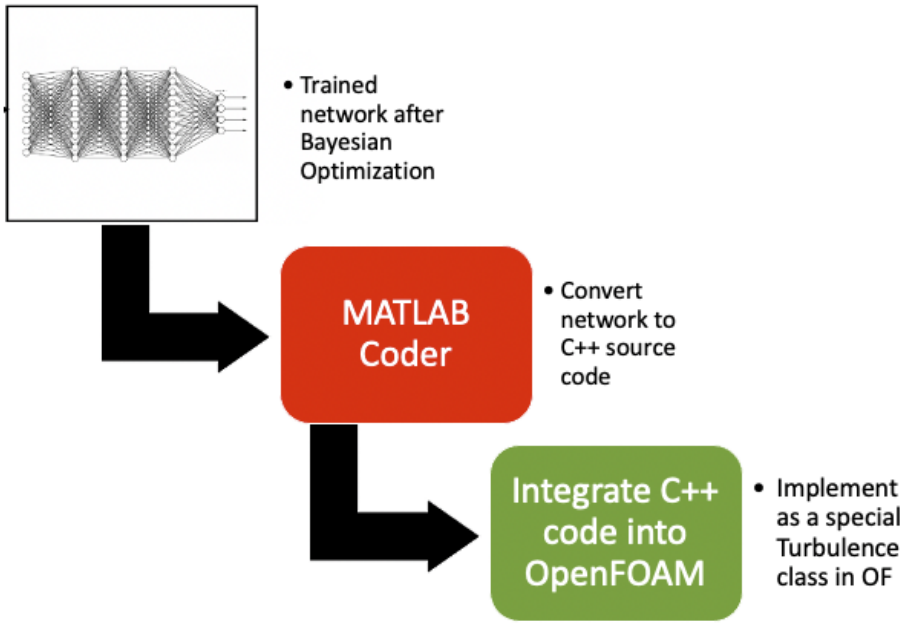

Figure 6.7. The overall workflow in implementing the trained network into OpenFOAM[78].

6.6 Results

The performance of the machine learning model are evaluated in an a-priori and a-posteriori setting. In the a-priori setting, the trained network is tested on an unseen dataset sampled from the same distribution of the training data. To evaluate an a-posteriori performance, the best network is integrated into the OFICE solver. The LES simulation with data-driven kinematic viscosity closure is then simulated for a five CADs and the data written out every 0.5 CAD. This is then compared to the ground truth dataset obtained from the experimentally-validated LES simulation [29].

6.6.1 Evaluating a-priori performance

We used a well-resolved and validated Large Eddy Simulation dataset [51] containing over 18 million datapoints, taken across five time-steps from a simulation of

a compressible flow within the cylinder of a gasoline engine generated using OFICE [51]. A subset of the dataset, about 20%, is set aside for a-priori testing. The number of function evaluations for the BayesOpt is set to 50. After each BayesOpt evaluation, an error is calculated on the test dataset defined as

$$\varepsilon = \frac{abs(y_{true} - y_{pred})}{y_{true}} \tag{6.12}$$

where y_{true} is the test data point, and the y_{pred} is the value predicted by the ANN. The best performing settings for all the possible combinations of optimizers and initializations are reported in Table 6.3. For each of our network evaluation, the architecture is optimized by self-repeating blocks of Dense Layer - Leaky ReLU activation. The p value of the Dropout layer is set to 0.2, per best practices [216]. For the specific case of optimizing the network architecture, the best set of Layer Width, W, and Layer Depth, D, the latter optimizing the number of the self-repeating blocks for a given BayesOpt evaluation, are identified and reported

The final objective of this effort is to not only identify the best settings for a given choice of solver and initialization as reported in Table 6.3, but to also understand the effective cost one can expect to pay when this network is coupled to a CFD solver to make run-time inferences across multiple timesteps and grid points, in a realistic CFD simulation. This inference cost is directly proportional to the size of the network, width (W) and depth (D). Therefore the total number of terms based on the formulation $N = I * W + D * W * W + W * O$, where I refers to the input features, in this case 14, O refers to the output features, which is 1, is also reported in Table 6.3.

Table 6.3. Best Performing hyper-parameters from autoML BayesOpt. The setting column indicates the choice of optimizer and weight initialization scheme used for the experiment. NN initializer indicates narrow-normal distribution

Setting	**Batch Size**	**LR**	**Width**	**Depth**	ε	N
ADAM/Glorot	1426	9.56e-04	91	10	1e-05	84094
ADAM/He	12719	3.77e-04	50	7	2e-04	18207
ADAM/NN	12553	1.65e-04	59	3	1e-03	11272
SGDM/Glorot	6814	0.0098	91	9	2e-03	75812
SGDM/He	1161	0.0098	89	6	1e-03	48778
SGDM/NN	9609	0.0015	64	5	7e-02	21381
RMSProp/Glorot	291	1e-04	78	4	8e-04	25432
RMSProp/He	11630	1.86e-05	55	9	1e-03	28004
RMSProp/NN	1085	2e-05	89	2	1e-02	17090

It is found that the ADAM-Glorot and ADAM-He combinations perform the best in terms of mean absolute errors, although the ADAM-Glorot configuration has the highest number of parameters. The RMPSProp on average, performs better than SGDM optimizer. This can be explained as RMSProp is an adaptive learning rate method and is capable of navigating regions of local optima and whereas that SGDM performs poorly navigating ravines and makes hesitant progress towards local optima [47]. It is observed that the Glorot initialization consistently has the best performance in terms of error, ε, which is an interesting result as He and ReLU activation have performed well for Convolutional Neural Networks [43].

The ADAM-Glorot combination network tested on a representative unseen dataset reveal very good performance, compared to the ground truth dataset in Figure 6.8.

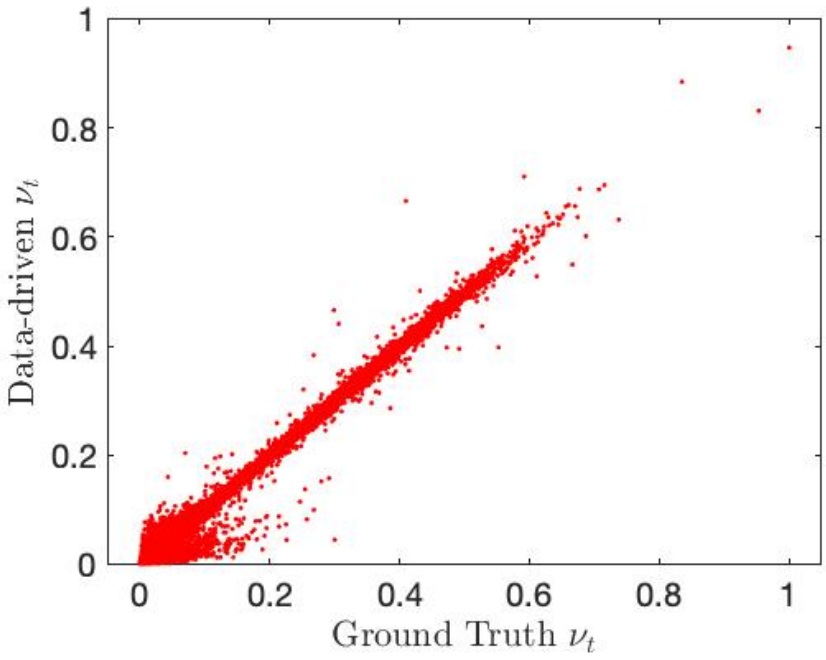

Figure 6.8. Performance of the data-driven model prediction to ground truth data indicates a good fit. Essentially the model has learnt the distribution well.

6.6.2 Evaluating a-posteriori performance

While accurately predicting the kinematic viscosity is a necessary first step, a more convincing measure of performance would be to evaluate the behavior in an actual CFD simulation. Previous studies in related areas, have shown the accumulation of errors when a machine-learnt network is integrated with a non-linear PDE solver and instances of solution divergence have been reported [31, 110]. For the a-posteriori performance, side-by-side contour plots of the center clip plane velocity [velocity magnitude including components of velocities] are shown in Figures 6.9, 6.10 and 6.11. The ground truth LES with the Smagorinsky closure model are presented on the left side, while the LES with the data-driven closure are presented on the right side of each image. The comparisons are done at 0.5 CAD, 2.5 CAD and at 5 CAD after start of simulation. These comparisons reveal that the machine-learnt closure terms are stable and produce qualitatively-similar results even after 5 CAD.

In addition to the qualitative center plane velocity comparisons, the quantitative comparisons are shown in Figures 6.12, 6.13, 6.14 . These include plotting the components of velocities for both the LES with machine-learnt closures (SGSnet) as well as from the LES with Smagorinsky closure predictions. For each set of images, the error distributions are plotted alongside. While the error distributions are similar across the three times of interest, the propagation of error from 0.5 CAD after start of simulation to 5 CAD after start of simulation do not indicate major differences. In other words, the error propagation seems to be limited for the derived quantity, in this case velocity from the Navier-Stokes equation and is in the range of 0.05%, which can be attributed to numerical precision as well as truncation errors. This further builds confidence in the quality of the model prediction.

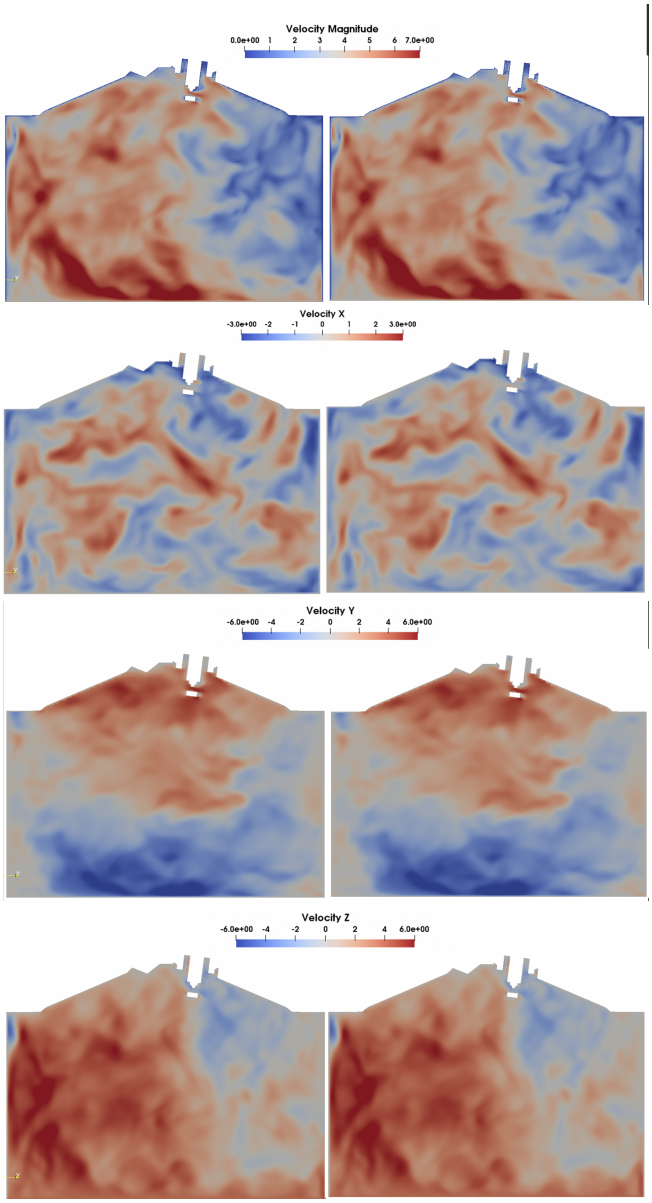

Figure 6.9. The a-posteriori velocity comparison at 0.5 CAD after start of simulation is shown in this plot. The ground truth data (on the left in each panel) and the data-driven closure based data (on the right in each panel) have been identically scaled.

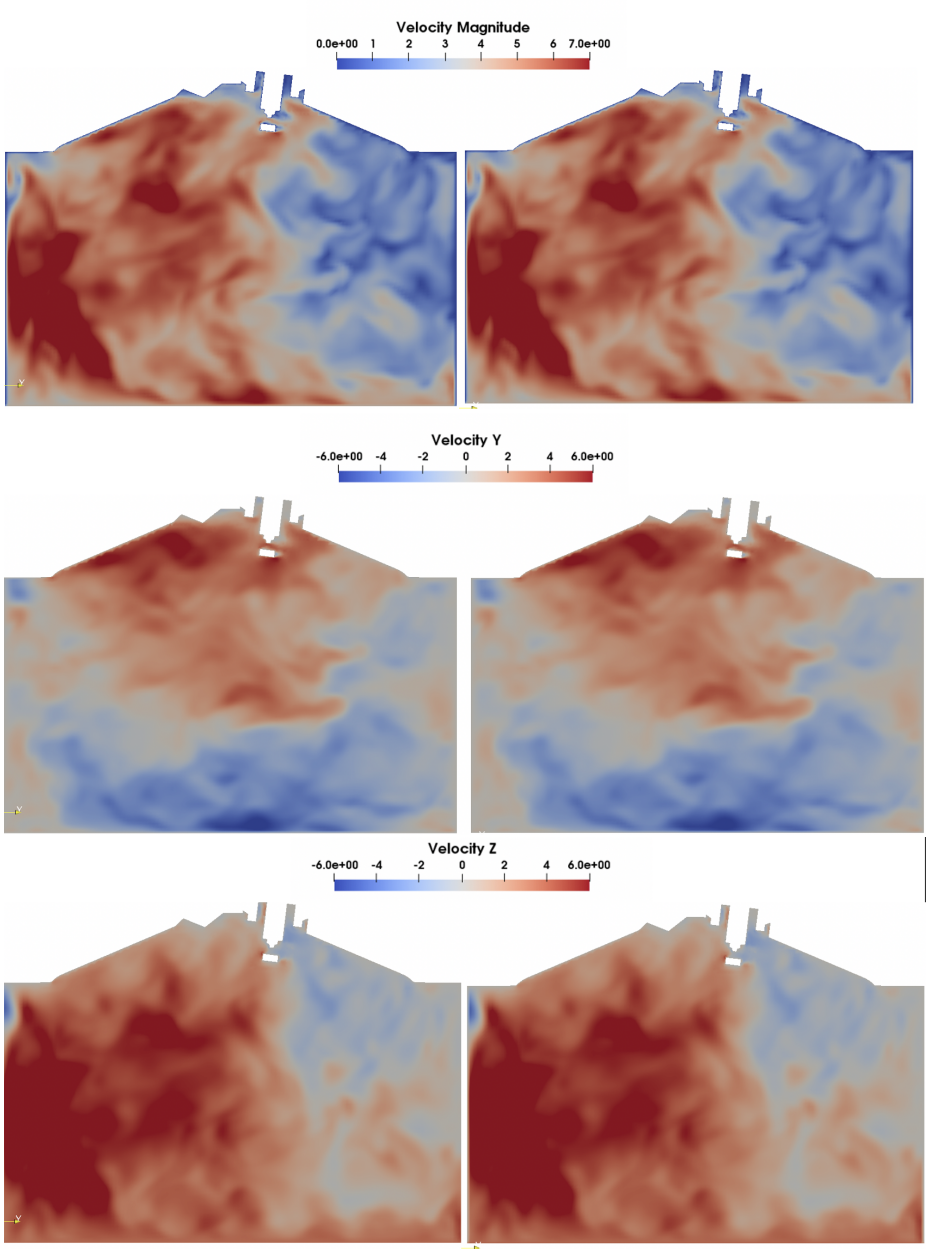

Figure 6.10. The a-posteriori velocity comparison at 2.5 CAD after start of simulation is shown in this plot. The ground truth data (on the left in each panel) and the data-driven closure based data (on the right in each panel) have been identically scaled.

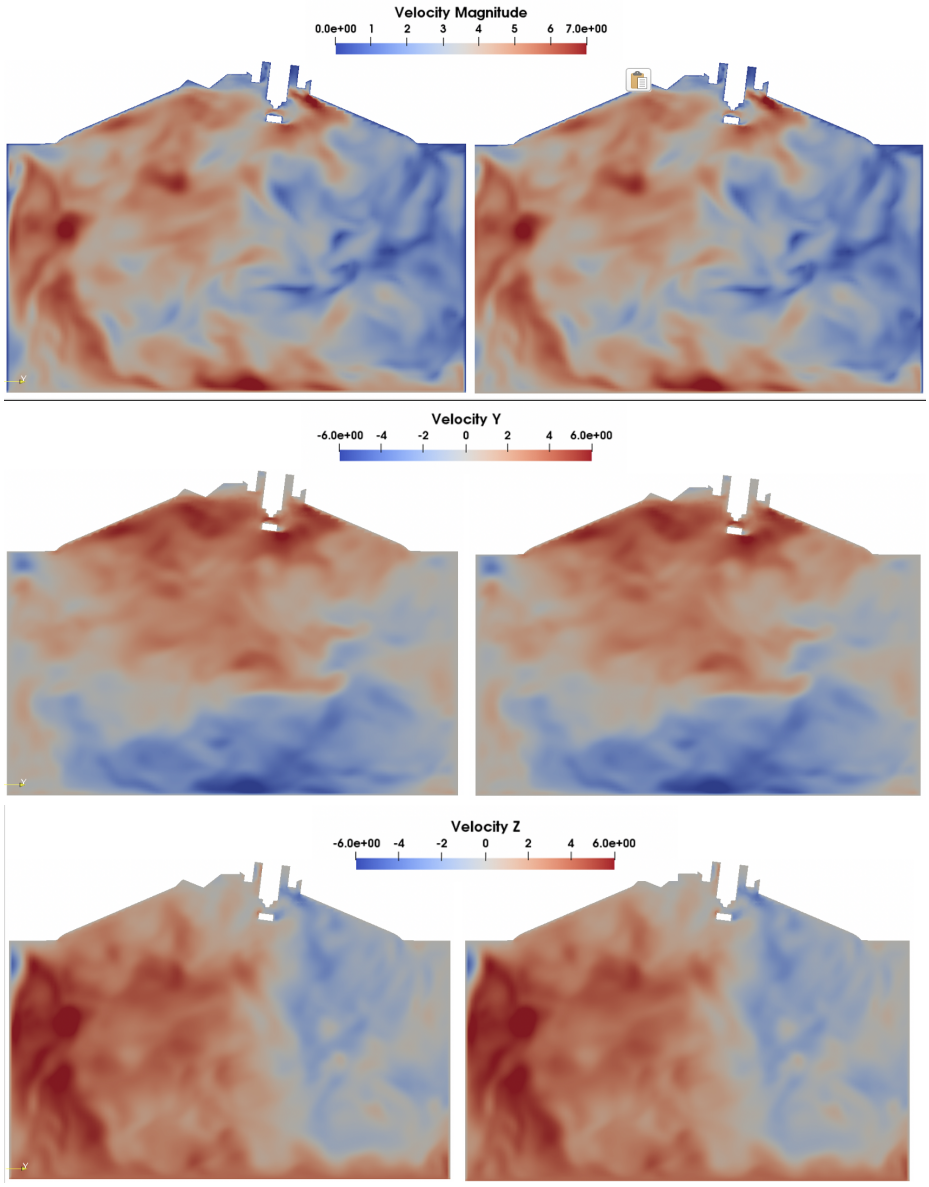

Figure 6.11. The a-posteriori velocity comparison at 5 CAD after start of simulation is shown in this plot. The ground truth data (on the left in each panel) and the data-driven closure based data (on the right in each panel) have been identically scaled.

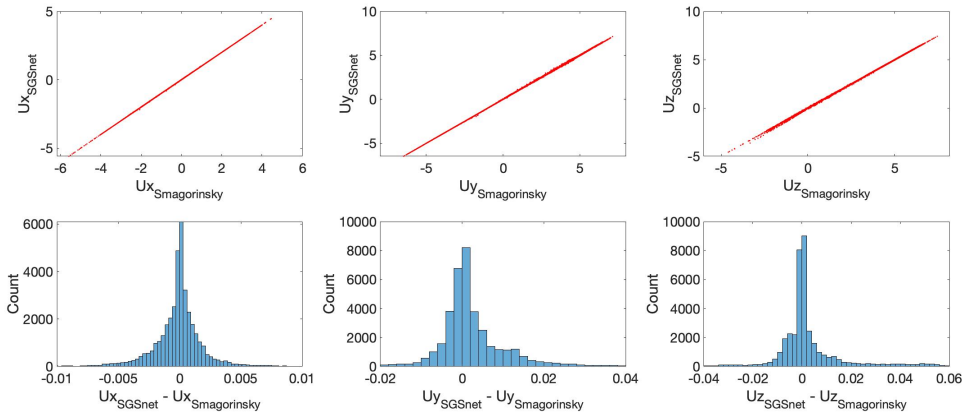

Figure 6.12. The quantitative comparisons at 0.5 CAD shows the machine-learnt closure model velocity predictions are consistent with the ground truth observations. The histograms show the distribution of the errors.

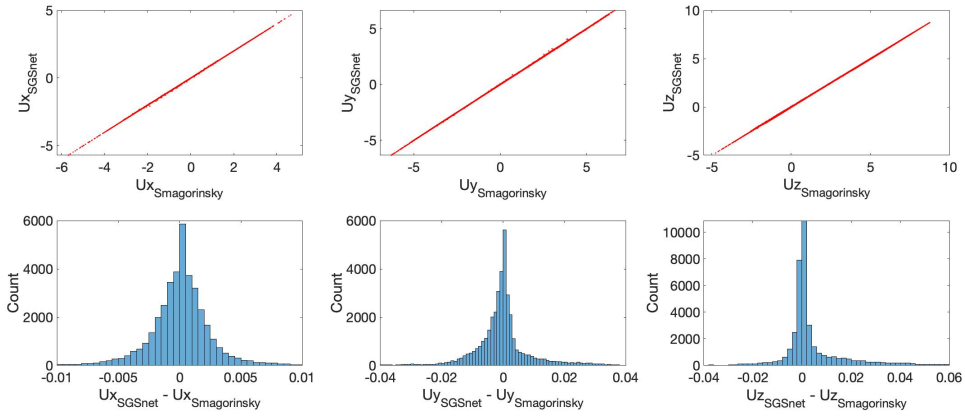

Figure 6.13. The quantitative comparisons at 2.5 CAD shows the machine-learnt closure model velocity predictions are consistent with the ground truth observations. The histograms show the distribution of the errors, and the extent of the error margins are similar to the 0.5 CAD indicating limited error propagation.

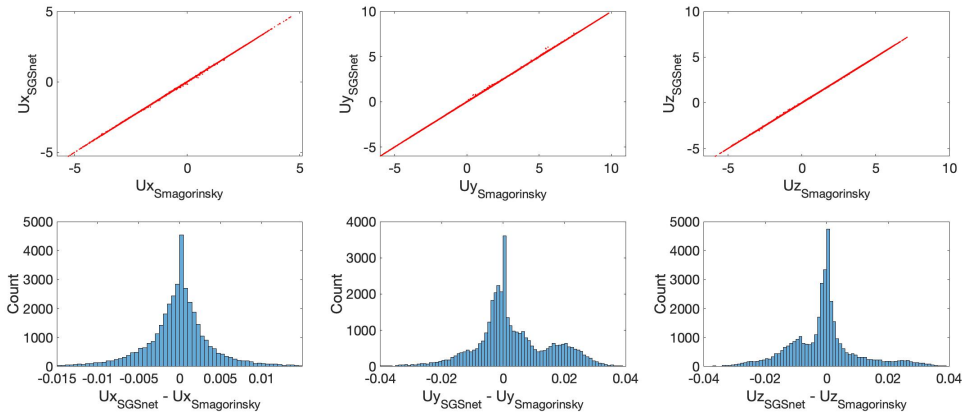

Figure 6.14. The quantitative comparisons at 5 CAD shows the machine-learnt closure model velocity predictions are consistent with the ground truth observations. The histograms show the distribution of the errors, and the extent of the error range only slightly increased from the 2.5 CAD data for Uy, indicating limited error propagation.

6.6.3 Interpretability

While the dominant representative power of a neural network is a result of the large number of nodes and their non-linear interactions, these models are often times deemed black boxes. In order to unravel the learning process better, the trained neural network is investigated in order to understand the layer-by-layer comparisons of the weight characteristics. The neural network nodes are comprised of weights and biases. These weights in each layer are initialized and when sufficiently trained, form a distribution of values which can be investigated to understand the training process better. In Figure 6.15, the distribution of the weights in Layer 2 and 3 are plotted. The data approximate a Gaussian curve, a behavior that is expected from a network initialized with Glorot initializer. The similarities in the distribution of the Layer 2/3 means that the initial layers have overlapping similarities in the learning. However, distributions only reveal one side of the story. In order to establish the extent of the similarity, a cosine similarity distribution is plotted in Figure 6.16. Cosine similarity of weights can be mathematically represented as:

$$cos(\theta_1, \theta_2) = \frac{\theta_2 \theta_1^T}{||\theta_2||||\theta_1||} \tag{6.13}$$

where θ_1, θ_2 are the weights of the ANN. A cosine similarity value of 1 means similar vectors. A value of 0 indicates an orthogonal vector that has no match. The similarity plots show that while the weight distributions are the same in both the layers, the weight values are not co-located; infact they are orthogonally located. This means that during the training process, the weights are updated differently in the subsequent layers to maximize the representation power of the network. This sheds light on the evolution of the complicated training process for a well-trained network. More analysis on the differences in the learning process for various optimizer-weight initializer setting has been discussed in the following section.

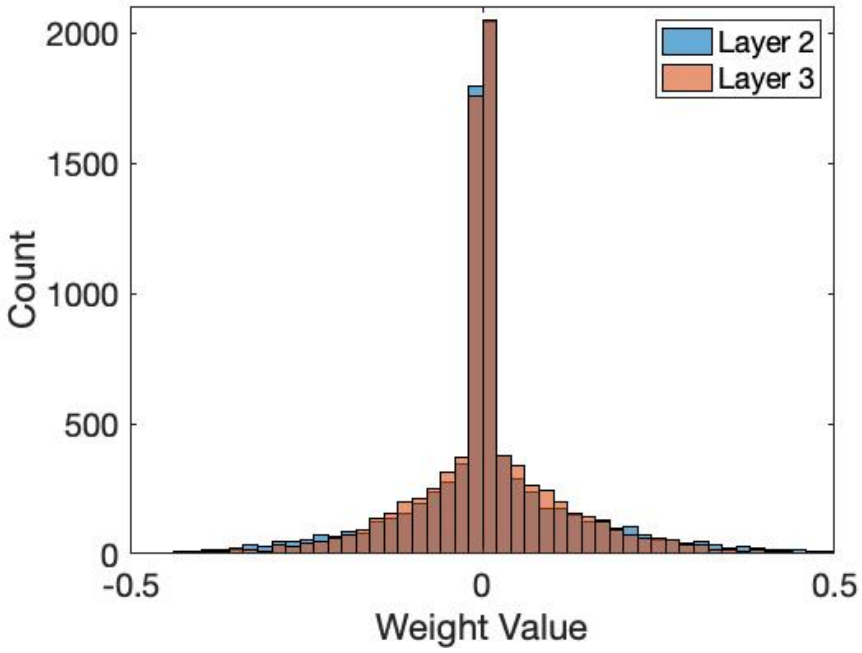

Figure 6.15. Comparisons of the weight histograms show the similarity in their distribution for consecutive layers.

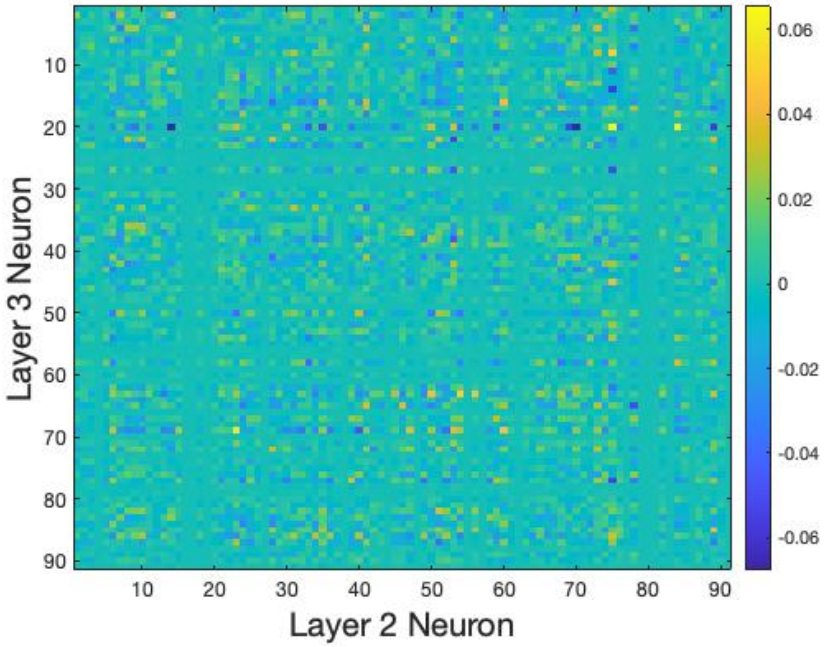

Figure 6.16. The similarity of the distribution in the consecutive layers seen from Figure 6.15 does not translate to the similarity in the weight space. The cosine similarity plot shows the two layers' weights are orthogonal to each other. This is a testament to the complicated representational powers of a deep neural network

We investigate the similarities in the weight and function space for similar as well as differently initialized trajectories, in an effort to better understand the commonalities in the optimization process. In order to do that, the simulations are check-pointed after every epoch and the cosine similarity among the weights are computed. From the left panel in Figure 6.17 , it is observed the checkpoints along a trajectory are largely similar in the weight space, in this case for ADAM-Glorot configuration. However, when compared to the same initialization, He, ADAM and SGDM we observe major differences in the trajectories in the weight space, and this can be attributed to the optimization methods and the stochasticity of the learning process therein. Therefore, we see that functions within a single trajectory exhibit higher similarity, and this similarity map is optimizer-initialization specific.

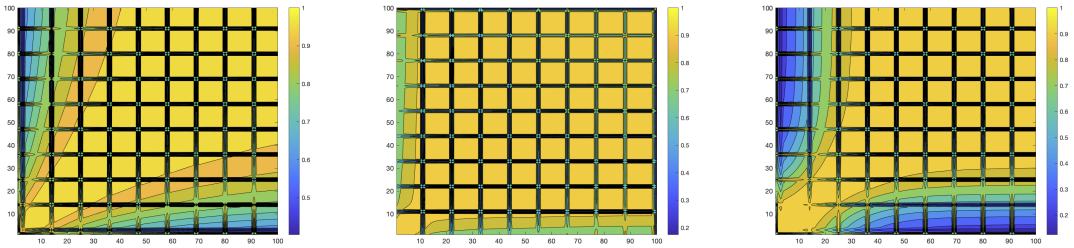

Figure 6.17. Cosine-similarity between checkpoints to measure weight space alignment for the ADAM- Glorot (left), SGDM-He (middle) and ADAM-He (right) configurations. This shows the stochasticity of the weights trajectories that can occur during the training process.

In addition to investigating the similarities in the weight space which are inherently high-dimensional and therefore non-intuitive, the use of dimensionality reduction methods is used, such as t-SNE [233, 91, 126] to observe the trajectory of the checkpoints in a 2D space in Figure 6.18. It is observed that the Glorot and He have overlapping similarities, which makes sense as they both have similar functional forms and theoretical analysis: they both find a good variance for the distribution from which the initial parameters are drawn and only differ in the type of distribution they use - Gaussian for He, and Uniform for Glorot.

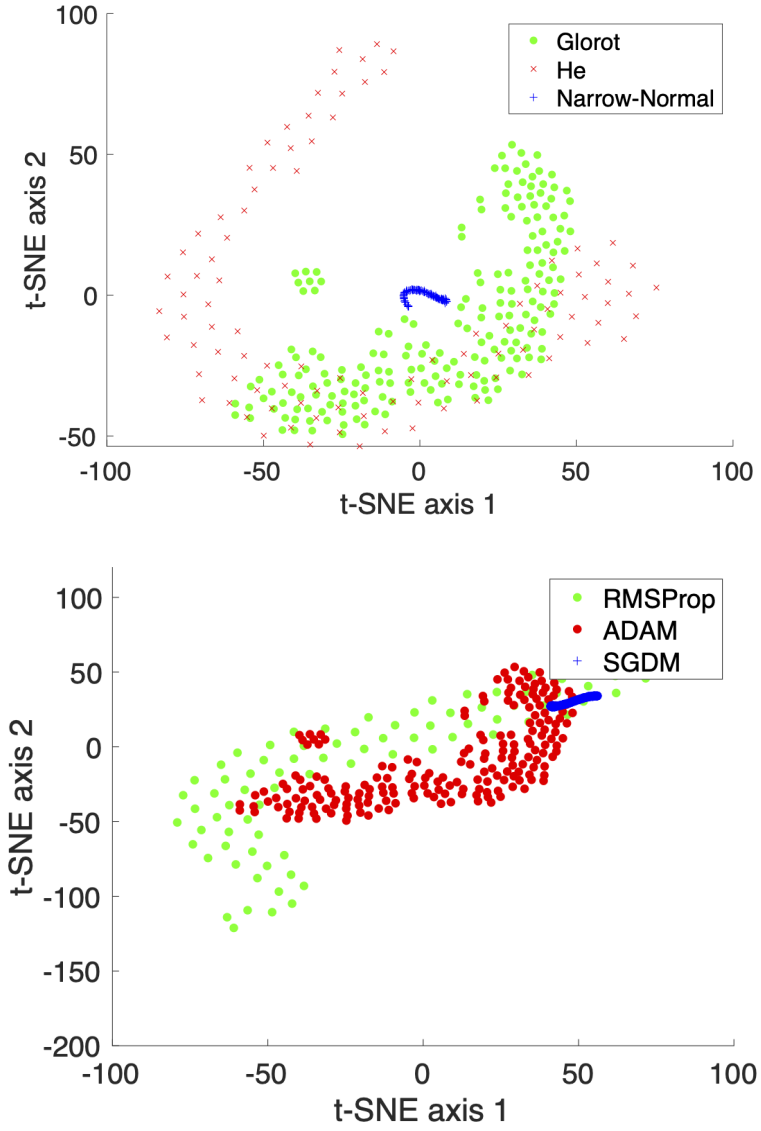

Figure 6.18. The left panel shows the 2D point rendering using t-SNE for different initialization for the same solver, ADAM. Glorot and He have similar functional space, whereas the narrow-normal indicates a very small region of exploration in the function space, explaining the high error observed in Table 6.4. The right panel shows the trajectories of the different optimizer, given same weight initialization and shows the difference in the function space exploration of each optimizer.

On the other hand, the t-SNE trajectory of the narrow-normal initialization which independently samples from the normal distribution with zero mean and a standard deviation of 0.01, thereby not incorporating information from the data, only occupies

a small region in the phase space. For the same initialization, and different solvers, it is observed all the three different solutions start off from the same point but quickly diverge and follow different trajectories, consistent with previously reported observations [110]. This shows that the functions explored by different optimizers are far away and this leads to the divergence and differences in predictions, while functions explored within a single trajectory tend to be much more similar.

6.6.3.1 Sensitivity Analysis

The principles of automatic differentiation [23] that enables the optimization of network weights during the training process can also be used to determine the sensitivities of the model output, in this case the kinematic viscosity to the input parameters. This provides a phenomenological interpretability of neural network mappings [171, 66] but falls short of determining causality as would be apparent from models obtained by symbolic approaches [171, 34]. We compute the sensitivities by calculating gradients of model output with respect to the inputs, on the test dataset. Figure 6.19 shows the plots of the sensitivity study and reveals that the shear stress terms are likely most sensitive and thereby prominent. This is not a surprising result since the Smagorinsky closure terms are indeed dependent on the shear stress. This further builds confidence in the predictions of the data-driven model.

Further analysis of the gradients of the shear stress terms (refer Equation 6.13) in form of plotting the pair-wise joint probability density functions (pdfs) in Figure 6.20 show the relatively minor differences in the distributions between the off-diagonal components of the shear stress tensors. This further proves that no directional preference is awarded in the learning process.

$$Q \equiv \frac{\partial \nu_t}{\partial (S_{ij})} \tag{6.14}$$

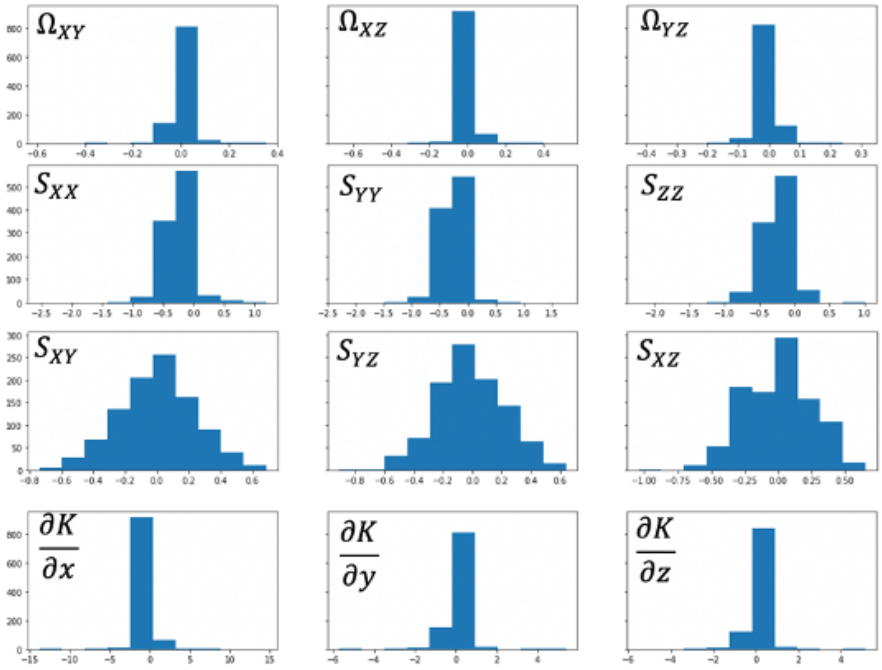

Figure 6.19. The sensitivity study, based on automatic differentiation, by taking gradients of model outputs to inputs show the relative importance of different model inputs. Each plot show histograms of the sensitivity gradients. It appears the shear stress terms are the most sensitive parameters affecting model output predictions

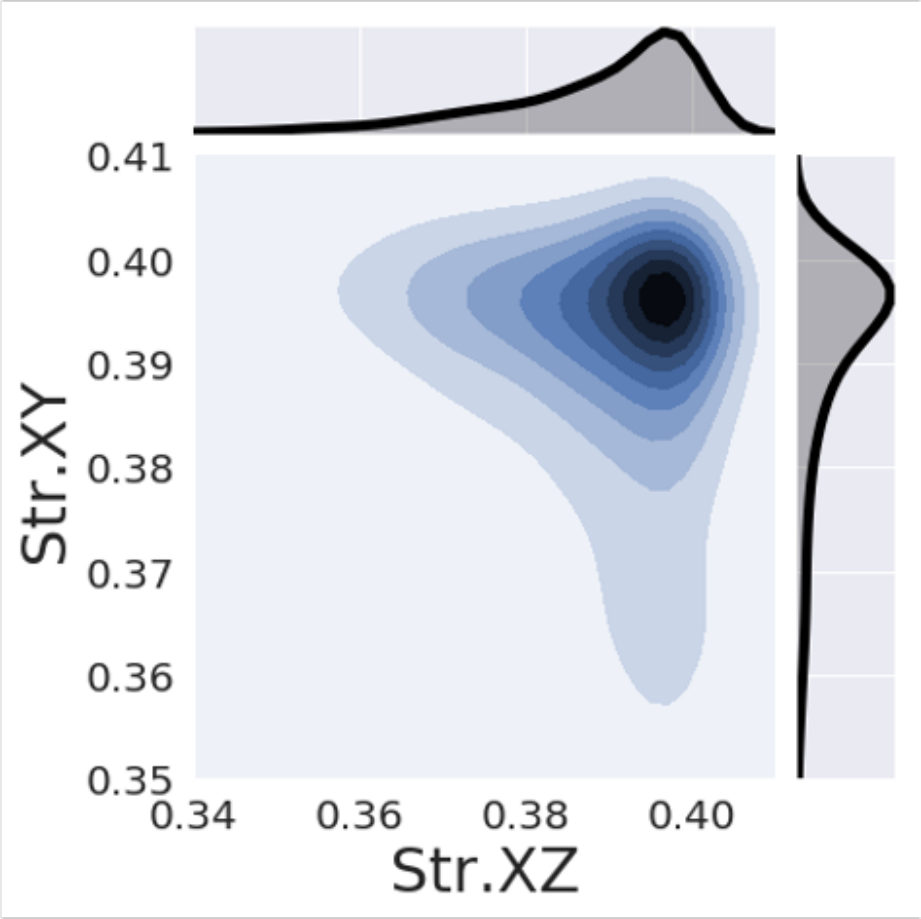

Figure 6.20. The joint pdf estimation of the two most sensitive parameters reveal the similarities in their sensitivities.

6.6.3.2 Accelerating the inference engine

Post network training, the MATLAB-trained neural network is converted into its C++ equivalent using the Codegen toolbox. This C++ code is then used as an inference engine in conjunction with OpenFoam. This step is necessitated as OpenFoam is a C++ based CFD solver. In converting the code using Codegen, we set the inference miniBatch, in other words the number of cells that the inference engine uses at each pass, a-priori. The default is using a miniBatch of 1, implying that for a simulation with N cells, there will be N number of inference engine evaluations at each timestep. This obviously will drastically slow down the process of the inference and is

undesirable. In order to accelerate this we use higher miniBatches such that we now need N/mB number of inference evaluations, where mB is the inference miniBatch and is more than 1.

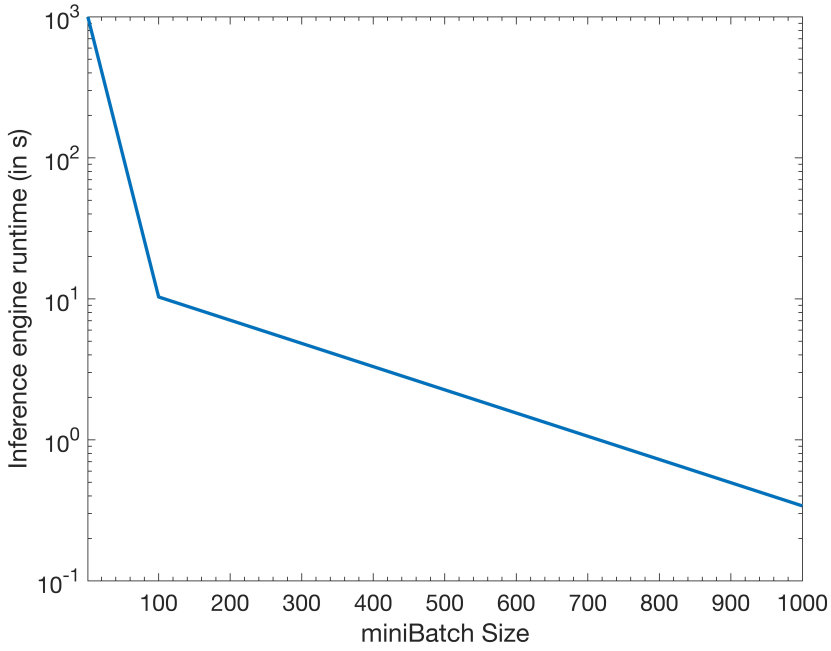

Figure 6.21. Using larger batch sizes results in large savings in the inference engine cost. This has potential to accelerate the CFD inference engine by orders of magnitude

Additionally we show results from the pruning study for an eight-layer feedforward network with 100 neurons in each layer. We use two thresholds to compare the performance of the pruning algorithm. First is a threshold of pruning 50% parameters and the second a threshold of pruning 90% parameters. For the 50% pruned case, this essentially means that after the pruning algorithm is applied the network has 50% less parameters than the original network (and similar deductions for the 90% pruned case). These thresholds although somewhat arbitrary, will demonstrate the possible compression of the network and the effect on the inference performance. The pruning is done on all the network layers, fc_1 to fc_9 in an iterative fashion. Tables 6.4 and 6.5 shows the relative number of parameters pruned in each setting and the

corresponding Mean Squared Error (MSE) measure on the unseen validation dataset, randomly sampled from the original dataset. It follows that for both the pruning thresholds, the network performance (MSE) improves over the baseline, determined by the original network MSE on the same data. This outcome is not entirely surprising, as overparameterization and the propensity of deep neural networks to learn noise for regularization, and better generalization can lead to lower baseline scores [20]. Pruning can often remove the learnt 'noise' which improves performance at the cost of robustness and generalizability.

For scientific surrogate models, while the network inference is key in accelerating computational speed, the most important metric is the accuracy of the model output. This model accuracy therefore gains relevance, as the ML models are coupled to non-linear PDE solvers which makes the overall system extremely sensitive to the accuracy of these data-driven sub-models. The 50% pruned network shows remarkable consistency with the original network characteristics with the added advantage of an at least 5x speedup [refer Table 6.7]. The loss in the saliency score (*saliency* is the gradient of loss) is about an order of magnitude for the 50% threshold, compared to three orders of magnitude for the 90% pruning (refer Table 6.6). This makes the former a stronger candidate for pruning thresholds for the given problem.

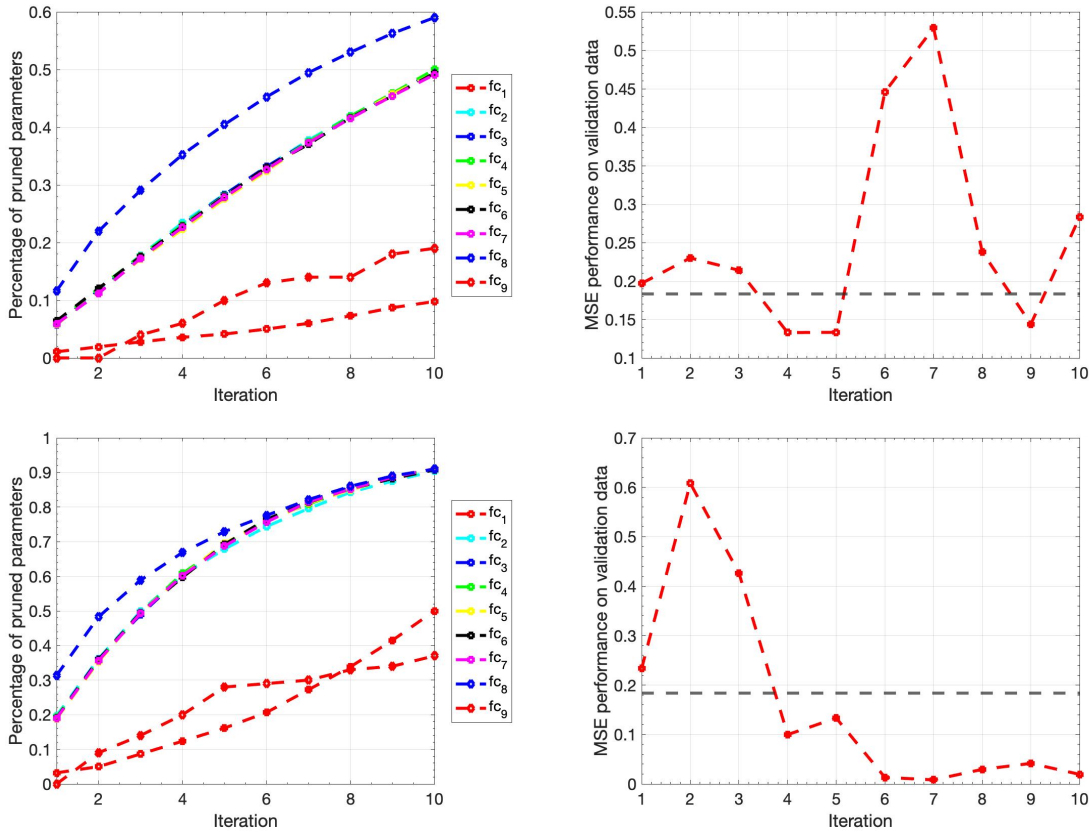

Figure 6.22. [L-R] As the pruning iteration increases, the larger share of parameters are pruned from deeper layers preventing layer collapse. Pruning 50% parameters from the network shows a slight improvement over original MSE (marked in dotted horizontal line) of the full network. Pruning 90% parameters show significant improvement over original MSE baseline. This can be attributed to the removal of the noise in the learning process. The gain in speedup is compensated by the loss in generalizability for these models. However, for many surrogate modeling applications this may be acceptable.

The final state for the layer wise pruning, in Table 6.4, indicates the shallower layers lose fewer parameters while the deeper layers lose more parameters as a percentage of the total per-layer connections. This is consistent with the SynFlow algorithm and essentially ensures there is no layer collapse in the network.

Layer	Total parameters	50% pruned network	90% pruned network
fc_1	1400	137	700
fc_2	10000	4949	9045
fc_3	10000	4956	9099
fc_4	10000	5000	9082
fc_5	10000	4938	9109
fc_6	10000	4938	9074
fc_7	10000	4912	9108
fc_8	10000	5901	9096
fc_9	100	19	37
MSE	0.1874	0.133	0.029

Table 6.4. The total number of pruned out parameters for each threshold (50% and 90%) at the end of the final iteration. This means at the end, the networks have 50% and 10% parameters remaining in comparison to the base network. The MSE performance improves over the baseline (original network).

Apart from the discussions regarding the compressed network's inference performance in terms of validation dataset accuracy and the speed up, Table 6.5 discusses the footprint of the original network compared to the compressed networks in terms of space occupied by the weights and the run time memory. As expected, with a higher compression ratio, the memory occupied by weights and the runtime memory requirements (number of operations used as a proxy for runtime memory estimate) are reduced.

Network	# Weights	Memory (Weights)	# Operations	MSE*
Original	71500	0.27 MB	142199	0.1840
Unpruned	71500	0.14 MB	142199	**0.1680**
50% Pruned	35750	0.14 MB	70699	0.2694
90% Pruned	14300	0.05 MB	27799	**0.0339**

Table 6.5. The effect in terms of runtime memory and space occupied by network weights is inversely proportional to the compression ratio. The MSE is tested on an unseen validation dataset containing 10,000 samples. The **bolded** networks indicate better than baseline performing settings.

The conservation of the saliency score (gradient of loss) is shown in Table 6.6 for both pruning thresholds. The 50% pruned network shows a loss of an order of magnitude compared to three for the higher compression ratio setting.

Epoch/Network	50% pruned	90% pruned
1	$1.09e^8$	$1.09e^8$
2	$1.06e^8$	$8.40e^7$
3	$9.92e^7$	$4.26e^7$
4	$8.80e^7$	$1.69e^7$
5	$7.52e^7$	$6.09e^6$
6	$6.20e^7$	$2.46e^6$
7	$4.98e^7$	$1.13e^6$
8	$3.90e^7$	$5.20e^5$
9	$3.00e^7$	$2.45e^5$
10	$2.27e^7$	$1.10e^5$

Table 6.6. The saliency score for both pruned networks show that the 50% pruning conserves much of the original information.

Among the primary motivation to explore the feasibility of network compression is to improve network inference performance. Table 6.7 shows the performance of the inference engine in a simulation-relevant environment.

Network	SS: 1000	SS: 1000000	Speedup	Overall speedup
Original	0.19 s	1.5 s	1x	3-4x
50 % pruned	0.14 s	1.1 s	1.5x	4.5-6 x
90 % pruned	0.06 s	0.47 s	3x	9-12x

Table 6.7. The inference speed-up post pruning for different data sizes shows up to 5-10x overall improvement in network inference without appreciable loss in generalizability. The results from different samples are shown here to demonstrate consistency in inference throughput. **SS** means sample size.

6.7 Conclusions

Data-driven deep learning models have emerged as a promising alternative to build, calibrate, or replace existing models for complex physical processes including fluid turbulence. In this study, we build a deep learning model for an algebraic closure,

based on Smagorinsky, for a transient, cold flow compressible LES simulation. To effectively model the different regimes of a combustion cycle, data are sampled from various CAD in the training dataset. In order to determine the best parameters, that generalize well to an unseen test dataset, a Bayesian Optimization based autoML method is used. The autoML evaluation reveals the ADAM optimizer and Glorot initialization to be the best performing settings for our dataset. An a-priori testing on an unseen dataset, sampled from the same distribution as the training data, shows the network predictions to be comparable to the ground truth and the errors to the order of 0.03

While regressing the Algebraic closure is a first necessary step in the study, it is not the final goal. The network converted to C++ code using MATLAB Coder is then integrated into OFICE [51]. This LES with data-driven closure are compared to the original Smagorinsky closure predictions for up to 5 CAD. The predicted velocities are comparable, and a quantitative analysis reveal the errors to be in the range of 0.01-0.05%. In the literature, it has been reported that an error accumulation occurs when a data-driven model is coupled to a non-linear PDE solver. We do not see the evidence of that in our limited a-posteriori testing.

In order to further our understanding of the black-box training process, we analyze the evolution of the trained network's weights using cosine similarity. We observe the weights in the consecutive layers, with the same distribution of data are orthogonal to each other. This reveals the complicated training process for a deep neural network. Finally, to develop a phenomenological intuition of the factors important to the machine-learnt network, a sensitivity study is conducted using principles of automatic differentiation by taking the gradients of the model output to its inputs. This reveals the shear stress terms to be the most important/sensitive input feature in the model. This observation makes sense since the canonical formulation of the kinematic viscosity algebraic closure is derived based on shear stress.

Overall, this study provides a pathway to building more complicated data-driven closures, relevant to predictive engine modeling. This also lays the groundwork for Coarse Grained – CFD (CG-CFD) by building surrogate models for estimating local errors due to lower mesh resolutions.

CHAPTER 7

LOCALLY-ENHANCED CFD

7.1 Introduction

Numerical modeling of fluids has a great many scientific and engineering applications. The issue of performing reliable, repeatable and tractable simulations is a long-standing goal for the CFD community. However, often times higher numerical fidelity is associated with an increase in the computational cost, thereby creating challenge for obtaining repeatable, reliable solutions. The necessary spatial and temporal resolution required to accurate model the physics and the entire range of scales is often out of reach for many computational problems. This is due to the fact that fluid motion is inherently multi-scale in nature due to turbulence, and as necessitated by the governing equations of fluid momentum Navier-Stokes. While solving the full Navier Stokes using Direct Numerical Simulation (DNS) results in the most accurate representation of the complicated, non-linear, non-local, multi-scale phenomenon, it is often computationally intractable. Reduced-order or engineering level solutions based on Reynolds Averaged Navier-Stokes (RANS) and Large Eddy Simulations (LES) alleviate this issue by resolving the larger integral length scales and modelling the smaller unresolved scales by introducing a linear operator to the Navier-Stokes equation to reduce the simulation complexity. This divides the governing equation into resolved and unresolved terms. Transport equations are solved for the resolved terms, while the unresolved terms are modeled as a function of the resolved field. These models however suffer from the difficulty of turbulence closure, a topic we discussed in detail in the previous chapter. These turbulence closure models are often

times empirical and perform well for problems in which the resolved scales drive the dominant transport processes [56]. Duraiswamy et al. [56] noted that for an ICE simulation, since combustion occurs exclusively at the small scales, the influence of chemical reactions and unresolved turbulence on the large-scale flow evolution requires careful consideration. The left panel in figure 7.1 shows a typical energy spectrum curve and the regions of operability for various full-order and reduced-order modeling approaches.

It is also well-understood that implicit filtering in LES is directly related to the spatial discretization used in the simulation [125]. This means that smaller the discretization, or finer the mesh, the simulation is expected to resolve more of the scales directly and therefore yield more accurate results. However for practical engineering applications the cost of repeatedly performing these fine-grid high-fidelity simulations for design space explorations soon outweighs the benefits. It is in this area, that we hope to make a modest contribution in enhancing the already available coarse mesh engineering model solution, using machine learning. In short, our goal is to help make the cost-accuracy trade-off favorable over current empirical methods by recovering the lost resolved field information missing from coarser mesh simulations.

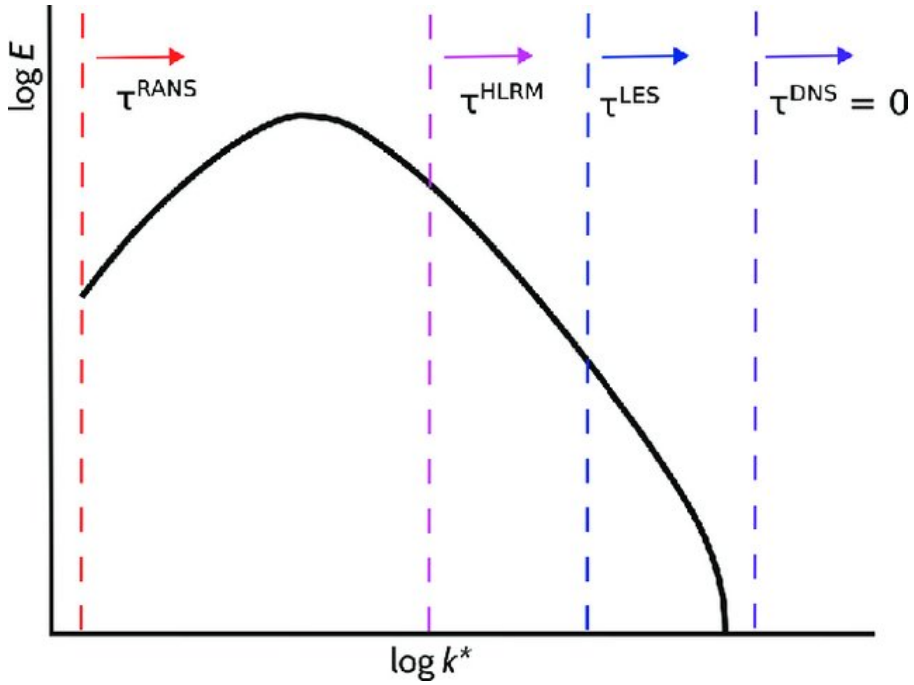

Figure 7.1. A typical energy spectrum curve with regions of operability for various modeling techniques. Image adopted from [18].

Coarsening the numerical grid, or decreasing the spatial discretization results in loss of information due to resolving fewer scales. To recover the lost information and correct the fluid flow variables, as a result of the numerical error as well as sub-grid error, a machine-learned surrogate model is proposed to be developed. Given sufficiently fine-mesh simulations, a surrogate model can be trained to predict the Coarse Grained-CFD local errors as a function of the coarse-grid local flow features. In this context, local is defined as in the vicinity of the cell of interest. This trained model can then be used in an inverse mapping scenario to a-priori predict the expected correction to be applied to the coarse-grained variables.

Multi-scale problems have often used coarse-graining approaches in the past. This is especially true for computational biology applications such as in Molecular Dynamics (MD) [30, 185]. The range of interacting scales span magnitudes in such applications, hence fully resolving all relevant information is often not computationally

tractable, similar to observations from fluid modeling. In essence, this framework involves an integrated, iterative approach to couple information from different scales (atomistic scales to meso scales). The primary steps, which coincide with key areas of method development, include developing first-pass coarse-grained models guided by experimental results, performing numerous large-scale coarse-grained simulations, identifying important interactions that drive emergent behaviors, and finally reconnecting to the smaller scale by performing high-fidelity simulations guided by the coarse-grained results. The coarse-grained modeling can then be extended and refined, with the entire loop repeated iteratively if necessary.

However the translation of such a modeling paradigm to CFD is not straightforward, due to the unique challenges. Primarily the nature of scale interactions is non-linear in fluid modeling compared to the linear interactions in MD. The non-linearity in the Navier-Stokes equations, governing fluid momentum, is introduced by the advection term, and the range of scales are limited by the viscosity term. In addition to that, the non-local nature of fluid turbulence, within limits of incompressibility, adds to the challenge in scale interactions.

The term *Super-resolution* or *coarse graining* involves the inference of a high-resolution information from low-resolution measurements, leveraging the statistical structure of high-resolution training data [33]. In the recent past, Machine Learning has played an important role in recovering lost information and enhancing solution quality especially from an image processing perspective [115]. This technique was first developed for applications related to deep image processing [115]. The super-resolution (SR) reconstruction of turbulent flows from low-resolution coarse flow field data has seen recent interest in the fluid modeling community [97, 106, 231, 29]. While a variety of methods have been used in enhancing the low-res input, most noticeably Generative Adverserial Networks (GANs) [70] have been prominent. Generative Adversarial Networks, or GANs for short, are an approach to generative modeling using

deep learning methods, such as convolutional neural networks. A full review of GANs can be found in the literature [74]. The GAN model architecture involves two sub-models: a generator model for generating new examples and a discriminator model for classifying whether generated examples are real, from the domain, or fake, generated by the generator model. A schematic of a typical GAN model is shown below

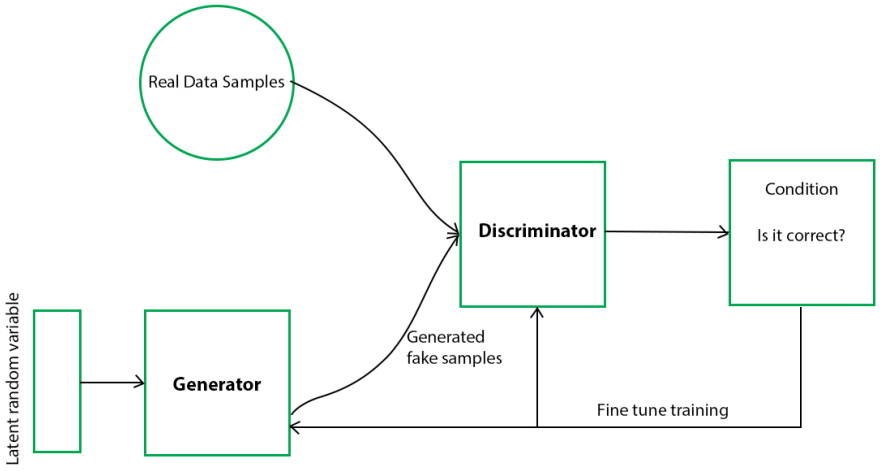

Figure 7.2. Schematic of a simple GAN model. Adopted from [70].

However, the application of these methods for problems in fluids suffer from some severe shortcomings. These include:

- GANs are notoriously hard to train, especially with the presence of two competing networks (such as the discriminator and generator networks). While there exists solutions to make the training process more stable, there are few convergence guarantees.

- Since we deal with non-linear dynamical systems in our study, enforcing the physical and numerical constraints are critical. To accomplish this task on a GAN is non-trivial for complicated 3D turbulence problems.

- The high inference cost at run time is another cause of concern as noted by Bode et al. [29].
- The SR paradigm has been so far applied to simplified Cartesian grids, and simulation settings such as isotropic turbulence [29] and 2D Rayleigh-Bernard [97] and the path forward to applying it for engineering problems is not clear.
- The global approach for SR is not compatible to locally variables flows.

7.1.1 Locally enhanced CFD

The principal aim of this study is to make the cost-accuracy trade-off favorable. It is in the same vein that Kochkov et al. [108] demonstrated acceleration of LES simulations using ML based enhancement for the missing information in coarser meshes. Previous work in this area [108, 239] showed the ML models have the ability to effectively super-resolve the missing information for applications ranging to 2D turbulence [108] and tracers in climate models [31, 239]. Our work is an extension of the study by Kochkov et al. [108] in using machine learning to accelerate CFD simulations. The authors [108] used the concept of learned-corrections or nudge to the coarse-grid simulations and this is similar to what we propose. Apart from some key differences in the implementation philosophy, a critical improvement over the previous work includes extending this approach to engineering relevant problems and to full 3D simulations, compared to canonical 2D simulations proposed in the original study.

Since we propose to use local cell level information to build functional relationships between coarse meshes and fine-grid output, our method is aptly called *local-enhanced* CFD. The locally enhanced velocity within each cell, would then have the following functional form:

$$\mathbf{u_e} = \mathbf{u_c} + \mathrm{LC}(\mathbf{u_c}) \tag{7.1}$$

where $\mathbf{u_e}$ is the enhanced velocity, $\mathbf{u_c}$ is the coarse grid velocity, and LC is the learned correction provided by the machine learning algorithm, during inference time.

For a model physical system governed by a set of non-linear equations, the relationship between the true solution, S_f, from a fine mesh simulation and the coarse mesh predictions can be expressed as:

$$S_f = S_{LF}((\vec{\mathbf{u}_{\mathbf{LF}}}, \delta_{LF})) + \varepsilon \tag{7.2}$$

where S_{LF} represents the output of the low fidelity simulation, $\mathbf{u_{LF}}$, and δ_{LF} represent the model variables - in our case fluid velocity, and the coarse mesh size used in the low fidelity simulation, and ε the simulation error (lost information) due to sub-grid model error arising due to use of simplistic physical models coupling the resolved and unresolved terms, and the mesh error. Functionally, it can be represented as:

$$\varepsilon = \varepsilon_\delta + \varepsilon_{sgs} = \mathbf{u_{f\to\Delta}} - \mathbf{u_c} \tag{7.3}$$

where subscripts δ and sgs are errors due to mesh discretization and the sub-grid errors respectively. The term $\mathbf{u_{f\to\Delta}}$ is the fine to coarse mapped velocity, and $\mathbf{u_c}$ is the coarse mesh velocity. The mesh induced error is a direct result of the spatial discretization and the sub-grid error due to not resolving enough scales and using simplified models (as a function of the resolved scales) to predict this behavior. Both types of error affect the numerical solution of the CFD simulations, which can often lead to non-negligible error in the outputs of interest. More recently, error surrogate models based on machine learning techniques have received much attention, largely because of their non-intrusive nature and fast on-line evaluations. Several contributions have been made in error modeling for parameterized reduced-order models (ROM) [53, 147], and the ideas have been extended to estimate discretization-induced errors [59]. This locally enhanced approach is an extension of this line of investiga-

tion. Efforts have also been devoted to predicting the errors in flow solutions and the outputs of interest obtained on coarse computational meshes [180, 79], and the models have been used to guide the selection of a set of a priori meshes [79, 19] . However, all of these studies either are limited to a-priori testing or for a rare few that integrate the neural networks with a CFD code, are used for problems such as decaying turbulence in a box or 2D Rayleigh-Bernard.

When comparing the fine- and coarse-grid data, the number of cells in both grids is not the same. Thus, to compute the local grid-induced error, it is necessary to perform mapping of the fine-grid data ϕ_f onto the coarse grid Δ. In other words, ϕ_f is replaced by $\phi_{f\to\Delta}$ which is the fine-grid field of ϕ mapped on a grid whose cell length is Δ. This mapping, or interpolation, constitutes a source of error as some details of the flow field profile are lost due to interpolation. This mapped field is still expected to be more accurate than the field computed by a coarse grid. The error at the local cell level therefore can be defined in Equation 7.3.

After generating an error database using a large corpus of coarse mesh simulations and its corresponding error due to the mapped fields, a machine learning model is ready to be trained. Once appropriately trained, the data-driven surrogate model will then be used to locally enhance the solution using inverse mapping techniques.

The basic assumptions for the application of the coarse grained approach is that the coarse mesh simulation is able to capture/resolve the basic flow features. It would be inconceivable to use ultra coarse representations of the physics such that any important detail is not resolved by the coarse mesh, and thereby extrapolating the mapping abilities for the machine learning algorithm. It has been discussed widely in the literature that most machine learning algorithms extrapolate poorly [79] for most scenarios without explicit physics encoding [202].

Our goal is to produce solutions to the Navier-Stokes equations with diminished sensitivity to mesh resolution. In particular, we will focus on the velocity field since

for the constant density Navier-Stokes the velocity field and its derivatives sufficiently determines the pressure field. For low Mach number flow, such as in consideration here, the initial pressure and density fields are neglected as part of the feature selection.

This approach has several advantages:

- Incorporating localized interactions (in the cell neighborhood) make the solution more accurate.
- The boundary effects are expected to be resolve better, compared to globally enhanced GANs without physical constraints.
- Training and iterating over best network design is manageable.
- Inference cost would be cheaper compared to expensive GANs.
- This approach is problem specific, and therefore it would possibly generalize better within the set of operating conditions as well as grid configurations.

7.2 Methods

7.2.1 Modified Governing equations

Since design space explorations using Direct Numerical Simulations are computationally intractable for many engineering relevant flows, we use very fine mesh LES as our ground truth data. For a review of the LES governing equations, please review Chapter 6. Functionally, the mass conservation and the governing equations of momentum appear in the following form:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 \tag{7.4}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{\nabla P}{\rho} + \nu \nabla^2 \mathbf{u} \tag{7.5}$$

where $\mathbf{u}$ is the fluid velocity vector, P is the fluid pressure, ρ is the fluid density, ν is the kinematic viscosity, and ∇^2 is the Laplacian operator.

The machine-learned nudge (or 'bump') is integrated into the Navier-Stokes governing equations by adding a source term, $\mathbf{S}$. The modified governing equations functionally is shown as below.

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{\nabla P}{\rho} + \nu \nabla^2 \mathbf{u} + \frac{\mathbf{S}}{\tau} \tag{7.6}$$

The term τ is used as a time-scale (or relaxation) factor. In other words, it is used to relax the amount of extra information (machine-learnt nudge) that is added to the system of governing equations. This is primarily done to ensure numerical stability for the non-linear PDE solution. Add too much source, and the mass conservation has a hard time keeping the solution stable and converging. Add too little source, and the solution barely changes. This is therefore a hyperparameter in the modeling setup and we empirically investigate effects of different relaxation factors. A more scientific intuition or explanation is therefore warranted, and is a subject of future work. Once the sources are determined for a given cell, we propose to integrate them in two different ways.

- Explicit treatment

  - In this method the machine learning model predictions are solved explicitly. This functionally would appear as Eq 7.8 above. While this is expected to be numerically less stable due to stability restrictions, it is the easiest to implement numerically. The source term in this scenario is expected to be a vector with

the dimensions of velocity. For better numerical stability, this necessitates using robust solvers such as Preconditioned Bi-Conjugate Gradient solvers.

- Implicit treatment

  - In this method, the model predictions are added to the diagonal of the velocity vector matrix. Per limitations in OpenFOAM [95], it can only be implemented as a scalar, C. It is more stable numerically, however it limits the extent of functionality especially for shear driven flows. Functionally the modified Navier-Stokes equations is as follows

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{\nabla P}{\rho} + \nu \nabla^2 \mathbf{u} + C \cdot \mathbf{u} \tag{7.7}$$

  here $C = \frac{\varepsilon}{\mathbf{u}.\tau}$

Furthermore, the locally enhanced CFD method has two primary sources of error. The first is the use of CFD approximation itself. We use the very fine mesh LES as ground truth data, that itself is a source of error due to still not being the most accurate available solution i.e. DNS. We are limited by the computational tractability as well as the motivation to enhance engineering level design space explorations. The second is the interpolated fine to coarse mesh mapping. To compare error at each cell, the total cell count between the ground truth and coarse mesh has to be same. So there needs to be some sort of mapping between the finer mesh (larger cell count) ground truth data and the coarser mesh (lower cell count) dependent variables. This is achieved by using OpenFOAM's [95] in-built *mapFields* functionality.

7.2.2 Step-by-step implementation details

Our choice of machine learning algorithm for the local enhancement are the neural networks. We have previously discussed the details of the algorithm in Chapter 6. The implementation level details of the locally-enhanced approach can be divided

into two separate parts, details of which will be discussed separately. The first can be classified as the machine learning training, or training part as referred to from hereon. The second is the machine learning inference, or inference part.

7.2.2.1 Training part

The training can be divided into three steps

- Step 1: Run a fine-mesh simulation. This simulation typically consists of very large number of cells and therefore very accurate. For examples, where there are experimental validations available, these have been compared to experimental observations.

- Step 2: Run a variety of simulations with different coarser mesh configurations. This step essential creates a dataset for the model to learn with as inputs.

- Step 3: Use OpenFoam's [95] in-built mapping functionalities to map the fine data generated in Step 1 onto the coarser stencil from Step 2. This would be our ground truth the model aspires to achieve. In machine learning jargon, this is called the target data.

Finally we will have a local cell-level nudge needed for the coarser stencil simulations in order to retrieve the lost information due to discretization error and sub-grid error.

7.2.2.2 Inference part

Once the neural network model has been sufficiently trained on a large corpus of data, it is ready to be used as an anti-function or inverse mapping function. In mathematical terms, for data belonging to sets $\mathbf{S}$ and $\mathbf{T}$, a mapping can be defined as:

$$f : \mathbf{S} \rightarrow \mathbf{T} \tag{7.8}$$

and its inverse will be defined as

$$f^{-1} \in \mathbf{T} \rightarrow \mathbf{S} \tag{7.9}$$

In case of the locally enhanced flows, **S** is the cell-wise enhancement and the **T** is the coarse mesh simulation.

7.2.2.3 Field mapping strategy

We use the *mapFields* utility within OpenFoam to map the fine mesh data onto the coarser mesh. This method maps volume fields from one mesh to another, reading and interpolating all fields present in the time directory of both cases [87]. There are three available methods - *mapNearest*, *cellInterpolate*, and *cellInterpolateWeight* and Hopken et al. [87] provides a review of the available tools.

7.2.3 Inputs for ML model

The Helmholtz decomposition theorem states that every smooth vector field **u**, defined everywhere in space and vanishing at infinity together with its first derivatives, can be decomposed into a rotational part Ω and an irrotational part $\Delta\phi$ [211]. In simpler words, the motion of a fluid element can be defined in terms of three fundamental parts [211].

- A pure translation motion
- A pure strain along the principal axes
- Rotation rate, corresponding to vorticity

Therefore in choosing our inputs we prioritize using the strain rate tensor and the rotation rate tensor. This is particularly useful to preserve Galilean and rotational

in-variance, to prevent any directional preferences that the model may learn. We appropriately non-dimensionalize tensors, as well as use non-dimensional quantities such as cell Reynolds number and wall distance as inputs. The non-dimensionalization factors were inspired from [242] and tabulated in Table 7.1.

Quantity	**Raw Input**	**Normalization Factor**
Velocity gradient	$\nabla \mathbf{u}$	$\frac{\sqrt{k}}{\delta_c}$
Strain-rate tensor	S	$\frac{\sqrt{k}}{\delta_c}$
Rotation-rate tensor	Ω	$\frac{\sqrt{k}}{\delta_c}$
Wall distance	$\mathbf{Y}$	δ_c

Table 7.1. Non-dimensionalization of Input features

where R_c is the cell Reynolds number, measured as $R_c \equiv \frac{\rho u \delta_c}{\mu}$ where δ_c is the cube root of the cell volume, S is the shear-rate tensor, Ω is the rotation-rate tensor and $\mathbf{Y}$ is the non-dimensional measure of the wall distance. The non-dimensionalization factors for each term are discussed in Table 7.1. In addition to non-dimensionalizing the data, to scale the features of the flow, the data is further normalized to scale the range between 0 and 1 to follow machine learning model training practices.

The strain-rate and rotation-rate tensors are derived quantities from velocity gradient and using them altogether might indicate unnecessary repeating information for the model. However, in early testing phase for this framework, models trained on all three terms outperformed models just based on the velocity gradient. This empirical observation led to the retention of these set of inputs for the rest of this study. Investigating sensitivity of the model learning on inputs, is a subject of future study.

In addition to normalizing the data, it is critically important to non-dimensionalize each individual component as well. The importance of the non-dimensionalization step is shown in Figure 7.3. Each individual panel consists of data taken across multiple simulations with different mesh configurations - coarser to finer from left to

right. The panels on the left side are datasets that are not normalized, whereas the panel on the right side are datasets that have been appropriately normalized (per Table 7.1. Looking at a particular field, for example Str_0 it is evident that as the mesh becomes finer the range of learning for the model increases correspondingly. For finer mesh configurations, the datasets are order of magnitude larger over the smallest mesh configurations. Using this dataset would limit the models ability to extrapolate and learn the effects due to discretization. On the other hand, upon normalization $StrPrime_0$ retains the essential physics yet the range of data is actually similar for all mesh sizes. This is easier for the model to learn and is therefore recommended.

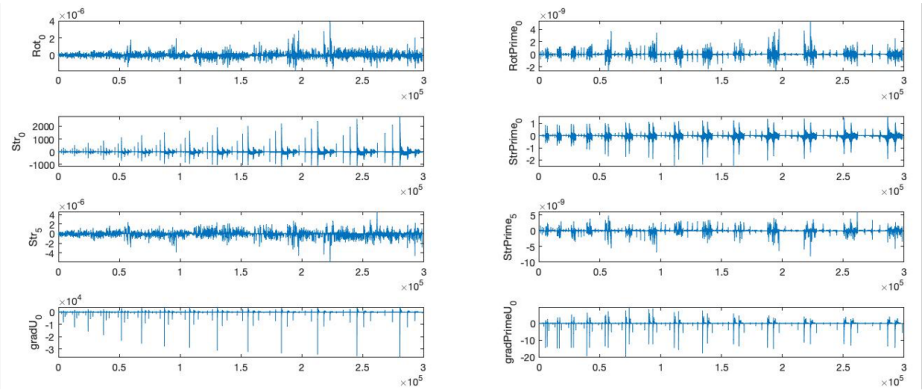

Figure 7.3. Effects of normalizing inputs on a toy engine simulation dataset. Indicates normalization helps in model's ability to learn.

7.2.4 Machine Learning model for mapping

The fundamental notion behind developing a machine learnt surrogate model to predict local enhancement, is to build a functional map between a set of inputs and targets. Mathematically, it can be defined as $out = \mathbf{F}(in)$. Once sufficiently trained on a large dataset, the goal is to approximate the inverse map $\mathbf{F^{-1}}$ for the expected enhancement for a different set of inputs. Compared to classical programming, deep learning algorithms are exploiting the underlying rules that drive these functional maps. A simple schematic adopted from [42] is shown in Figure 7.4.

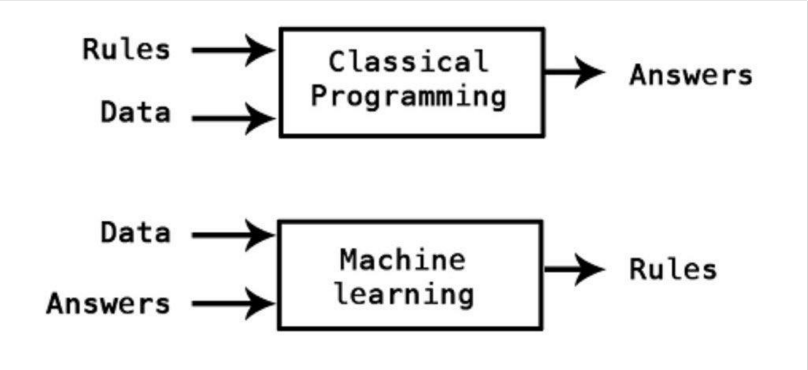

Figure 7.4. Functional difference between classical programming and deep learning. Adopted from [42]

Inspired by the famous words of Abraham Maslow, *"If the only tool you have is a hammer, you tend to treat everything as if it were a nail"* the important question that needs to be answered, is if the Machine Learning algorithm shares similar inductive biases to the data and therefore actually reliable to perform this mapping task. Mitchell et al. [139] argued that neural networks have inductive biases. To design and train an appropriate network, it is therefore important to look at the inductive biases of the problem itself and align it with the algorithm used. Since in this study, our goal is to use a functional map without any spatial or temporal information encoded, a simple feed forward network would be appropriate. Feed-forward networks are known to have comparatively larger number of parameters in contrast to convolutional networks. However, unlike applications in computer vision where models are deployed on edge devices (ex. FPGAs), our choice of network is not limited to reducing number of parameters due to memory limitations as CFD is generally run on large HPCs. It is however dependent on reducing number of total operations for a faster inference, when the network is integrated into OpenFoam. Therefore our chosen network architecture is an simple feed-forward dense neural network that has similar inductive biases of our problem of interest – i.e. to determine functional relationship between

parameters to predict a series of outputs – and lower number of operations compared to a spatial relationship aware Convolutional Neural Network (CNN).

7.2.4.1 Customized Loss functions

Engineering-relevant CFD simulations are inherently transient and often times this leads to presence of outliers in the input data. This is especially true for scenarios involving moving geometries as well as complex and intermittent physics such as combustion. While in machine learning there are different best practices to deal with such outlier data, often times it involves ignoring them as it leads to poor training abilities for the model. Using this framework for building a surrogate model might lead to loss of important physics and therefore lead to a degrading performance of the model itself. One method to alleviate this is to use a customized loss function. Compared to the mean-squared-error or L2 loss, that fits well to mean behavior, a mean-absolute-error or L1 loss, tends to fit the outliers better. Our proposed loss makes use of both of these separate entities in a weighted fashion. The weighting between the loses is based on data distribution of outliers.

$$loss = 0.7 * \mathrm{L2} + 0.3 * \mathrm{L1} \tag{7.10}$$

7.2.5 Quantitative metrics

In addition to qualitative metrics to measure performance we define a quantitative criterion to measure success for the locally enhanced approach. The first is defined as the *Cell volume weighted L2 norm* defined as:

$$L2_{mapped} = \sum \delta v * (\mathbf{U}^{+} - \mathbf{U})^2 \tag{7.11}$$

$$L2_{net} = \sum \delta v * (\mathbf{U}^{*} - \mathbf{U}^{+})^2 \tag{7.12}$$

here δv is the cell volume, $\mathbf{U}^+$ is the modified velocity due to the mapping, $\mathbf{U}^*$ is the velocity predicted by the network and $\mathbf{U}$ is the coarse mesh velocity. We choose to focus on the velocity error as it is the metric we use to locally enhance the coarse mesh simulation. A lower L2 norm metric would establish the improvement in accuracy of the enhanced result compared to the coarse mesh simulation.

7.3 Results and Discussions

In this section the results from the passive scalar study are presented first. The goal of the passive scalar study is to identify the best field mapping/interpolation strategy for the fine mesh to the coarse mesh stencil. The best mapping strategy is then proposed to be used for each subsequent OpenFOAM test study. For each OpenFOAM study, the training data includes simulations carried out with multiple Re numbers. The Re numbers are varied by modulating the dominant direction velocity. This is done to extend the model's learning ability to include effects of the (non-dimensionalized) large scale flow features. Each Re number is further studied for multiple coarse mesh resolutions. The different coarse meshes are characterized by a reduction factor. The reduction factor (RF) is defined as the ratio of the fine mesh cells to the coarse mesh, so in effect coarser meshes have higher RF. This enables the model to learn the discretization error as a function of the large scale flow characteristics. Instead of using the cell-wise error as the target the neural networks have been trained on the mapped field velocity directly as this provides better stability to the learning process. During inference time however, cell wise error can easily be retrieved from the coarse mesh prediction and the inference from the network. This indirect method of estimating error is found to be more stable numerically.

7.3.1 Passive scalar study

One of the two main sources of error for this method, is the interpolation of fields due to mapping of fine mesh results onto the coarser stencil. It is therefore pertinent to invest energy into characterizing the available mapping methods in OpenFOAM. Since spatio-temporal transported quantities such as fluid velocity are difficult to contrast across different mapping algorithms as that the ground truth data (mapped fields) is an interpolated 'estimate'. Therefore a simple study involving a cell-center based passive scalar in a 2D bounded box is proposed. This passive scalar is not advected and takes on a constant value as a function of the local cell center co-ordinates (more in Eq 7.17). This problem setup is appropriate since the mapped field interpolation error is primarily dependent on the discretization/mesh characteristics.

The fine mesh scalar field is predicted at a higher resolution and the mapping estimates this field at a coarser (less frequent) resolution. However, the coarse mesh predictions at these locations are still highly accurate since this field is dependent only on the spatial location of the cells, and there's no advection in the passive scalar. Therefore, the best mapping algorithm will have the least L2-error when compared with the coarse mesh results. This idea at a high-level is demonstrated on a 1D problem in figure 7.5. In the case of the finer mesh with a higher spatial resolution the observations are recorded at frequent intervals. This is represented as the red curve in the adjoining figure. For the mapped/interpolated field on a coarser mesh, the fine mesh field is then interpolate on a subset of coarse mesh points (shown as vertical dashed lines on the spatial (x) axis). The best interpolation algorithm will have the least error between the fine mesh representation and the mapped/interpolated quantity. This in figure 7.5 is represented by the green 'x'. To quantitatively estimate this we use the, previously introduced, L2-error norm as a metric. The scalar field (F) is mathematically formulated as below.

$$\mathrm{F} = sin(X_c) + cos(X_c) \tag{7.13}$$

where X_c is the x-cell center location and F is the scalar field that takes on a constant value based on the x- cell center location.

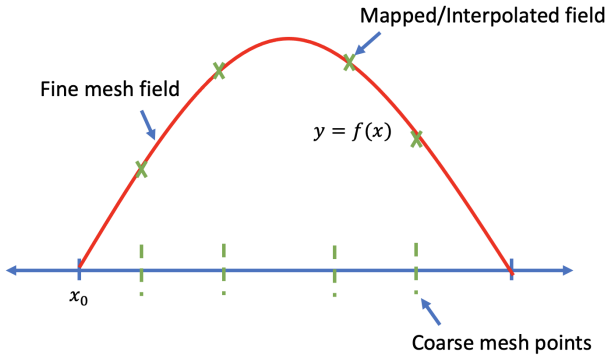

Figure 7.5. One-dimensional schematic of the passive scalar study

In Figure 7.6 shows the passive scalar field (here noted as tracer in the legend axis).

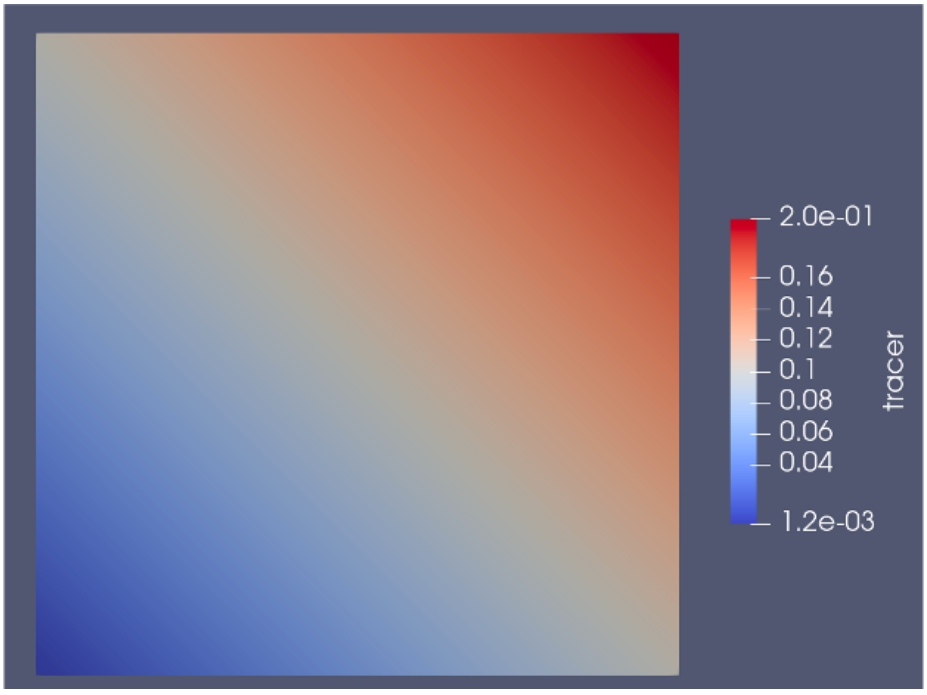

Figure 7.6. The cell center based passive scalar field in a 2D bounded box

Case	L2-norm
coarse - mapNearest	0.020
coarse - cellInterpolate	0.0143
coarse - cellInterpolateWeight	$3.2E^{-6}$

Table 7.2. The L2-norm of tracer for the different in-built OpenFoam mapping strategies shows that the cellInterpolateWeight method yields the least error and therefore recovers most of the information

The table 7.2 reports the difference between the mapped field and the coarse mesh. It shows that the *cellInterpolateWeight* method has the least reproducibility error and therefore used for the rest of the study.

7.3.2 lid driven cavity study

7.3.2.1 Setup

The lid-driven cavity is a well-known benchmark problem for viscous incompressible, laminar fluid flow. A cavity driven flow is a simple simulation environment involving isothermal, incompressible flow in a two-dimensional square domain. The geometry is shown in Figure 7.7 in which all the boundaries of the square are walls. The top wall moves in the x-direction at a speed of 1 m/s while the other three are stationary. The lower left corner has a reference static pressure of 0. The flow is laminar and will be solved on a uniform mesh using the icoFoam solver within the OpenFoam code. It is the simplest possible problem in OpenFoam and will be used to develop the end-to-end workflow.

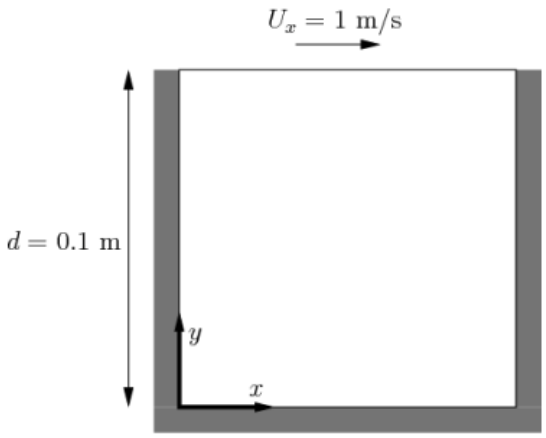

Figure 7.7. Schematic of a typical cavity driven flow. Source [95].

7.3.2.2 a-priori results

While the physics might not be very complicated and it is infact a simple problem from the perspective of numerical CFD, it does provide the opportunity for faster and iterative end-to-end workflow development. That remains the main focus for this particular study. For this test case, three Re numbers were simulated - 100, 150, and 200. And for each simulated Re, six different coarse mesh configurations were studied. This study involves reduction factors ranging from 1.2 to 8, for a fine mesh cell total of 400. The total number of cells in training were about 40000. Post the Bayesian optimization training step, a network with 3 layers and 23 neurons in each layer was determined to provide the best generalization on a validation dataset. This is somewhat expected as the physics in the problem is simple (laminar flow), and the datasets not very large. The a-priori results are shown in Figure 7.8 and the network R^2 is 0.91, which is generally considered a good fit to the data.

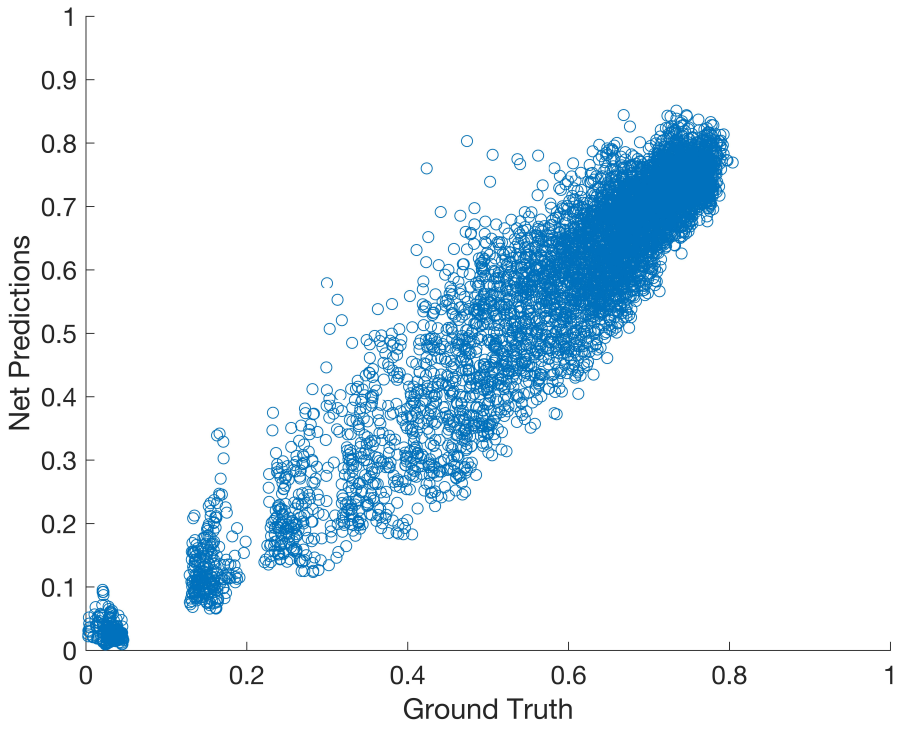

Figure 7.8. A-priori testing of network performance against ground truth data, with an R^2 of 0.91.

7.3.2.3 a-posteriori results

The best network's weights and biases were integrated into the icoFoam solver with a miniBatchSize of 50 (the minimum coarse mesh cell total, reduction factor of 8). The inference engine is used to predict the velocity magnitude at each cell level as a function of the coarse mesh inputs. This inference prediction is then used to calculate the 'nudge' needed at each cell level to recover the missing information. Typically, as the mesh coarsens, the magnitude of the 'nudge' increases as there is more information to recover. The network predictions for the velocity magnitude is shown in Figure 7.9 and the scatter plot comparison of the individual components and the kinetic energy for the fine to mapped coarse data and the learned correction is shown in Figure 7.10

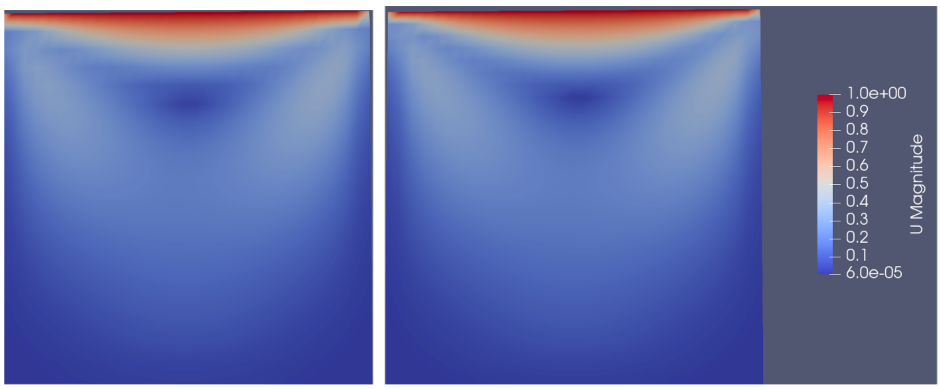

Figure 7.9. Velocity magnitudes for the coarse mesh (left) and the enhanced CFD mesh (right). Table 7.3 quantifies the performance further. Both the results are for a reduction factor of 2.

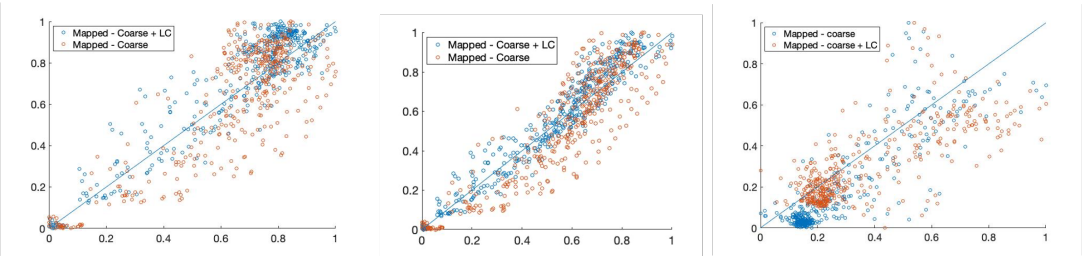

Figure 7.10. The scatter plots for comparisons between the mapped ground truth data and the learned correction enhanced data shows closeness in the predictions for a subset of the cases. Table 7.3 quantifies the performance further. All results shown here are for a reduction factor of 2.

Since the runtime for these simulations are of the order of 1-3 seconds depending on the coarse mesh resolution, the compute-accuracy tradeoff plots are not shown. Table 7.3 reports on the L2 norm of the network predictions, compared to the ground truth data for a reduction factor of 2.

Case	UMag
Mapped - Coarse	0.001213
Mapped - Coarse + LC	0.000667
% improvement	45.01

Table 7.3. The velocity magnitude L2 Norm between mapped/coarse and mapped/net simulations

7.3.3 Turbulent channel flow

7.3.3.1 Setup

The next set of simulations proposed to be studied is the turbulent channel flow (as shown in Figure 7.11). This involves a step up in the simulation complexity for the mapping algorithm, and therefore an important test for the applicability of the algorithm.

The turbulent channel flow is a flow between two infinite parallel plates driven by a pressure gradient. The flow is fully defined by the friction velocity-based Reynolds number $\mathrm{Re}_\tau = u_\tau \delta / \nu$, where u_τ is the friction velocity, ν is the kinematic viscosity, and δ is the half-width of the channel. The computational cost of the simulation grows with Re_τ, therefore in this tutorial we will use the lowest Reynolds number for which DNS data is available, namely Re_τ=395. This will allow using a mesh fine enough to resolve a big part of the turbulent structures present in the flow, yet small enough for the case to be computed in a reasonable time on a single workstation. The big challenge in the LES simulation is that in the near-wall region of a turbulent boundary layer – the necessary resolution required of a high quality LES renders such simulations prohibitively expensive unless a high degree of empiricism is introduced into the (sub-grid) modeling process [56]. Then the challenge for the machine learning algorithm is to not only learn the mean flow behavior but the near wall behavior accurately. To simulate an infinite domain, periodic boundary conditions are commonly applied in the stream- and spanwise directions. The pressure gradient is then introduced via an extra forcing term in the momentum equations.

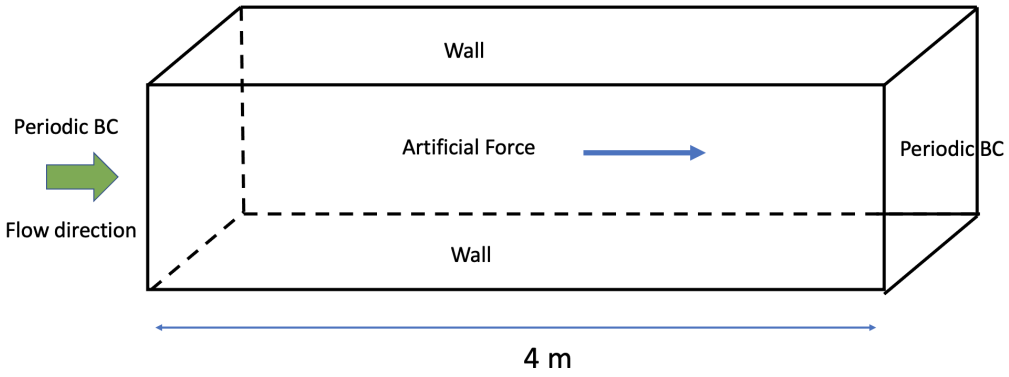

Figure 7.11. The schematic of the turbulent channel flow study

The turbulent channel flow is a statistically-developing internal flow through parallel smooth walls. The x-axis is the mean flow direction, the y-axis is the wall normal and the z-axis is the spanwise direction with statistically homogeneous flow with periodic boundaries. In order to keep the flow constantly going in the channel, and not letting the viscous effects of the walls slow it down, a body force is applied in the mean flow (x) direction. Therefore most of the contribution to the velocity field is in the x-direction, with diminishing contributions from the other two axes. This means that a neural network trained on the x-component of velocity (or the error therein) can be used as a surrogate for the entire velocity magnitude, due to the large contributions. The turbulence model chosen in generating the fine mesh and coarse mesh data is the LES wall adapting local eddy viscosity model (WALE) model [153]. To the interested reader, Ben-Nasr et al. [28] has conducted an excellent review between popular sub-grid scale models and shown WALE to be an appropriate model for such a wall-bounded study.

7.3.3.2 a-priori results

The training data are sampled across three different Re numbers of 290, 395 and 500. Further each Re number consists of fourteen different coarse mesh configurations. The reduction factor for the channel study ranged from 1.12 to 4.5, for fine

mesh of 60000 cells. The training dataset comprised of about 8 million points. The automated machine learning step, yielded a network with eight layers and 48 neurons in each layer. Figure 7.12 shows the a-priori network performance compared to the ground truth for an intermediate Re number. Figure 7.13 shows the histogram of the normalized velocity predictions and the percentage difference between the predicted and the ground truth data. The a-priori test performance has a R^2 of 0.8460, which indicates a decent fit. The large range of learning, in terms of the near wall behavior, mean flow characteristics across different discretization and large scale flow configurations, are some of the challenges in the model's learning ability.

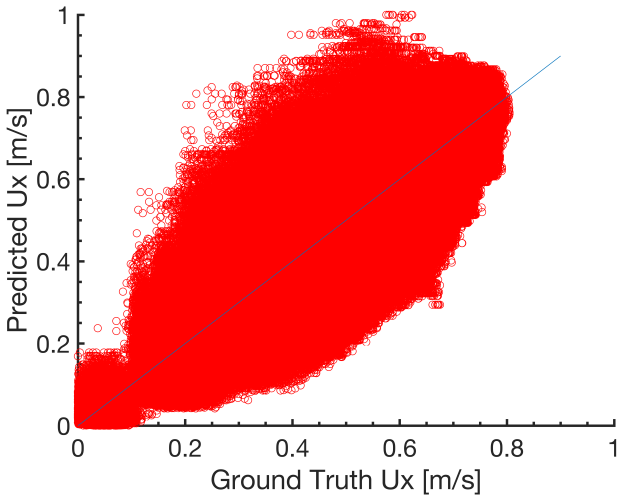

Figure 7.12. The apriori results from the channel study show a good R^2 fit

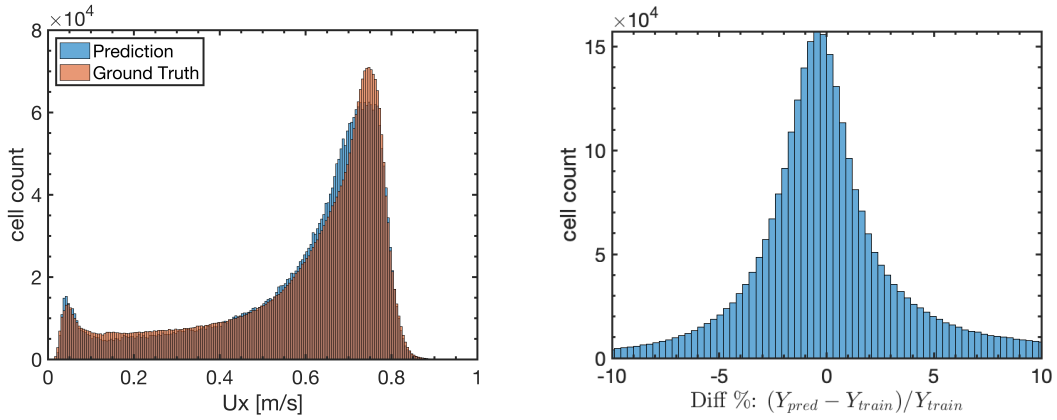

Figure 7.13. Histograms indicate close network predictions (left panel) and relative lower errors (right panel)

7.3.3.3 a-posteriori results

The trained network is coupled to the OpenFoam solver pimpleFoam, using a batch size of 13000. The qualitative performance for the velocity magnitude is shown at a mid-clip plane in Figure 7.14. This snapshot is taken at time t=1000s, for a reduction factor of 2. The comparisons indicate the coarse mesh simulations (middle panel) fails to accurately resolve the near wall effects. On the other hand, the error surrogate model enhanced coarse simulation recovers a large degree of lost information near the walls. Figure 7.15 presents the velocity magnitude difference between the mapped (ground truth) and the CFD simulations. The top panel shows the difference between mapped and coarse mesh simulation, and the observation in the near wall region behavior is consistent to the earlier result. The lower panel is the difference (shown on the same scale) between the mapped and the locally enhanced coarse mesh simulation. It is clearly evident that the locally enhanced simulation is able to recover some of the lost information.

To further demonstrate the improvement, Figure 7.16 shows the histogram of the difference (mapped field - CFD simulation) fields. This reveals the histogram of the

locally enhanced simulation has a narrower width around the zero point, indicating that information has been recovered in the process of adding the source.

While the previous figures show instantaneous results at the end of the simulation, for a transient simulation the time-averaged measurements might reveal key outcomes. Using the data time-averaged for the entire duration of the simulation (1000s), Figure 7.17 shows the difference between the mapped fields and the CFD simulations. The top panel shows the coarse mesh results and the bottom panel shows the locally enhanced simulations. For the coarse mesh simulations, the near wall behavior is the region of the highest discrepancy. This is consistent from earlier observations. For the locally enhanced simulations, the regions of maximum error are localized in a very small region near the upper wall, with a lower error magnitude compared to the coarse mesh results. For the lower wall, the locally enhanced simulations yield significantly better to coarse mesh simulations. From figure 7.15 this can be attributed to the large degree of transients at the upper wall from the instantaneous snapshot shown earlier.

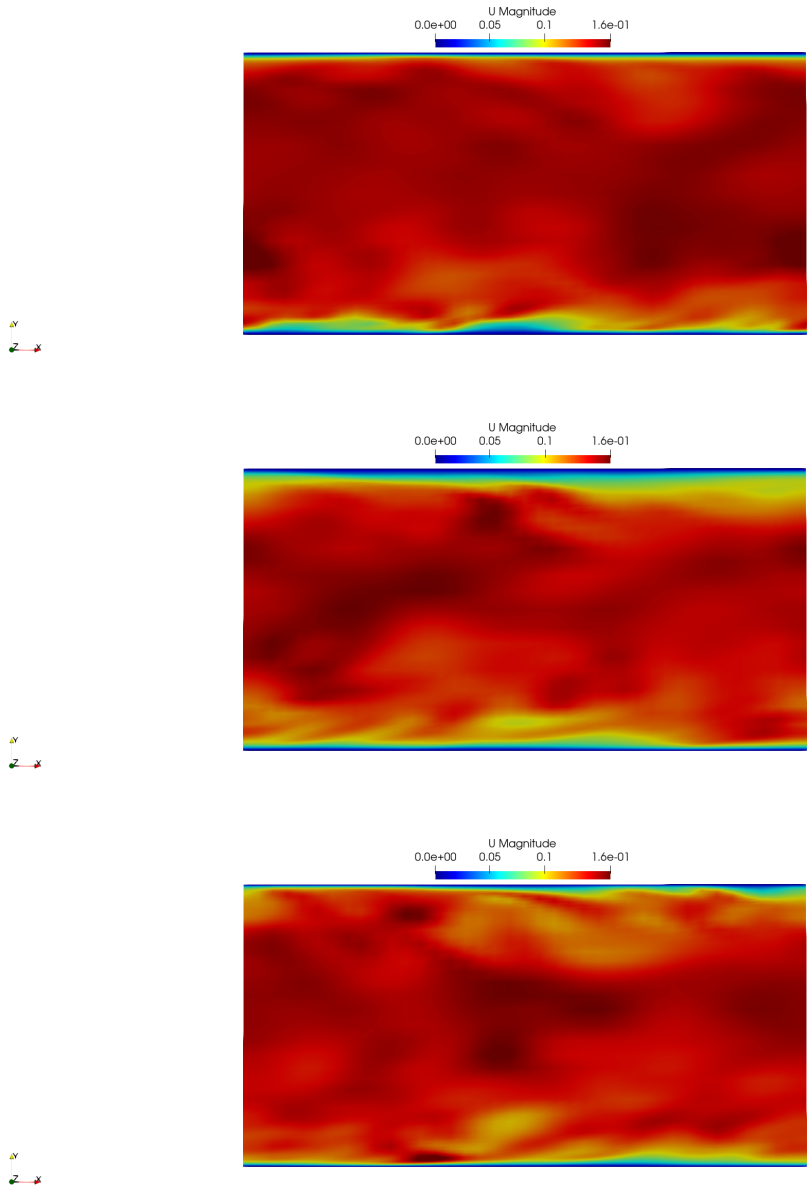

Figure 7.14. The panel at the top is from the fine to coarse mesh mapping, the middle panel is from the coarse mesh simulation, and the bottom panel is from the network enhanced simulation. Each plan shows the mid-clip plane colored by the velocity magnitude (scaled similarity). It is clearly evident that the network enhanced (bottom panel) recovers missing information (ground truth in the top panel) compared to the coarse mesh (middle panel) simulation.

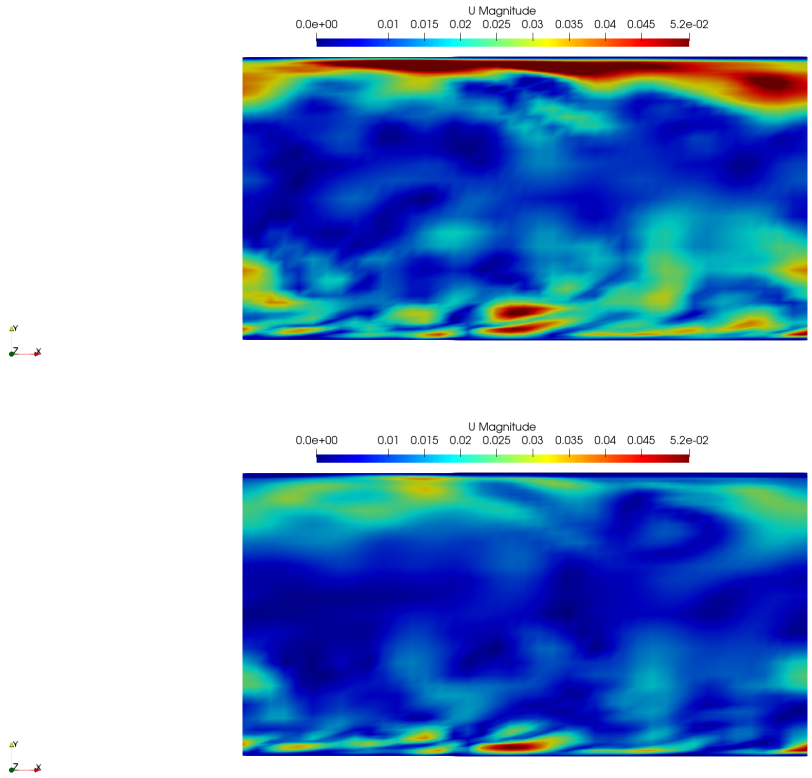

Figure 7.15. The differences in the velocity magnitude further confirms the earlier observation that the network enhanced (bottom panel) recovers missing information, and therefore has lower velocity magnitude differences.

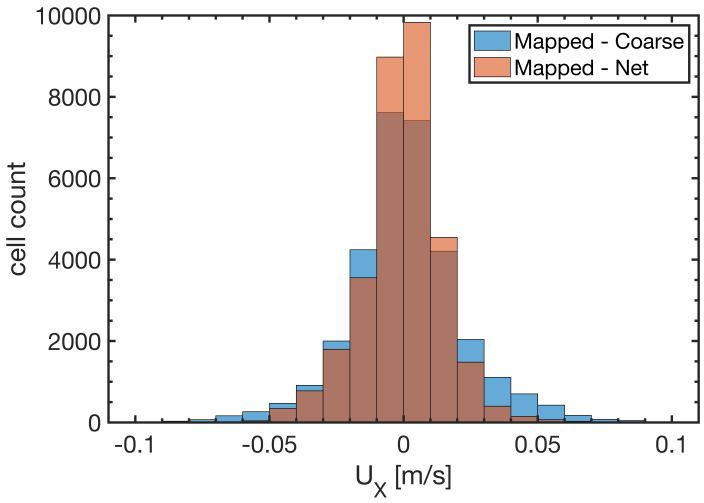

Figure 7.16. Histogram of velocity differences reveal narrow distribution for the network enhanced simulation, therefore confirming that lost information is retrieved.

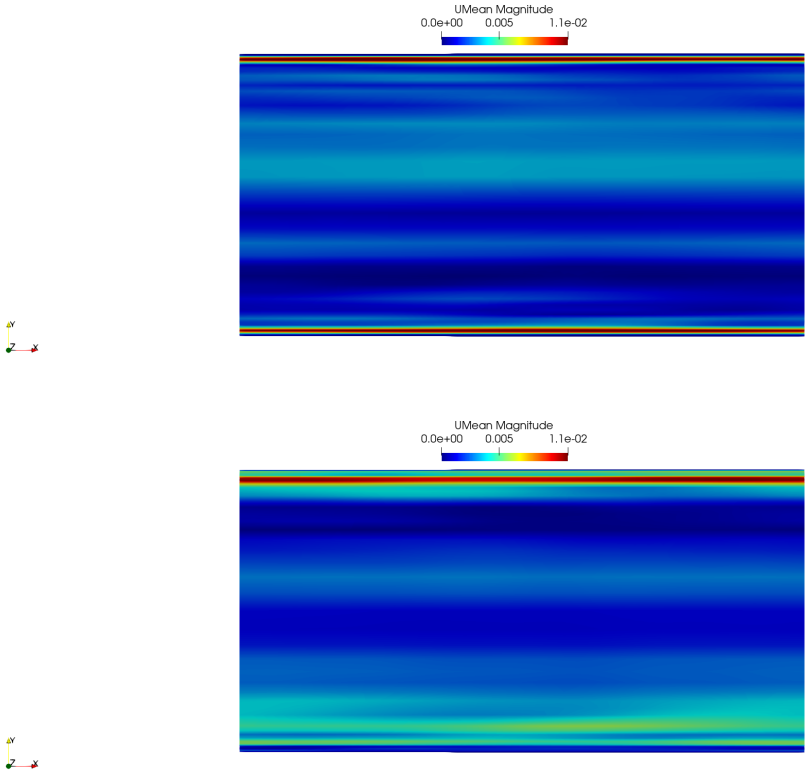

Figure 7.17. Time-averaged velocity difference mid-clip plane indicates the network enhanced simulation has recovered lost information

Further analysis is done using line plots at three equally spaced locations. The instantaneous and mean x-direction velocity performance are reported. These probes are placed at the center of the channel at 1, 2 and 3 m from the channel entrance (total length of the channel is 4 m). Figure 7.18 shows the instantaneous x-component of velocity at the three different locations - 1m (top panel), 2m (middle panel) and 3m (bottom panel). It is difficult to identify a particular trend, and therefore a more relevant comparison are the time-averaged plots. One important takeaway from this however is that the addition of source does not lead to abnormal behavior, especially in the near wall region. This is important as during the network training process, the cells associated with the near-wall region are limited and therefore represents a challenge in the learning process. Figure 7.19 shows the time-averaged x-component of velocity plots for the three locations. For all the three panels representing the 1m probe (top), 2m probe (middle) and 3m probe (bottom), the takeaway is consistent. It is that the locally enhanced simulations improve the solution performance and recover lost information, especially in the near-wall region and in the mean flow. Figure 7.20 represents the instantaneous snapshots of the Turbulent kinetic energy and the Reynolds stress tensor (in the near wall region) shows the close improvement in the prediction for the learned correction model. For the turbulent channel flow, the near wall region is the most challenging to resolve and very important from the perspective of viscous dissipation, and energy generation.

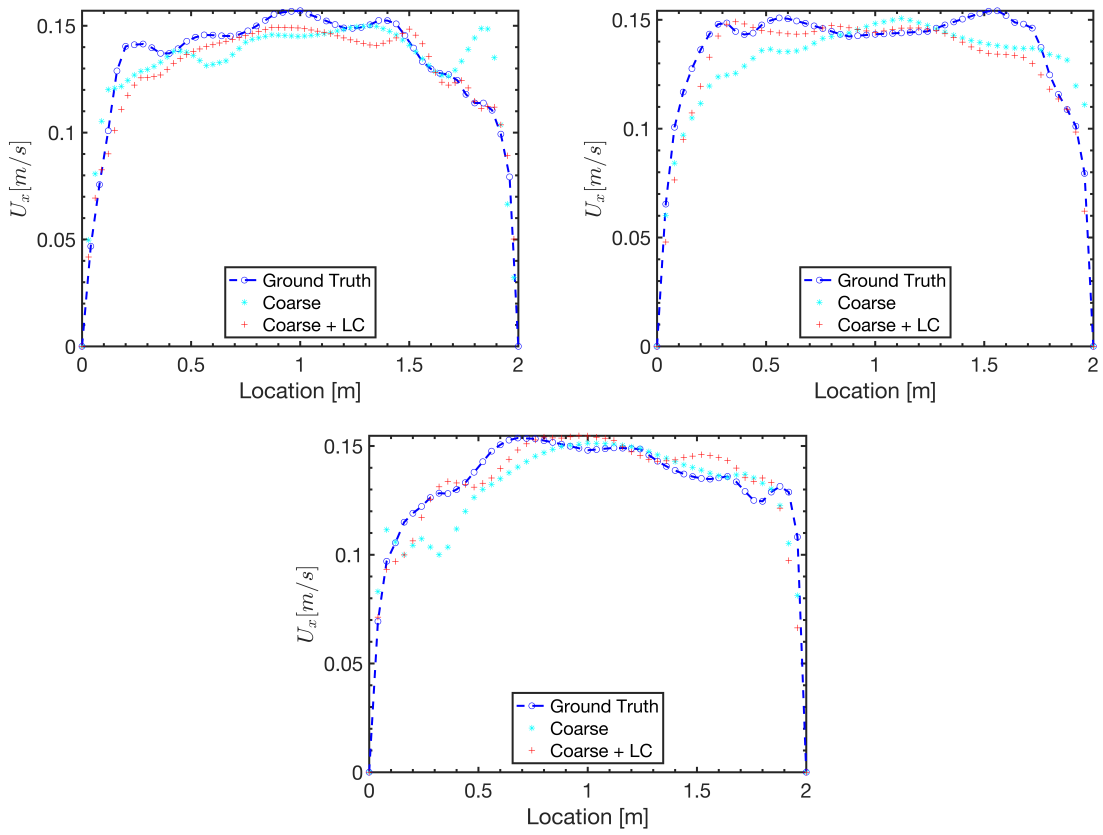

Figure 7.18. Instantaneous velocity plotted at x=1 m (top panel), x=2 m (middle panel) and x=3 m (bottom panel) show the locally enhanced simulation to spatially track the ground truth data better over the coarse mesh

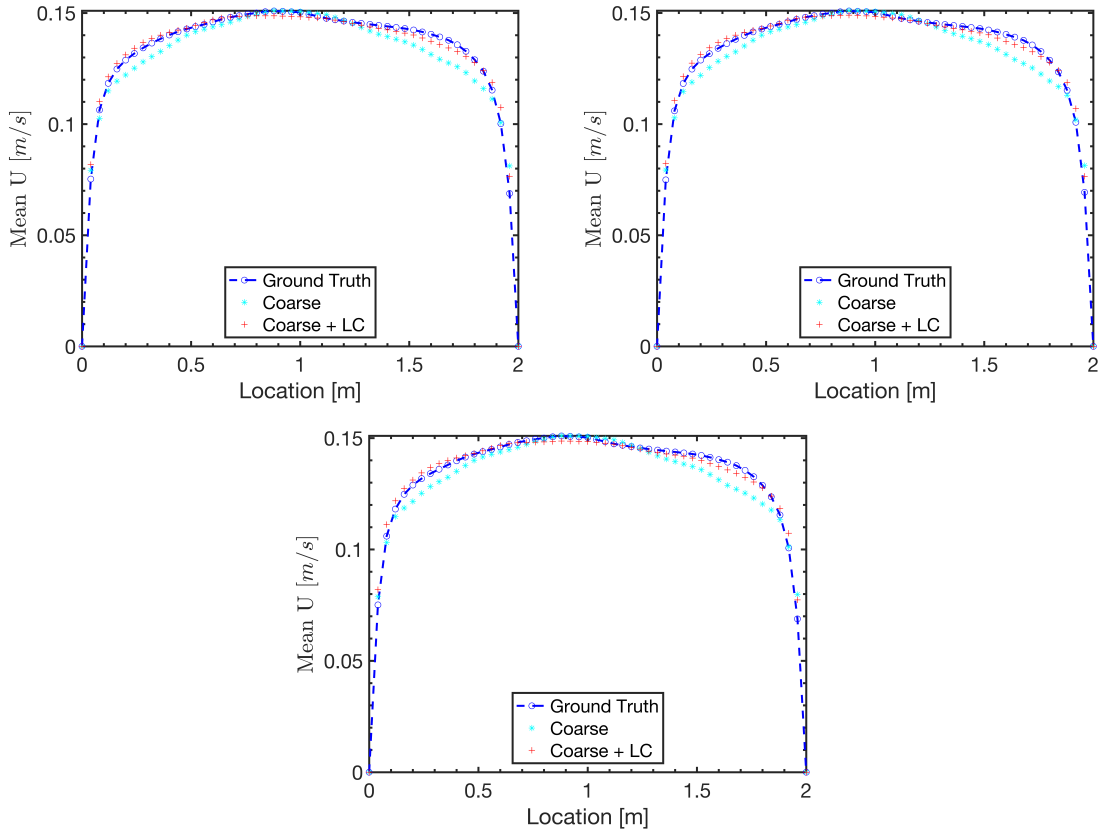

Figure 7.19. Time-averaged velocity plotted at x=1 m (top panel), x=2 m (middle panel) and x=3 m (bottom panel) show the locally enhanced simulation to spatially track the ground truth data more accurately over the coarse mesh

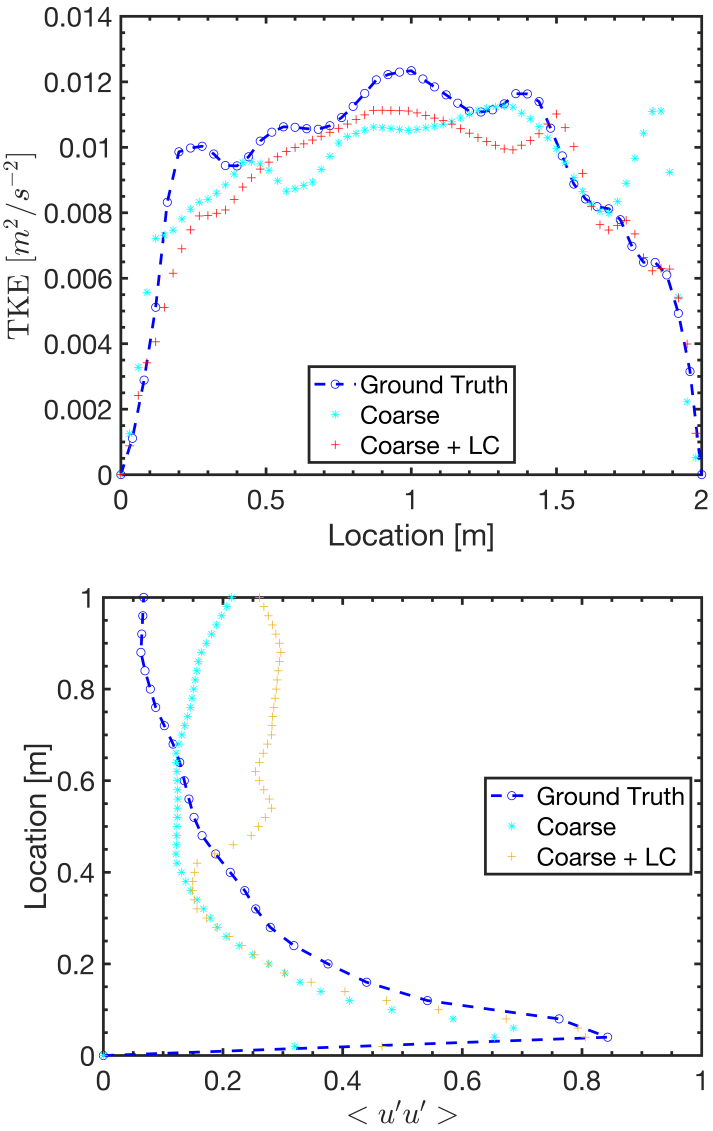

Figure 7.20. The top panel with the instantaneous TKE and the bottom panel with the Reynolds stress show the network enhanced simulations improve the near-wall behavior

In addition to the qualitative diagnostics, the L2-norm is calculated for the entire range of reduction factors. The L2-norm plot is shown in figure 7.21. In comparing the coarse and the locally enhanced simulation performance for each reduction factor, it is clearly evident that the addition of the source term improves the simulation fidelity significantly. The maximum benefit is observed for the higher degree of

coarseness. This is understandable since for a very coarse mesh, the loss of details is proportionately higher and therefore the learned correction model has a larger impact in the recovery.

Figure 7.22 represents the compute cost versus accuracy trade-off for the turbulent channel flow study. The radii of each of the circles represent the reduction factor. The larger the reduction factor (or coarser the mesh) larger the radii of the circle. Comparing circles of similar sizes gives a measure of the performance gains with the local enhancement approach. The general trend is that using the local enhancement approach, there is a potential for massive gains at a moderate increase in inference cost - about 10%. One other way to look at this is to compare the coarse mesh circles (blue) with the local enhanced circles (orange) along the Y-axis. In other words to obtain a L2-norm of 0.35, the coarse mesh simulations took about 1000s [wall time], whereas similar levels of performance were obtained at a fraction of the compute cost in approximately 300s, thereby indicating a speed-up of over 3x for similar fidelity solution. The speedup can be further improved by using a finer mesh LES, which is expected to be more accurate and expensive to compute. This would result in a higher information retrieval at a fraction of the cost making the cost-accuracy tradeoff even more beneficial. Kochkov et al. [108] used 2D DNS dataset for their ground truth and reported 40-80x speeds. The cost to perform DNS on this channel flow is orders of magnitude higher than the fine mesh LES used and therefore there are performance gains yet to be realized using this approach.

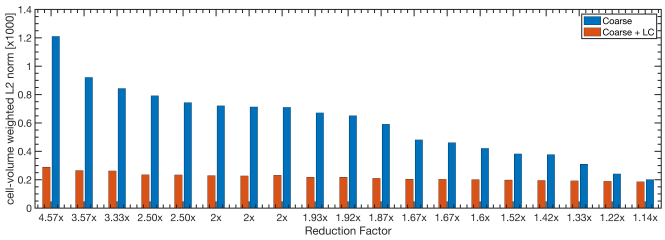

Figure 7.21. The error comparison shows a clear improvement in using the local enhanced CFD framework

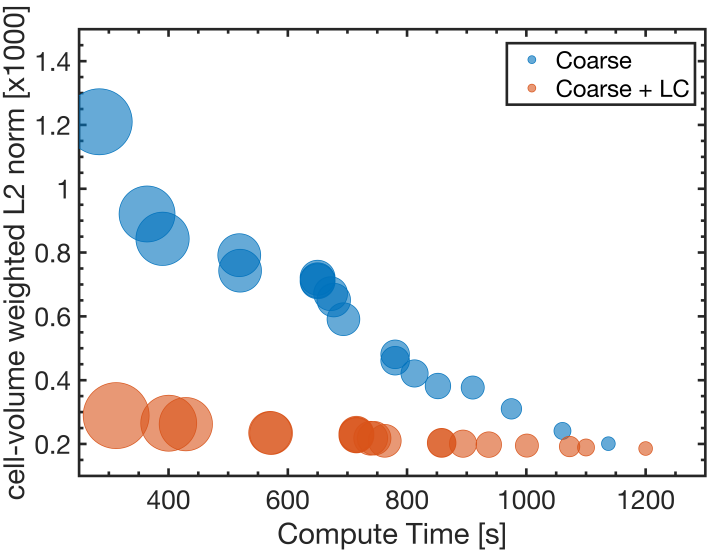

Figure 7.22. The timing plot shows the relative improvement in the cost versus accuracy, as a result of the local enhancement

RF	% improvement
4.57	76.11
3.33	70.25
2.50	68.51
2	67.31
1.52	48.14
1.14	7.67

Table 7.4. As the mesh coarsens, the expectation is that the network enhancement has more potential to contribute, as the coarse mesh resolves less features thereby providing the neural network to add more of the lost information. The results here across a range of reduction factors indicate an improvement as the mesh coarsens consistent with the expected outcome.

7.3.4 Simplified engine study

Our last test case involves a setting relevant to Internal Combustion Engine (ICE) modeling, that is the coldEngineFoam. This solver for modeling cold-flow in an internal combustion engines by moving a piston freely and is a transient, compressible, solver with the energy equation turned on. In this simulation, the transient mesh simulation would make the problem challenging in comparison to the previous cases.

The metrics of interest are the stability of the proposed approach in terms of uncertainity quantification in terms of error (CFD + ML predictions to the fine to coarse mapped solution) propagation, numerical stability among others. Quantifying the error in the form of L2 norm for velocity discrepancy is proposed to be studied.

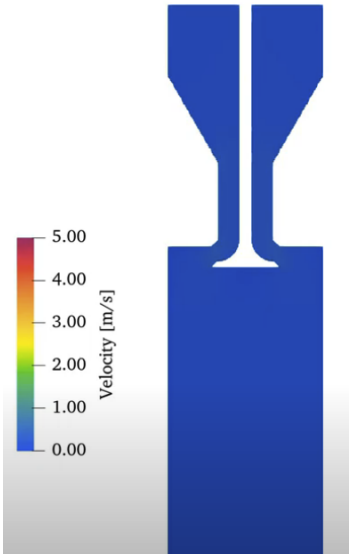

Figure 7.23. Schematic of the typical example for coldEngineFoam. The piston moves in the Z-direction and the image is a mid-clip in the Y-direction.

7.3.4.1 a-priori results

To reduce the numerical complexity there are no sprays or combustion involved, however this simple piston motion representing a compression stroke has significant computational challenges compared to the previous studies. First, this simulation operates in a low-mach number compressible setting in contrast to the incompressible flows studied so far. In addition to that there are challenges with regards to the transient mesh motion, and the energy equation is additionally solved for calculating in-cylinder temperatures. For the scope of this study we enhance the local cell velocity, similar to the previous examples. The network training steps involves using data across three Reynolds numbers and twelve reduction factors ranging from 1.12x to 4.58x. The total number of data points used in training is about 4 million.

The flow in this 2D study is dominated by the X and the Z components. Therefore we train a multi-output feed forward network predicting both of these velocity components. Figures 7.24 and 7.25 show the a-priori performance of the network for the x-component and the z-component respectively. The R^2 value for x-component network is 0.877 and for z-component is 0.943, which indicates a good model fit. For training this network, an additional set of inputs were used. Advection-dominated transient flows are classified as hyperbolic PDEs. The first-order wave equation is a good example. For the case of constant advection speed, the solution at any point is equal to the solution from somewhere just upstream at a slightly earlier time. So adding the timestep velocity information at the previous time provides the necessary and sufficient condition for modeling these hyperbolic systems seen in fundamentally transient flows like the IC engine.

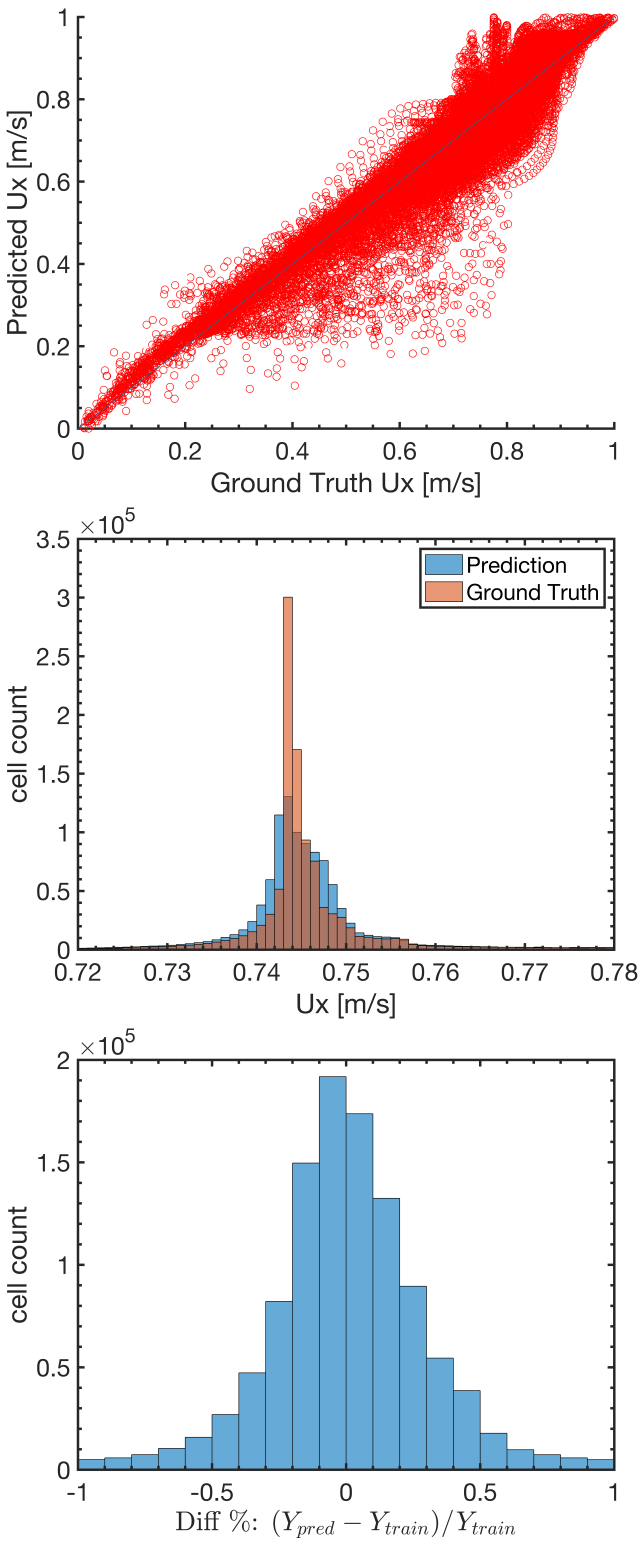

Figure 7.24. The apriori tests for the Ux velocity indicate a good fit to the data. The error histograms reveal relatively little errors

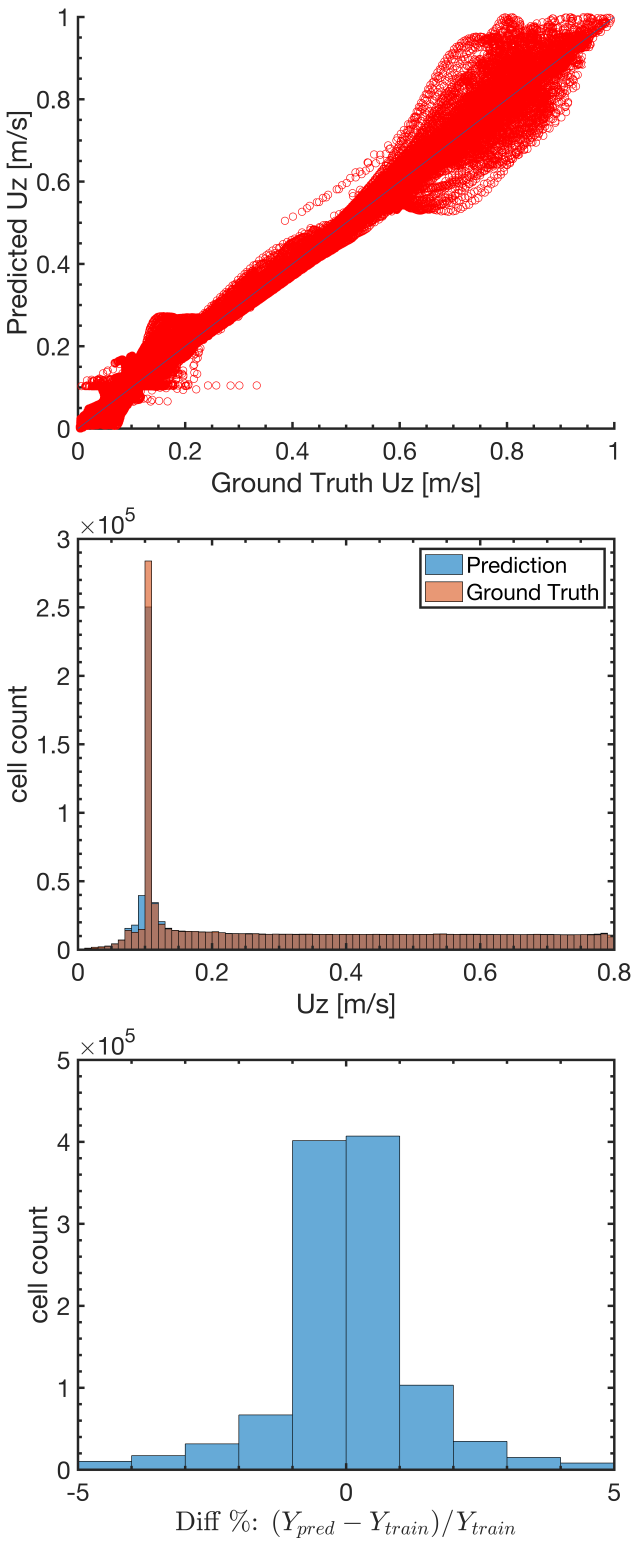

Figure 7.25. The apriori tests for the Uz velocity indicate a good fit to the data. The error histograms reveal relatively little errors

7.3.4.2 a-posteriori results

The turbulence model chosen in generating the fine mesh and coarse mesh data is the LES wall adapting local eddy viscosity model (WALE) model [153]. The velocity magnitude comparisons for the three different scenarios at the end of the simulation for a reduction factor of 2.2 is shown in Figure 7.26. The top panel represents the mapped field, the middle panel the coarse mesh field, and the bottom panel the local enhanced field. It is evident that the quality of the coarse mesh simulations do not fully resolve transient flow features. In contrast, adding the local enhancement recovers some of the lost details. Figure 7.27 plots the differences between the mapped field and the CFD simulations. The top panel representing the coarse mesh differences reveal regions of missing information which is recovered to some degree by the local enhancement.

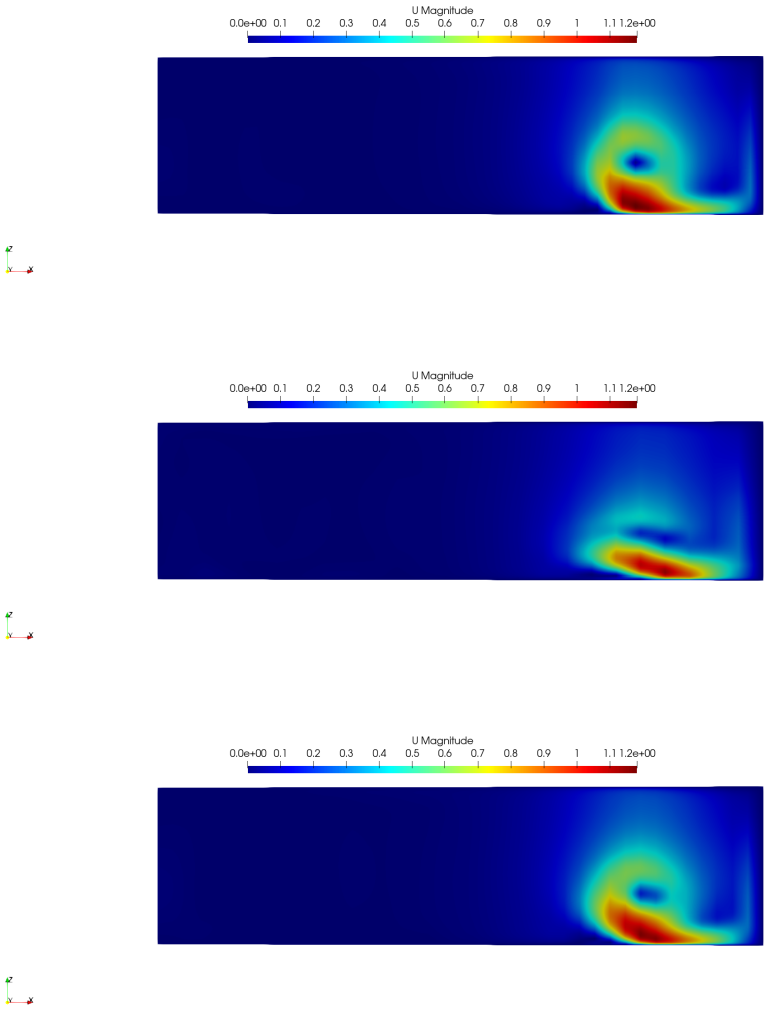

Figure 7.26. The network enhanced simulations (bottom panel) track the transient features (top panel) better compared to the coarse mesh simulations (middle panel)

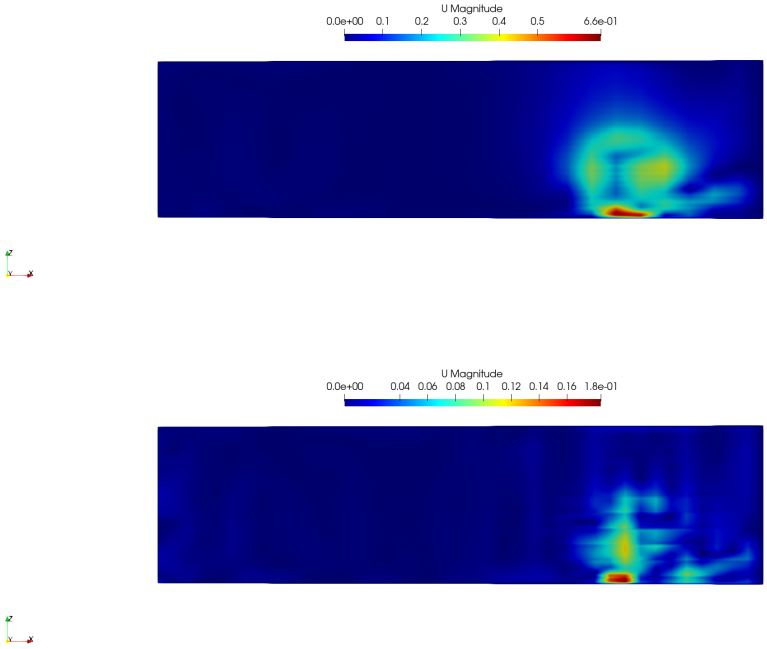

Figure 7.27. The velocity magnitude differences, compared to ground truth data, are shown in this image. The top panel is from the coarse mesh simulation and the bottom panel from the network enhanced simulation. It is evident that the network enhanced simulations (bottom) has less discrepancy with the ground truth data, compared to the coarse result (top panel), and therefore is able to recover part of the lost information.

To reveal the improvement in performance, a line plot across the domain is taken at the end of the simulation. At this time instant, in the z-component of velocity, there are a set of counter rotating vortices that amplifies the challenge for the machine learning model. Figure 7.28 reveals the locally enhanced simulation spatially tracks the ground truth mapped data significantly better over the coarse mesh simulations, therefore enhancing the model fidelity.

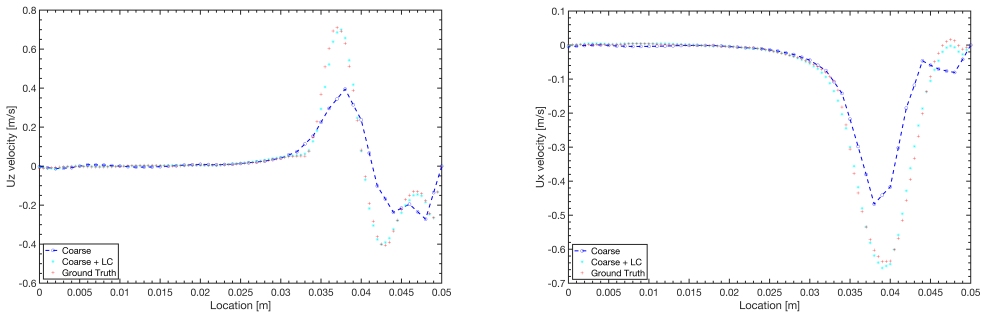

Figure 7.28. The line probe plots indicate the locally enhanced CFD spatially tracks the ground truth data better compared to coarse mesh simulations.

Along with the qualitative metrics, the L2 norm quantitative metric has been used. Figure 7.29 represents the results in the form of a bar chart across all reduction factors. It is to be noted that the Y-axis is on the logarithmic scale, to accommodate the range of data. As we move from higher to lower reduction factors, the L2 norm reduces, as is expected. However the local enhanced simulations represents a significant decrease in the L2 norm over the corresponding coarse mesh results for the each reduction factor.

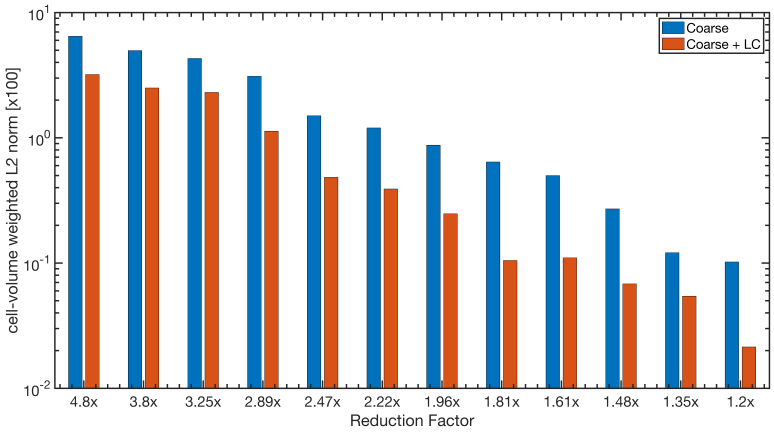

Figure 7.29. The engine study L2 norm shows the relative improvement in using local enhanced approach

The compute cost-accuracy trade off plot, shown in Figure 7.30, reveals a similar story to the previous result. Here the radii of the bubbles indicate the reduction factor, and the comparisons should be made between the bubbles of the same size. For each of the similar-sized bubbles, the local enhancement method improves the fidelity of predictions by around 50 % with only a marginal increase in computational cost, about 10-15 %. The other way to look at this result is that for reaching a $L2$ norm of 0.007, the coarse mesh simulation takes about 500s, whereas the local enhanced simulation reaches that level of fidelity in around 360s, thereby saving over 1.4x the compute time. This makes the compute-accuracy trade off curve favorable (more in Table 7.5

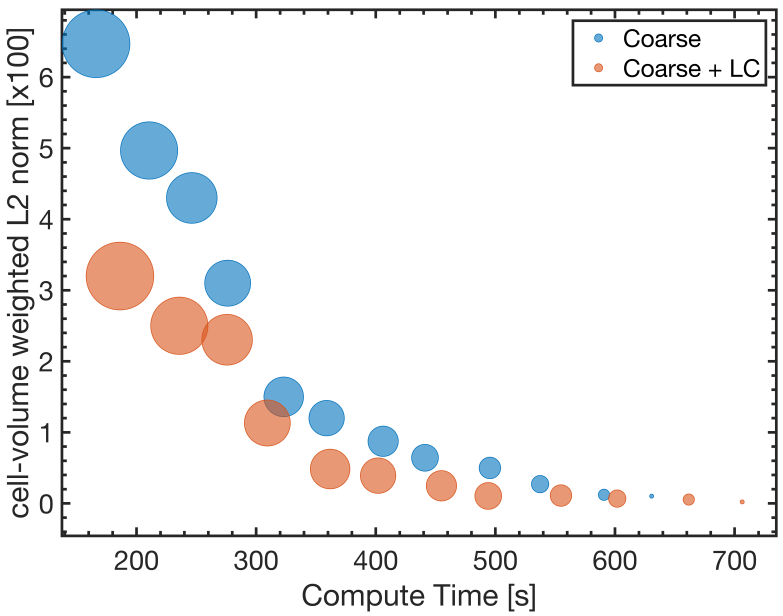

Figure 7.30. The cost-accuracy tradeoff plot shows improvement in using the local enhanced approach

RF	% improvement
4.8	79.01
3.8	55.04
2.89	77.94
2.22	67.03
1.81	59.1
1.48	42.13
1.26	33.81

Table 7.5. With larger reduction factor, the expectation is for a larger recovery in using the network enhanced approach. Consistent trends in improvement over coarse mesh is seen as the RF increases, thereby suggesting local enhanced CFD is able to recover lost information.

7.4 Out of Distribution performance

The real measure of the impact of this method is in estimating the performance on an out-of-distribution dataset. This is critical since retraining the network with additional data is expected to be expensive and add to the overall cost for using this approach. In the previous section, the demonstrated performances were of an in-distribution nature. The out-of-distribution (OOD) performance also reveals the generalizability and robustness for the inference engine. For each of the OOD tests, we keep the mesh configuration constant. In other words we use the same mesh resolution for all cases and change the Re of the simulation. Since the model is trained on a large corpus of reduction factors, it is expected that the model will be able to generalize better for different mesh configurations. However, the performance against different large scale flow features based on Re is expected to be more challenge. The OOD performance for the channel and the engine inference are done separately. For each set of tests, the Re equal to 1 is from the baseline study shown in the previous section. This is scaled to 0.5 x higher and lower than the baseline to estimate the performance in the nearby vicinity.

The top panel in Figure 7.31 shows the OOD performance for the turbulent channel flow study. It is evident that for Re lower than the baseline, the OOD performance is satisfactory. However, for the higher Re studies the local enhanced simulations have lower L2 errors compared to the coarse mesh simulations. And in terms of percentage improvement there is a significant decrease from the earlier Re numbers. It is very well known that machine learning models extrapolate poorly. So for the range of conditions that involve inputs that are out of the range seen during training, the network performance during inference is expected to deteriorate. Similar inferences can be deduced for the bottom panel showing the OOD performance for the engine compression stroke study. It is therefore important that the model limitations in using this method for a range of operating conditions seen during training is understood for a reliable OOD performance.

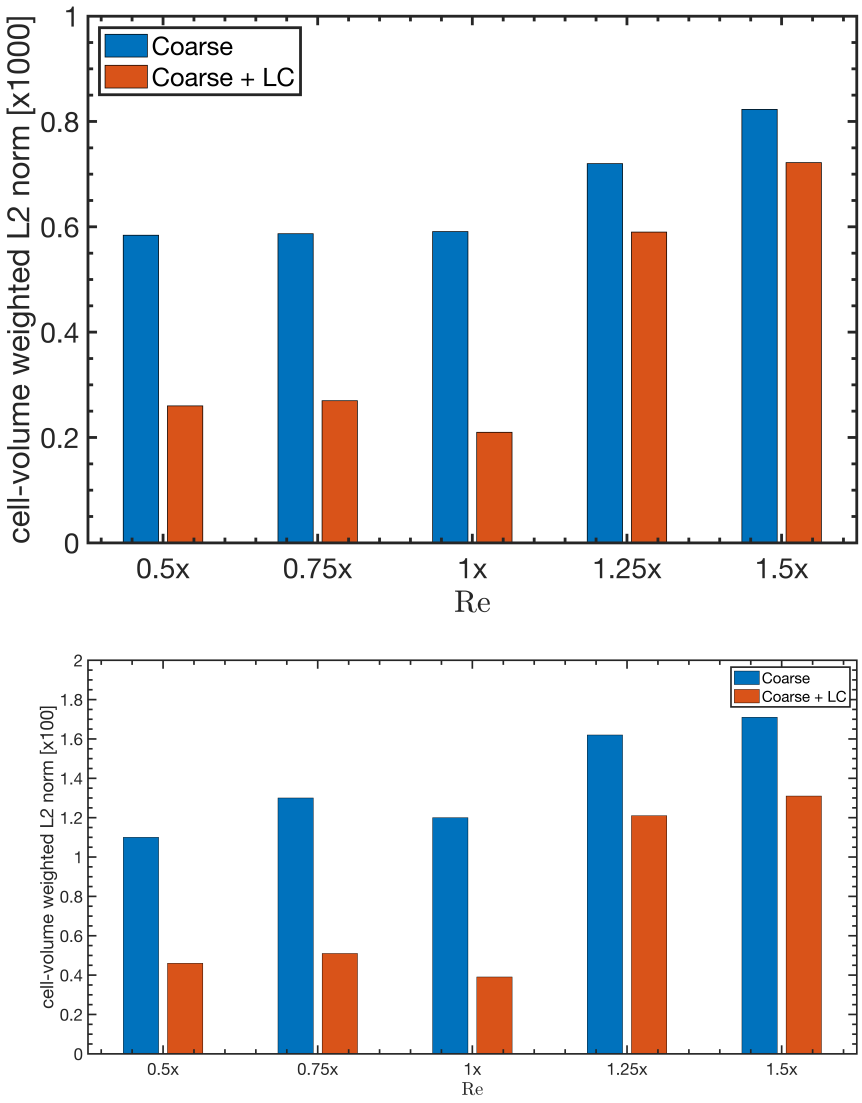

Figure 7.31. The out of distribution performance for the channel (top) and engine (bottom) studies show similar trends. In the immediate vicinity of the training dataset, the performances are reasonable so long the inputs are within the data-distribution of the learning.

7.5 Brief Inferences

The local enhanced model presents a novel paradigm for integration machine learning based discretization error estimation approach into CFD. We present a viable approach in training the network based on CFD data, integrating learned model into an

open source CFD code, and presenting metrics to validate the performance of such a framework on a variety of test settings. Some important observations and inferences include:

- We demonstrate the development of a machine learning error surrogate model coupled with full 3D CFD
- Cost-accuracy trade-off is favorable, inference can be further improved using pruning
- The ML enhanced simulations are numerically stable, provided the appropriate relaxation factor is chosen
- There's modest generalization ability, and can be improved further using transfer learning
- Given a nominal configuration, we developed a methodology that designers can use to efficiently characterize dynamics and explore the effects of parametric changes on quantities of interest
- In the current formulation for network enhancement, there is no explicit mass conservation step imposed. This may be necessitated due to addition of extra source terms into the simulation. The conservation is implicitly implemented at the next time step.

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

Low-mach number terrestrial propulsion systems such as medium and heavy-duty Internal combustion engines are expected to comprise of a large proportion of on-road vehicles well into the mid-21st century. Numerical simulations, including CFD, provides a viable pathway in improving the performance and efficiency of these engines. Since performing DNS for these applications are prohibitively expensive, the focus of this thesis has been on developing reliable, robust engineering level reduced order models. This thesis can broadly be divided into two different sections. Spray atomization is one of the limiting processes for combustion in ICEs and is therefore a major area of focus for the first part of this thesis. In the first section, CFD simulations using diffused interface approaches such as the Homogeneous Relaxation Model (HRM) has been used to model the phase change arising due to fuel cavitation. From these simulations on single/multi-hole cylindrical and conical nozzles, different mechanisms for the onset of cavitation have been identified. In a subsequent study, a new primary atomization model based on mixing limited physics has been proposed and implemented in a full 3D Navier-Stokes solver. In the second part of this thesis, data-driven methods for surrogate modeling are explored. The application areas include turbulence closure and error estimation due to grid coarsening.

The key findings of this thesis are presented in the next section, and suggested future work and exploratory studies are discussed in the concluding section of this chapter.

8.1 Summary of Conclusions

The internal nozzle flow represents a unique challenge in building reliable computational tools. Until very recently, CFD modelers relied on near nozzle, dense core region experimental data to validate their models for internal flows. These experimental observations are highly transient and obtaining dense core data is significantly challenging due to a multitude of reasons detailed in Chapter 3 and 4 of this thesis. In Chapter 3 and 4, state-of-the-art experimental studies conducted by the the community (including research collaborators) presented the unprecedented opportunity to validate numerical model predictions within the internal nozzle itself. Once sufficiently validated, the CFD studies using in-house HRMFoam [195] code revealed important mechanisms for the onset of cavitation in a variety of nozzles. For a cylindrical single hole nozzle, it was determined that the shape of the inlet corner radii has a significant effect on the cavitation propensity. Furthermore, it was discovered that the manufacturing defects of these nozzles have a sizeable impact on creation of local cavitation sites. This is consistent with observations from experimental studies. Previous studies in the literature suggested that conical nozzles are expected to suppress cavitation. However, upon closer investigation of real injector geometries and simulating transient injection behavior, it was discovered that even for conical nozzles there are conditions that favor cavitation. These include in-cylinder thermodynamic conditions and the effect of transient needle motion. This important discovery has been validated extensively against experimental studies. The mechanisms for bulk cavitation has been identified post the needle closure event. It was also found that the presence of vapor in the injector has a significant effect on the subsequent injection cycles, an observation validated against experimental studies. Unlike single hole injectors where the flow direction and the nozzle hole axis are aligned, for real world multi-hole nozzles the off-axis location of the nozzle hole means that the accelerated flow passing through the injector often has to turn at sharp angles to exit the nozzle.

Other studies show that a sharp inlet corner causes flow separation due to negative pressure gradients in the needle seat and inlet corners, thereby promoting localized cavitation zones due to the sudden pressure drops.

In the second part of spray atomization studies, the focus shifts to modeling external sprays. Current spray atomization models are based on a high degree of empiricism due to the many constants that are present in them. We propose extending a one-dimensional mixing-limited physics based model [205, 150], and integrated it to a full 3D Navier Stokes solver. We extended the original model [205] by adding the physics of vaporization based on a 1D vaporization model by Desantes et al.[49]. The model was validated against experimentally obtained vapor penetration data along the spray axis and at transverse locations downstream of the nozzle exit. This was done for a range of spray nozzles and fuels. The comparable performance to highly optimized standard Lagrangian-Eulerian models is encouraging, as this model inputs are obtained experimentally and has no tunable constants.

The last part of the thesis is focused on exploring machine learning techniques relevant for ICE modeling. The primary objective is to explore these methods for building surrogate models. In the first part, we build a data-driven surrogate LES closure model. In building this model, we employ a Bayesian optimization based auto machine learning (autoML) method. It is widely understood that the choice of network architecture and training hyper parameters (such as learning rate, batch size) has a leading order effect on the trained network performance. Using random or grid search on networks training on millions of datapoints is non-trivial. Besides the nature of scientific datasets - complex, high-dimensional and the non-convex nature of optimization means a-priori estimating parameters requires human expertise. This autoML method alleviates all these shortcomings by automating the task of network architecture search and identifying best performing hyper parameters using Bayesian optimization. We demonstrate the effectiveness of this approach by estimat-

ing the post-training learning process and visualizing the loss manifold explored by this method. Models trained using these methods generally outperform hand-tuned models and is therefore particularly suited for these high-dimensional applications. Using the trained surrogate model, and integrating it with OpenFoam the machine-learned closure model is not only numerically stable but also the predictions are highly accurate. The network inference is further accelerated using techniques like pruning and quantization by over 10-15x, thereby further reducing the cost overhead of using these methods.

In the final chapter of the thesis, a machine learning surrogate model to predict grid coarsening error was built and tested across different problems and settings. The model trained using automatic machine learning techniques to jointly optimize better performance on test data as well as generate smaller, efficient networks shows good agreement with the ground truth data in an a-priori and a-posteriori setting. This learned correction approach is close in spirit to the residual correction for discretized Navier-Stokes equations previously discussed in [108, 206, 230]. The network integrated CFD is numerically stable, provided an accurate time constant be chosen for the source term addition. Further acceleration of the inference engine is feasible using network pruning, although it might cause some loss in generalizability. Overall, noticeable performance gains (30-70%) across various settings are recorded for a modest increase in computational time (about 10-15%). This therefore makes the compute-accuracy trade-off more favorable than current methods. And this achieves the principal aim of our study, which is to accelerate these simulations as well as improving the accuracy for many such scenarios.

8.2 Suggested future directions

In this thesis, various methods to numerical model physical phenomena relevant to ICEs were developed, studied and reported. However, there still remain unexplored

areas and questions unanswered. In this brief section, some of these suggested future directions for this research are presented.

8.2.1 Internal nozzle flow

The transient simulations for the single hole injector, detailed in Chapter 4, yielded important findings including mechanisms for the onset of cavitation in smooth injectors. However, most real world injectors are multi-hole in nature. It is therefore prudent to study these multi-hole injectors in a transient, moving needle setting. The failure of the HRM model to accurately represent the near-nozzle asymmetry in Chapter 4 suggests that the mesh resolution and numerical models (turbulence) choices play an important role in these models and therefore must be carefully selected.

8.2.2 External flow

The ELMO model has demonstrated impressive results but there are some severe shortcomings in the current setup. The premise of this model, as designed currently, is to advect along the axis of the spray. While the assumption of primarily axial flow is generally accurate for most single hole spray applications, it hinders the ability for the spray to turn and prohibits plume-to-plume interactions which is a common feature for multi-hole gasoline sprays under flash-boiling conditions. To enable this feature in ELMO, the easiest approach would be to vectorize the advection equations such interactions along transverse axes are successfully incorporated.

ELMO's biggest strength is the lack of tunable constants. It is also its biggest weakness. This is evident in the regions of the spray where the physics deviate from the mixing limited paradigm and is therefore only applicable in the dense core region of the spray atomization. The ELMO model predicts the theoretical maximum energy exchange possible, thereby leading to lesser liquid fuel penetration. For sprays with lower Weber number the applicability of the mixing limited paradigm is not very

clear. Further exploration into extension of this model to a wider set of conditions might be an interesting area to further establish the model's capabilities.

8.2.3 Machine-learning for CFD

For the machine learned error-estimation surrogate model, there are two main sources of approximation. First is the use of very fine mesh LES data as the ground truth as performing DNS is computationally intractable. The other is the fine to coarse mesh mapping that leads to interpolation errors. For the latter, the interpolation error can be reduced further if a higher-order mapping method is used. This would require writing a new mapping utility in OpenFoam based on the mapFields library and will be computationally expensive than current methods. However, the higher cost associated with this will be spent during the training phase and this would most definitely lead to improving the model's performance during inference as the network will now be trained on a much better dataset.

The main advantage of this method is in studies with large mesh counts. The demonstration cases in Chapter 7, were for small to moderate cell counts, reasonable for a proof-of-concept study. However, the cost to perform high-fidelity simulations scales non-linearly with the total cell count. For example, given a reduction factor of 2 the theoretical expected speed up is about 8 times the fine mesh. Therefore, it might be reasonable to extend this framework for much larger CFD problems involving millions of cell points. This benefit would show up on the compute-accuracy trade off, making this framework even more favorable with higher fidelity output.

The use of the relaxation (time scale) factor enabled the numerical stability of the source term. However in the current paradigm, the choice of the relaxation is *ad-hoc*. It will therefore be valid to develop some scientific intuition about *a-priori* estimation of the relaxation factor.

The current error estimation model is a black box. In order to understand the learning better and determine robustness to OOD evaluations, performing sensitivity studies (as shown in Chapter 6) might be useful. This would also help in building the pathway to explainable grey-box physics surrogate models. The challenge for using ML in production is to generalize to constantly changing edge cases. Two main approaches can help alleviate this issue. First is the use of massive data because more data can lead to better generalization. The second and the most important is to build necessary infrastructure that allows models to learn to adapt in real-time with additional data. This is particularly relevant to observations reported in Chapter 7, regarding the OOD tests.

Lastly, loss functions play an important role in helping the learning process. Physics based loss as seen in Physics Informed Neural Networks [177] are a prime example. In our current setup, we use a mix of L1 and L2 loss terms, however they do not represent the physics of the problem. Physics based losses that might be relevant include using the target velocity $\mathbf{u}$, normalized by the local cell turbulent kinetic energy k. Functionally it can be represented as

$$\mathrm{L} = \frac{\mathbf{u}}{\mathrm{sqrt(K)}} \tag{8.1}$$

BIBLIOGRAPHY

[1] Engine Combustion Network. https://ecn.sandia.gov/. Accessed: 2020-10-20.

[2] Engine Combustion Network Spray C Injector. https://ecn.sandia.gov/diesel-spray-combustion/target-condition/spray-c-nozzle-geometry/. Accessed: 2020-10-20.

[3] Engine Combustion Network Spray D Injector. https://ecn.sandia.gov/diesel-spray-combustion/target-condition/spray-d-nozzle-geometry/. Accessed: 2020-10-20.

[4] Wikipedia Large Eddy Simulation. https://en.wikipedia.org/wiki/LargeEddySimulation, note = Accessed: 2020-10-20.

[5] Converge 2.4 manual, 2017.

[6] Abraham, John, and Magi, V. A virtual liquid source (VLS) model for vaporizing diesel sprays. *SAE transactions* (1999), 1363–1374.

[7] Agarwal, Arpit, and Trujillo, Mario F. A closer look at linear stability theory in modeling spray atomization. *International Journal of Multiphase Flow 109* (2018), 1–13.

[8] Als-Nielsen, Jens, and McMorrow, Des. *Elements of modern X-ray physics*. John Wiley & Sons, 2011.

[9] Andreini, Antonio, Bianchini, Cosimo, Puggelli, Stefano, and Demoulin, FX. Development of a turbulent liquid flux model for eulerian–eulerian multiphase flow simulations. *International Journal of Multiphase Flow 81* (2016), 88–103.

[10] Apte, SV, Gorokhovski, Mikhael, and Moin, Parviz. LES of atomizing spray with stochastic modeling of secondary breakup. *International Journal of Multiphase Flow 29*, 9 (2003), 1503–1522.

[11] Arai, M, Tabata, M, Hiroyasu, H, and Shimizu, M. Disintegrating process and spray characterization of fuel jet injected by a diesel nozzle. *SAE transactions* (1984), 358–371.

[12] Arienti, Marco, Doisneau, Francois, and Oefelein, Joseph. Modeling primary atomization of liquid fuels using a multiphase dns/les approach. Tech. rep., Sandia National Lab.(SNL-CA), Livermore, CA (United States), 2019.

[13] Arienti, Marco, and Sussman, Mark. Nozzle geometry effects on primary atomization. Tech. rep., Sandia National Lab.(SNL-CA), Livermore, CA (United States), 2012.

[14] Arienti, Marco, and Sussman, Mark. A numerical study of the thermal transient in high-pressure diesel injection. *International Journal of Multiphase Flow 88* (2017), 205–221.

[15] Ashgriz, Nasser. *Handbook of atomization and sprays: theory and applications*. Springer Science & Business Media, 2011.

[16] Baldwin, Eli T. Eulerian CFD modeling of multiphase internal injector flow and external sprays.

[17] Baldwin, Eli T, Grover Jr, Ronald O, Parrish, Scott E, Duke, Daniel J, Matusik, Katarzyna E, Powell, Christopher F, Kastengren, Alan L, and Schmidt, David P. String flash-boiling in gasoline direct injection simulations with transient needle motion. *International Journal of Multiphase Flow 87* (2016), 90–101.

[18] Bangga, Galih. *Three-dimensional flow in the root region of wind turbine rotors*. kassel university press GmbH, 2018.

[19] Bao, Han, Dinh, Nam T, Lane, Jeffrey W, and Youngblood, Robert W. A data-driven framework for error estimation and mesh-model optimization in system-level thermal-hydraulic simulation. *Nuclear Engineering and Design 349* (2019), 27–45.

[20] Bao, Zhenshan, Liu, Jiayang, and Zhang, Wenbo. Using distillation to improve network performance after pruning and quantization. In *Proceedings of the 2019 2nd International Conference on Machine Learning and Machine Intelligence* (2019), pp. 3–6.

[21] Battistoni, Michele, Poggiani, Claudio, and Som, Sibendu. Prediction of the nozzle flow and jet characteristics at start and end of injection: transient behaviors. *SAE International Journal of Engines 9*, 1 (2016), 84–97.

[22] Battistoni, Michele, Xue, Qingluan, and Som, Sibendu. Large-eddy simulation (les) of spray transients: start and end of injection phenomena. *Oil & Gas Science and Technology–Revue d'IFP Energies nouvelles 71*, 1 (2016), 4.

[23] Baydin, Atılım Günes, Pearlmutter, Barak A, Radul, Alexey Andreyevich, and Siskind, Jeffrey Mark. Automatic differentiation in machine learning: a survey. *The Journal of Machine Learning Research 18*, 1 (2017), 5595–5637.

[24] Beck, Andrea, Flad, David, and Munz, Claus-Dieter. Deep neural networks for data-driven les closure models. *Journal of Computational Physics 398* (2019), 108910.

[25] Beck, Andrea, and Kurz, Marius. A perspective on machine learning methods in turbulence modelling. *arXiv preprint arXiv:2010.12226* (2020).

[26] Beheshti, Novid, Burluka, Alexey A, and Fairweather, Michael. Assessment of σ- y liq model predictions for air-assisted atomisation. *Theoretical and Computational Fluid Dynamics 21*, 5 (2007), 381–397.

[27] Belhadef, A, Vallet, A, Amielh, Muriel, and Anselmet, Fabien. Pressure-swirl atomization: Modeling and experimental approaches. *International Journal of Multiphase Flow 39* (2012), 13–20.

[28] Ben-Nasr, O, Hadjadj, A, Chaudhuri, A, and Shadloo, MS. Assessment of subgrid-scale modeling for large-eddy simulation of a spatially-evolving compressible turbulent boundary layer. *Computers & Fluids 151* (2017), 144–158.

[29] Bode, Mathis, Gauding, Michael, Lian, Zeyu, Denker, Dominik, Davidovic, Marco, Kleinheinz, Konstantin, Jitsev, Jenia, and Pitsch, Heinz. Using physics-informed super-resolution generative adversarial networks for subgrid modeling in turbulent reactive flows. *arXiv preprint arXiv:1911.11380* (2019).

[30] Bond, Peter J, Holyoake, John, Ivetac, Anthony, Khalid, Syma, and Sansom, Mark SP. Coarse-grained molecular dynamics simulations of membrane proteins and peptides. *Journal of structural biology 157*, 3 (2007), 593–605.

[31] Brenowitz, Noah D, Beucler, Tom, Pritchard, Michael, and Bretherton, Christopher S. Interpreting and stabilizing machine-learning parametrizations of convection. *arXiv preprint arXiv:2003.06549* (2020).

[32] Brunton, Steven L, Hemati, Maziar S, and Taira, Kunihiko. Special issue on machine learning and data-driven methods in fluid dynamics, 2020.

[33] Brunton, Steven L, Noack, Bernd R, and Koumoutsakos, Petros. Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics 52* (2020), 477–508.

[34] Brunton, Steven L, Proctor, Joshua L, and Kutz, J Nathan. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences 113*, 15 (2016), 3932–3937.

[35] Burluka, AA, Borghi, R, et al. Development of a eulerian model for the "atomization" of a liquid jet. *Atomization and sprays 11*, 6 (2001).

[36] Canuto, VM, and Cheng, Y. Determination of the smagorinsky–lilly constant cs. *Physics of Fluids 9*, 5 (1997), 1368–1378.

[37] Carleo, Giuseppe, Cirac, Ignacio, Cranmer, Kyle, Daudet, Laurent, Schuld, Maria, Tishby, Naftali, Vogt-Maranto, Leslie, and Zdeborová, Lenka. Machine learning and the physical sciences. *Reviews of Modern Physics 91*, 4 (2019), 045002.

[38] Castellano, Giovanna, Fanelli, Anna Maria, and Pelillo, Marcello. An iterative pruning algorithm for feedforward neural networks. *IEEE transactions on Neural networks 8*, 3 (1997), 519–531.

[39] Chatterjee, Anindya. An introduction to the proper orthogonal decomposition. *Current science* (2000), 808–817.

[40] Chern, I-L, Glimm, J, McBryan, O, Plohr, B, and Yaniv, S. Front tracking for gas dynamics. *Journal of Computational Physics 62*, 1 (1986), 83–110.

[41] Chima, RODRICKV. A k-omega turbulence model for quasi-three-dimensional turbomachinery flows. In *34th Aerospace Sciences Meeting and Exhibit* (1995), p. 248.

[42] Chollet, Francois, et al. *Deep learning with Python*, vol. 361. Manning New York, 2018.

[43] Dahms, Rainer N, and Oefelein, Joseph C. On the transition between two-phase and single-phase interface dynamics in multicomponent fluids at supercritical pressures. *Physics of Fluids 25*, 9 (2013), 092103.

[44] Dal Santo, Niccolò, Deparis, Simone, and Pegolotti, Luca. Data driven approximation of parametrized PDEs by reduced basis and neural networks. *Journal of Computational Physics* (2020), 109550.

[45] de Bruyn Kops, Steve, Riley, James J, and Portwood, GD. Toward direct numerical simulations of the stratified turbulence inertial range. In *International Symposium on Stratified Flows* (2016), vol. 1.

[46] Demoulin, Francois-Xavier, Beau, Pierre-Arnaud, Blokkeel, Gregory, Mura, Arnaud, and Borghi, Roland. A new model for turbulent flows with large density fluctuations: Application to liquid atomization. *Atomization and Sprays 17*, 4 (2007).

[47] Dent, JC. A basis for the comparison of various experimental methods for studying spray penetration. *SAE Transactions* (1971), 1881–1884.

[48] Desantes, JM, Pastor, JV, García-Oliver, JM, and Pastor, JM. A 1D model for the description of mixing-controlled reacting diesel sprays. *Combustion and Flame 156*, 1 (2009), 234–249.

[49] Desantes, Jose M, Lopez, J Javier, Garcia, Jose M, and Pastor, Jose M. Evaporative diesel spray modeling. *Atomization and Sprays 17*, 3 (2007).

[50] Desantes, José M, Payri, Raúl, Garcia, Antonio, and Manin, Julien. Experimental study of biodiesel blends' effects on diesel injection processes. *Energy & Fuels 23*, 6 (2009), 3227–3235.

[51] Dias Ribeiro, Mateus, Mendonça Bimbato, Alex, Araújo Zanardi, Maurício, Perrella Balestieri, José Antônio, and Schmidt, David P. Large-eddy simulation of the flow in a direct injection spark ignition engine using an open-source framework. *International Journal of Engine Research* (2020), 1468087420903622.

[52] Doudou, A. Turbulent flow study of an isothermal diesel spray injected by a common rail system. *Fuel 84*, 2-3 (2005), 287–298.

[53] Drohmann, Martin, and Carlberg, Kevin. The romes method for statistical modeling of reduced-order-model error. *SIAM/ASA Journal on Uncertainty Quantification 3*, 1 (2015), 116–145.

[54] Du, Chengjun, Andersson, Sven, and Andersson, Mats. The effect of cavitation on the estimation of fuel injection rates based on momentum flux measurements. *Fuel 238* (2019), 354–362.

[55] Duke, Daniel, Swantek, Andrew, Tilocco, Zak, Kastengren, Alan, Fezzaa, Kamel, Neroorkar, Kshitij, Moulai, Maryam, Powell, Christopher, and Schmidt, David. X-ray imaging of cavitation in diesel injectors. *SAE international Journal of Engines 7*, 2 (2014), 1003–1016.

[56] Duraisamy, Karthik, Iaccarino, Gianluca, and Xiao, Heng. Turbulence modeling in the age of data. *Annual Review of Fluid Mechanics 51* (2019), 357–377.

[57] Ehleskog, Rickard, Ochoterena, Raúl L, and Andersson, Sven. Effects of multiple injections on engine-out emission levels including particulate mass from an hsdi diesel engine. Tech. rep., SAE Technical Paper, 2007.

[58] Finlayson, Sam. Induction, Inductive Biases, and Infusing Knowledge into Learned Representations. https://sgfin.github.io/2020/06/22/Induction-Intro/InduBias, 2020. [Online; accessed 19-February-2020].

[59] Freno, Brian A, and Carlberg, Kevin T. Machine-learning error models for approximate solutions to parameterized systems of nonlinear equations. *Computer Methods in Applied Mechanics and Engineering 348* (2019), 250–296.

[60] Ganti, Himakar, Kamin, Manu, and Khare, Prashant. Design space exploration of turbulent multiphase flows using machine learning-based surrogate model. *Energies 13*, 17 (2020), 4565.

[61] Ganti, Himakar, and Khare, Prashant. Data-driven surrogate modeling ofmultiphase flows using machine learning techniques. *Computers & Fluids* (2020), 104626.

[62] Garcia-Oliver, Jose M, Pastor, Jose M, Pandal, Adrian, Trask, Nathaniel, Baldwin, Eli, and Schmidt, David P. Diesel spray CFD simulations based on the σ-y eulerian atomization model. *Atomization and Sprays 23*, 1 (2013).

[63] Gelbart, Michael A, Snoek, Jasper, and Adams, Ryan P. Bayesian optimization with unknown constraints. *arXiv preprint arXiv:1403.5607* (2014).

[64] Gentz, Gerald, Thelen, Bryce, Litke, Paul, Hoke, John, and Toulson, Elisa. Combustion visualization, performance, and CFD modeling of a pre-chamber turbulent jet ignition system in a rapid compression machine. *SAE International Journal of Engines 8*, 2 (2015), 538–546.

[65] Germano, Massimo, Piomelli, Ugo, Moin, Parviz, and Cabot, William H. A dynamic subgrid-scale eddy viscosity model. *Physics of Fluids A: Fluid Dynamics 3*, 7 (1991), 1760–1765.

[66] Gilpin, Leilani H, Bau, David, Yuan, Ben Z, Bajwa, Ayesha, Specter, Michael, and Kagal, Lalana. Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)* (2018), IEEE, pp. 80–89.

[67] Glimm, James, Grove, John W, Li, Xiao Lin, Shyue, Keh-ming, Zeng, Yanni, and Zhang, Qiang. Three-dimensional front tracking. *SIAM Journal on Scientific Computing 19*, 3 (1998), 703–727.

[68] Glorot, Xavier, and Bengio, Yoshua. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (2010), pp. 249–256.

[69] Goodfellow, Ian, Bengio, Yoshua, Courville, Aaron, and Bengio, Yoshua. *Deep learning*, vol. 1. MIT Press, 2016.

[70] Goodfellow, Ian, Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair, Sherjil, Courville, Aaron, and Bengio, Yoshua. Generative adversarial nets. *Advances in neural information processing systems 27* (2014), 2672–2680.

[71] Gopala, Vinay R, and van Wachem, Berend GM. Volume of fluid methods for immiscible-fluid and free-surface flows. *Chemical Engineering Journal 141*, 1-3 (2008), 204–221.

[72] Gorokhovski, Mikhael, and Herrmann, Marcus. Modeling primary atomization. *Annu. Rev. Fluid Mech. 40* (2008), 343–366.

[73] Gray, Donald D, and Giorgini, Aldo. The validity of the boussinesq approximation for liquids and gases. *International Journal of Heat and Mass Transfer 19*, 5 (1976), 545–551.

[74] Gui, Jie, Sun, Zhenan, Wen, Yonggang, Tao, Dacheng, and Ye, Jieping. A review on generative adversarial networks: Algorithms, theory, and applications. *arXiv preprint arXiv:2001.06937* (2020).

[75] Guo, Hengjie, Torelli, Roberto, Rodriguez, Abian Bautista, Tekawade, Aniket, Sforzo, Brandon, Powell, Christopher, and Som, Sibendu. Internal nozzle flow simulations of the ecn spray c injector under realistic operating conditions. Tech. rep., SAE Technical Paper, 2020.

[76] Gürsoy, Doga, De Carlo, Francesco, Xiao, Xianghui, and Jacobsen, Chris. Tomopy: a framework for the analysis of synchrotron tomographic data. *Journal of synchrotron radiation 21*, 5 (2014), 1188–1193.

[77] Habchi, Chawki, and Bruneaux, Gilles. Les and experimental investigation of diesel sprays. In *12th Triennial International Conference on Liquid Atomization and Spray Systems,(ICLASS 2012), Heidelberg, Germany, September* (2012), pp. 2–6.

[78] Haghshenas, Majid, Mitra, Peetak, Dal Santo, Niccolo, Dias Ribeiro, Mateus, Mitra, Shounak, and Schmidt, David. Les turbulence model with learnt closure; integration of DNN into a CFD solver. *Bulletin of the American Physical Society* (2020).

[79] Hanna, Botros N, Dinh, Nam T, Youngblood, Robert W, and Bolotnov, Igor A. Machine-learning based error prediction approach for coarse-grid computational fluid dynamics (CG-CFD). *Progress in Nuclear Energy 118* (2020), 103140.

[80] Hay, N, and Jones, PL. Comparison of the various correlations for spray penetration. Tech. rep., SAE Technical Paper, 1972.

[81] He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, and Sun, Jian. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision* (2015), pp. 1026–1034.

[82] He, Xin, Zhao, Kaiyong, and Chu, Xiaowen. Automl: A survey of the state-of-the-art. *Knowledge-Based Systems* (2020), 106622.

[83] Heister, SD. Plain orifice spray nozzles. In *Handbook of Atomization and Sprays*. Springer, 2011, pp. 625–645.

[84] Higgins, Brian S, Mueller, Charles J, and Siebers, Dennis L. Measurements of fuel effects on liquid-phase penetration in dl sprays. *SAE transactions* (1999), 630–643.

[85] Hiroyasu, Hiro, Arai, Masataka, and Shimizu, Masanori. Effect of internal flow conditions inside injector nozzles on jet breakup processes. *Recent advances in spray combustion: Spray atomization and drop burning phenomena. 1* (1996), 173–184.

[86] Hiroyasu, Hiroyuki. Spray breakup mechanism from the hole-type nozzle and its applications. *Atomization and Sprays 10*, 3-5 (2000).

[87] Höpken, Jens, and Mooney, Kyle G. The openfoam® technology primer.

[88] Hornik, Kurt. Approximation capabilities of multilayer feedforward networks. *Neural networks 4*, 2 (1991), 251–257.

[89] Hsiang, L-P, and Faeth, Gerard M. Near-limit drop deformation and secondary breakup. *International journal of multiphase flow 18*, 5 (1992), 635–652.

[90] Islam, Moni. *Numerical and experimental investigations of high-pressure diesel sprays*. PhD thesis, Imperial College London (University of London), 2002.

[91] Jacobs, Robert A. Increased rates of convergence through learning rate adaptation. *Neural networks 1*, 4 (1988), 295–307.

[92] Jacobsohn, Gabriel L, Mohapatra, Chinmoy K, Grover, Ronald O, Duke, Daniel J, and Schmidt, David P. Comparison of turbulence modeling methods for evaluating GDI sprays with transient needle motion. Tech. rep., SAE Technical Paper, 2019.

[93] Janas, Peter. *Large eddy simulation of in-cylinder phenomena in spark ignition engines*. PhD thesis, 2017.

[94] Jasak, H, and Weller, HG. Interface tracking capabilities of the inter-gamma differencing scheme. *Department of Mechanical Engineering, Imperial College of Science, Technology and Medicine* (1995).

[95] Jasak, Hrvoje, Jemcov, Aleksandar, Tukovic, Zeljko, et al. Openfoam: A C++ library for complex physics simulations. In *International workshop on coupled methods in numerical dynamics* (2007), vol. 1000, IUC Dubrovnik Croatia, pp. 1–20.

[96] Jeong, Jinhee, and Hussain, Fazle. On the identification of a vortex. *Journal of fluid mechanics 285* (1995), 69–94.

[97] Jiang, Chiyu Max, Esmaeilzadeh, Soheil, Azizzadenesheli, Kamyar, Kashinath, Karthik, Mustafa, Mustafa, Tchelepi, Hamdi A, Marcus, Philip, Anandkumar, Anima, et al. Meshfreeflownet: A physics-constrained deep continuous space-time super-resolution framework. *arXiv preprint arXiv:2005.01463* (2020).

[98] Jin, Haifeng, Song, Qingquan, and Hu, Xia. Auto-keras: An efficient neural architecture search system. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (2019), pp. 1946–1956.

[99] John, Heywood. Internal combustion engine fundamentals, 1988.

[100] Jordan, Michael I, and Mitchell, Tom M. Machine learning: Trends, perspectives, and prospects. *Science 349*, 6245 (2015), 255–260.

[101] Joshi, Ameya. Review of vehicle engine efficiency and emissions. Tech. rep., SAE Technical Paper, 2020.

[102] Kastengren, A, Powell, Christopher F, Liu, Zunping, Moon, Seoksu, Gao, Jian, Zhang, Xusheng, and Wang, Jin. Axial development of diesel sprays at varying ambient density. In *22nd Annual ILASS-Americas Conference, Cincinnati, OH, Paper* (2010), no. 106.

[103] Kastengren, Alan, Powell, Christopher F, Arms, Dohn, Dufresne, Eric M, Gibson, Harold, and Wang, Jin. The 7BM beamline at the APS: a facility for time-resolved fluid dynamics measurements. *Journal of synchrotron radiation 19*, 4 (2012), 654–657.

[104] Kastengren, Alan, Tilocco, F Zak, Duke, Daniel, Powell, Christopher F, Zhang, Xusheng, and Moon, Seoksu. Time-resolved x-ray radiography of sprays from engine combustion network spray a diesel injectors. *Atomization and Sprays 24*, 3 (2014).

[105] Kastengren, Alan L, Tilocco, F Zak, Powell, Christopher F, Manin, Julien, Pickett, Lyle M, Payri, Raul, and Bazyn, Tim. Engine combustion network (ecn): measurements of nozzle geometry and hydraulic behavior. *Atomization and Sprays 22*, 12 (2012), 1011–1052.

[106] Kim, Byungsoo, Azevedo, Vinicius C, Thuerey, Nils, Kim, Theodore, Gross, Markus, and Solenthaler, Barbara. Deep fluids: A generative network for parameterized fluid simulations. In *Computer Graphics Forum* (2019), vol. 38, Wiley Online Library, pp. 59–70.

[107] Kingma, Diederik P, and Ba, Jimmy. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[108] Kochkov, Dmitrii, Smith, Jamie A, Alieva, Ayya, Wang, Qing, Brenner, Michael P, and Hoyer, Stephan. Machine learning–accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences 118*, 21 (2021).

[109] Kutz, J Nathan. Deep learning in fluid dynamics. *Journal of Fluid Mechanics 814* (2017), 1–4.

[110] Kwa, Anna, Brenowitz, Noah, Bretherton, Christopher Stephen, Clark, Spencer, Henn, Brian M, McGibbon, Jeremy, Perkins, Andre, and Watt-Meyer, Oliver. Performance of a random forest parameterization in predicting the diurnal cycle of precipitation. In *AGU Fall Meeting 2020* (2020), AGU.

[111] Lebas, Romain, Blokkeel, Gregory, Beau, P-A, and Demoulin, F-X. Coupling vaporization model with the eulerian-lagrangian spray atomization (ELSA) model in diesel engine conditions. Tech. rep., SAE Technical Paper, 2005.

[112] Lebas, Romain, Menard, Thibault, Beau, Pierre-Arnaud, Berlemont, Alain, and Demoulin, François-Xavier. Numerical simulation of primary break-up and atomization: Dns and modelling study. *International Journal of Multiphase Flow 35*, 3 (2009), 247–260.

[113] LeCun, Yann, Bengio, Yoshua, and Hinton, Geoffrey. Deep learning. *nature 521*, 7553 (2015), 436–444.

[114] LeCun, Yann, Denker, John S, Solla, Sara A, Howard, Richard E, and Jackel, Lawrence D. Optimal brain damage. In *NIPs* (1989), vol. 2, Citeseer, pp. 598–605.

[115] Ledig, Christian, Theis, Lucas, Huszár, Ferenc, Caballero, Jose, Cunningham, Andrew, Acosta, Alejandro, Aitken, Andrew, Tejani, Alykhan, Totz, Johannes, Wang, Zehan, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 4681–4690.

[116] Lemmon, Eric W, Huber, Marcia L, and McLinden, Mark O. Nist reference fluid thermodynamic and transport properties–refprop, 2002.

[117] Leonard, Athony. Energy cascade in large-eddy simulations of turbulent fluid flows. In *Advances in geophysics*, vol. 18. Elsevier, 1975, pp. 237–248.

[118] Leroux, Samuel, Dumouchel, Christophe, and Ledoux, Michel. The stability curve of newtonian liquid jets. *Atomization and sprays 6*, 6 (1996).

[119] Leung, Timothy F, Groth, Clinton P, and Hu, John. Evaluation of an eulerian-lagrangian spray atomization (ELSA) model for nozzle flow: Modeling of coupling between dense and disperse regions. In *47th AIAA Thermophysics Conference* (2017), p. 4352.

[120] Li, Li, Hoyer, Stephan, Pederson, Ryan, Sun, Ruoxi, Cubuk, Ekin D, Riley, Patrick, and Burke, Kieron. Kohn-sham equations as regularizer: building prior knowledge into machine-learned physics. *arXiv preprint arXiv:2009.08551* (2020).

[121] Lin, SP, and Reitz, RD. Drop and spray formation from a liquid jet. *Annual review of fluid mechanics 30*, 1 (1998), 85–105.

[122] Lin, Sung-Piau. *Breakup of liquid sheets and jets*. Cambridge university press, 2003.

[123] Ling, Julia, Kurzawski, Andrew, and Templeton, Jeremy. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics 807* (2016), 155–166.

[124] Liu, Chenxi, Zoph, Barret, Neumann, Maxim, Shlens, Jonathon, Hua, Wei, Li, Li-Jia, Fei-Fei, Li, Yuille, Alan, Huang, Jonathan, and Murphy, Kevin. Progressive neural architecture search. In *Proceedings of the European Conference on Computer Vision (ECCV)* (2018), pp. 19–34.

[125] Lund, TS. The use of explicit filters in large eddy simulation. *Computers & Mathematics with Applications 46*, 4 (2003), 603–616.

[126] Maaten, Laurens van der, and Hinton, Geoffrey. Visualizing data using t-sne. *Journal of machine learning research 9*, Nov (2008), 2579–2605.

[127] Magnotti, Gina, Battistoni, Michele, Saha, Kaushik, and Som, Sibendu. Predicting cavitation erosion propensity and severity in fuel injection systems. *Bulletin of the American Physical Society 63* (2018).

[128] Magnotti, Gina M, Battistoni, Michele, Saha, Kaushik, and Som, Sibendu. Influence of turbulence and thermophysical fluid properties on cavitation erosion predictions in channel flow geometries. *SAE International Journal of Advances and Current Practices in Mobility 1*, 2019-01-0290 (2019), 691–705.

[129] Malalasekera, W, and Versteeg, HK. *An introduction to computational fluid dynamics: the finite volume method*. Pearson Prentice Hall Upper Saddle River, NJ, 2007.

[130] Manin, J, Pickett, LM, and Yasutomi, K. Stereoscopic high-speed microscopy to understand transient internal flow processes in high-pressure nozzles. *Experimental Thermal and Fluid Science 114* (2020), 110027.

[131] Manin, Julien, Yasutomi, Koji, and Pickett, Lyle M. Transient cavitation in transparent diesel injectors. Tech. rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2018.

[132] Martinez-Cantin, Ruben. Bayesopt: A bayesian optimization library for nonlinear optimization, experimental design and bandits. *The Journal of Machine Learning Research 15*, 1 (2014), 3735–3739.

[133] Matusik, Katarzyna E, Duke, Daniel J, Kastengren, Alan L, Sovis, Nicholas, Swantek, Andrew B, and Powell, Christopher F. High-resolution x-ray tomography of engine combustion network diesel injectors. *International Journal of Engine Research 19*, 9 (2018), 963–976.

[134] Maulik, Romit, San, Omer, Rasheed, Adil, and Vedula, Prakash. Subgrid modelling for two-dimensional turbulence using neural networks. *Journal of Fluid Mechanics 858* (2019), 122–144.

[135] McCarthy, MJ, and Molloy, NA. Review of stability of liquid jets and the influence of nozzle design. *The Chemical Engineering Journal 7*, 1 (1974), 1–20.

[136] Meldi, M, Lucor, Didier, and Sagaut, P. Is the smagorinsky coefficient sensitive to uncertainty in the form of the energy spectrum? *Physics of Fluids 23*, 12 (2011), 125109.

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

[138] Milan, Petro Junior, Torelli, Roberto, Lusch, B, and Magnotti, GM. Data-driven model reduction of multiphase flow in a single-hole automotive injector. *Atomization and Sprays 30*, 6 (2020).

[139] Mitchell, Tom M. *The need for biases in learning generalizations*. Department of Computer Science, Laboratory for Computer Science Research ..., 1980.

[140] Mitra, Peetak, Haghshenas, Majid, Dal Santo, Niccolo, Daly, Conor, Mitra, Shounak, and Schmidt, David. Towards building robust neural network models for fluid simulations. *Bulletin of the American Physical Society* (2020).

[141] Mitra, Peetak, Matusik, Katarzyna, Duke, Daniel, Srivastava, Priyesh, Yasutomi, Koji, Manin, Julien, Pickett, Lyle, Powell, Christopher F, Arienti, Marco, Baldwin, Eli, et al. Identification and characterization of steady spray conditions in convergent, single-hole diesel injectors. Tech. rep., SAE Technical Paper, 2019.

[142] Mitra, Peetak, Venkatesan, Vaidehi, Jangid, Nomit, Nambiar, Ashwati, Kumar, Dhananjay, Roa, Vignesh, Santo, Niccolo Dal, Haghshenas, Majid, Mitra, Shounak, and Schmidt, David. Network compression for machine-learnt fluid simulations. *arXiv preprint arXiv:2103.00754* (2021).

[143] Mjolsness, Eric, and DeCoste, Dennis. Machine learning for science: state of the art and future prospects. *science 293*, 5537 (2001), 2051–2055.

[144] Mockus, Jonas, Tiesis, Vytautas, and Zilinskas, Antanas. The application of bayesian methods for seeking the extremum. *Towards global optimization 2*, 117-129 (1978), 2.

[145] Mohapatra, Chinmoy, Jacobsohn, Gabriel, Baldwin, Eli, and Schmidt, David. Modeling sealing in transient injector simulations. In *Fluids Engineering Division Summer Meeting* (2017), vol. 58042, American Society of Mechanical Engineers, p. V01AT04A009.

[146] Mohapatra, Chinmoy K, Schmidt, David P, Sforozo, Brandon A, Matusik, Katarzyna E, Yue, Zongyu, Powell, Christopher F, Som, Sibendu, Mohan, Balaji, Im, Hong G, Badra, Jihad, et al. Collaborative investigation of the internal flow and near-nozzle flow of an eight-hole gasoline injector (engine combustion network spray g). *International Journal of Engine Research* (2020), 1468087420918449.

[147] Moosavi, Azam, Ştefănescu, Răzvan, and Sandu, Adrian. Multivariate predictions of local reduced-order-model errors and dimensions. *International Journal for Numerical Methods in Engineering 113*, 3 (2018), 512–533.

[148] Mueller, Charles J, Nilsen, Christopher W, Ruth, Daniel J, Gehmlich, Ryan K, Pickett, Lyle M, and Skeen, Scott A. Ducted fuel injection: A new approach for lowering soot emissions from direct-injection engines. *Applied energy 204* (2017), 206–220.

[149] Mukundan, Anirudh Asuri, Tretola, Giovanni, Ménard, Thibaut, Herrmann, Marcus, Navarro-Martinez, Salvador, Vogiatzaki, Konstantina, de Motta, Jorge César Brändle, and Berlemont, Alain. DNS and LES of primary atomization of turbulent liquid jet injection into a gaseous crossflow environment. *Proceedings of the Combustion Institute* (2020).

[150] Musculus, Mark PB, and Kattke, Kyle. Entrainment waves in diesel jets. *SAE International Journal of Engines 2*, 1 (2009), 1170–1193.

[151] Naber, Jeffrey D, and Siebers, Dennis L. Effects of gas density and vaporization on penetration and dispersion of diesel sprays. *SAE transactions* (1996), 82–111.

[152] Neroorkar, KD, Mitcham, CE, Plazas, A, Grover, T, and Schmidt, D. Simulations and analysis of fuel flow in an injector including transient needle effects. In *ILASS-Americas 24th Annual Conf Liquid Atomization and Spray Systems* (2012).

[153] Nicoud, Franck, and Ducros, Frédéric. Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow, turbulence and Combustion 62*, 3 (1999), 183–200.

[154] Olsson, Elin, and Kreiss, Gunilla. A conservative level set method for two phase flow. *Journal of computational physics 210*, 1 (2005), 225–246.

[155] Örley, Felix, Hickel, Stefan, Schmidt, Steffen J, and Adams, Nikolaus A. Large-eddy simulation of turbulent, cavitating fuel flow inside a 9-hole diesel injector including needle movement. *International Journal of Engine Research 18*, 3 (2017), 195–211.

[156] Pai, Madhusudan G, and Subramaniam, Shankar. A comprehensive probability density function formalism for multiphase flows. *Journal of Fluid Mechanics 628* (2009), 181.

[157] Papadopoulos, Nikolaos, and Aleiferis, Pavlos. Numerical modelling of the in-nozzle flow of a diesel injector with moving needle during and after the end of a full injection event. *SAE International Journal of Engines 8*, 5 (2015), 2285–2302.

[158] Park, Cheolwoong, Kook, Sanghoon, and Bae, Choongsik. Effects of multiple injections in a hsdi diesel engine equipped with common rail injection system. Tech. rep., SAE Technical Paper, 2004.

[159] Pastor, José V, López, J Javier, García, José M, and Pastor, José M. A 1D model for the description of mixing-controlled inert diesel sprays. *Fuel 87*, 13-14 (2008), 2871–2885.

[160] Pastor, Jose V, Payri, Raul, Garcia-Oliver, Jose M, and Briceno, Francisco J. Analysis of transient liquid and vapor phase penetration for diesel sprays under variable injection conditions. *Atomization and Sprays 21*, 6 (2011), 503–520.

[161] Payri, R, Gimeno, J, Cuisano, J, and Arco, Javier. Hydraulic characterization of diesel engine single-hole injectors. *Fuel 180* (2016), 357–366.

[162] Payri, Raul, Gimeno, Jaime, Bardi, Michele, and Plazas, Alejandro H. Study liquid length penetration results obtained with a direct acting piezo electric injector. *Applied energy 106* (2013), 152–162.

[163] Payri, Raul, Viera, Juan P, Gopalakrishnan, Venkatesh, and Szymkowicz, Patrick G. The effect of nozzle geometry over internal flow and spray formation for three different fuels. *Fuel 183* (2016), 20–33.

[164] Pickett, Lyle M, Manin, Julien, Genzale, Caroline L, Siebers, Dennis L, Musculus, Mark PB, and Idicheria, Cherian A. Relationship between diesel fuel spray vapor penetration/dispersion and local fuel mixture fraction. *SAE International Journal of Engines 4*, 1 (2011), 764–799.

[165] Pickett, Lyle M, Manin, Julien, Payri, Raul, Bardi, Michele, and Gimeno, Jaime. Transient rate of injection effects on spray development. Tech. rep., SAE Technical Paper, 2013.

[166] Pickett, Lyle M, and Siebers, Dennis L. An investigation of diesel soot formation processes using micro-orifices. *Proceedings of the Combustion Institute 29*, 1 (2002), 655–662.

[167] Pickett, Lyle M, Siebers, Dennis L, and Idicheria, Cherian A. Relationship between ignition processes and the lift-off length of diesel fuel jets. *SAE transactions* (2005), 1714–1731.

[168] Pitsch, Heinz. Large-eddy simulation of turbulent combustion. *Annu. Rev. Fluid Mech. 38* (2006), 453–482.

[169] Pope, Stephen B. Turbulent flows, 2001.

[170] Portwood, Gavin D, Mitra, Peetak P, Ribeiro, Mateus Dias, Nguyen, Tan Minh, Nadiga, Balasubramanya T, Saenz, Juan A, Chertkov, Michael, Garg, Animesh, Anandkumar, Anima, Dengel, Andreas, et al. Turbulence forecasting via neural ode. *arXiv preprint arXiv:1911.05180* (2019).

[171] Portwood, Gavin D, Nadiga, Balasubramanya T, Saenz, Juan A, and Livescu, Daniel. Analysis and interpretation of out-performing neural network residual flux models. *arXiv preprint arXiv:2004.07207* (2020).

[172] Pratama, Raditya Hendra, Sou, Akira, Katsui, Tokihiro, and Nishio, Shigeru. String cavitation in a fuel injector. *Atomization Sprays 27*, 3 (2017), 189–205.

[173] Qian, J. Universal equilibrium range of turbulence. *The Physics of Fluids 27*, 9 (1984), 2229–2233.

[174] Quan, Shaoping, Lou, Jing, and Stone, Howard A. Interactions between two deformable droplets in tandem subjected to impulsive acceleration by surrounding flows. *Journal of fluid mechanics 684* (2011), 384.

[175] Rachakonda, Sampath K, Paydarfar, Arman, and Schmidt, David P. Prediction of spray collapse in multi-hole gasoline direct-injection fuel injectors. *International Journal of Engine Research 20*, 1 (2019), 18–33.

[176] Rachakonda, Sampath K, Wang, Yue, Grover Jr, Ronald O, Moulai, Maryam, Baldwin, Eli, Zhang, Gaoming, Parrish, Scott, Diwakar, Ramachandra, Kuo, Tang-Wei, and Schmidt, David P. A computational approach to predict external spray characteristics for flashing and cavitating nozzles. *International Journal of Multiphase Flow 106* (2018), 21–33.

[177] Raissi, Maziar, Perdikaris, Paris, and Karniadakis, George E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics 378* (2019), 686–707.

[178] Raissi, Maziar, Yazdani, Alireza, and Karniadakis, George Em. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science 367*, 6481 (2020), 1026–1030.

[179] Rasmussen, Carl Edward, and Nickisch, Hannes. Gaussian processes for machine learning (gpml) toolbox. *The Journal of Machine Learning Research 11* (2010), 3011–3015.

[180] Rauser, Florian, Korn, Peter, and Marotzke, Jochem. Predicting goal error evolution from near-initial-information: A learning algorithm. *Journal of computational physics 230*, 19 (2011), 7284–7299.

[181] Reitz, RD, Ogawa, H, Payri, R, Fansler, T, Kokjohn, S, Moriyoshi, Y, Agarwal, AK, Arcoumanis, D, Assanis, D, Bae, C, et al. Ijer editorial: The future of the internal combustion engine, 2020.

[182] Reitz, Rolf D, and Chigier, Norman. Regimes of jet breakup and breakup mechanisms (physical aspects). *Prog. Astronaut. Aeronaut 166* (1996), 109.

[183] Reitz, Rolf D, and Diwakar, R. Effect of drop breakup on fuel sprays. *SAE transactions* (1986), 218–227.

[184] Rider, William, and Kothe, Douglas. Stretching and tearing interface tracking methods. In *12th Computational Fluid Dynamics Conference* (1995), p. 1717.

[185] Rudd, Robert E, and Broughton, Jeremy Q. Coarse-grained molecular dynamics and the atomic limit of finite elements. *Physical review B 58*, 10 (1998), R5893.

[186] Rulli, Federico, Barbato, Alessio, Fontanesi, Stefano, and d'Adamo, Alessandro. Large eddy simulation analysis of the turbulent flow in an optically accessible internal combustion engine using the overset mesh technique. *International Journal of Engine Research* (2020), 1468087419896469.

[187] Saha, Kaushik, Quan, Shaoping, Battistoni, Michele, Som, Sibendu, Senecal, PK, and Pomraning, Eric. Coupled eulerian internal nozzle flow and lagrangian spray simulations for GDI systems. Tech. rep., SAE Technical Paper, 2017.

[188] Saha, Kaushik, Srivastava, Priyesh, Quan, Shaoping, Senecal, PK, Pomraning, Eric, and Som, Sibendu. Modeling the dynamic coupling of internal nozzle flow and spray formation for gasoline direct injection applications. Tech. rep., SAE Technical Paper, 2018.

[189] Salvador, Francisco Javier, Gimeno, Jaime, Pastor, José Manuel, and Martí-Aldaraví, Pedro. Effect of turbulence model and inlet boundary condition on the diesel spray behavior simulated by an eulerian spray atomization (esa) model. *International Journal of Multiphase Flow 65* (2014), 108–116.

[190] Schlichting, Hermann, and Gersten, Klaus. *Boundary-layer theory*. Springer, 2016.

[191] Schmidhuber, Jürgen. Deep learning in neural networks: An overview. *Neural networks 61* (2015), 85–117.

[192] Schmidt, David P. *Cavitation in diesel fuel injector nozzles*. PhD thesis, University of Wisconsin–Madison, 1997.

[193] Schmidt, David P, and Corradini, ML. The internal flow of diesel fuel injector nozzles: a review. *International Journal of Engine Research 2*, 1 (2001), 1–22.

[194] Schmidt, David P, Corradini, ML, and Rutland, Christopher J. A two-dimensional, non-equilibrium model of flashing nozzle flow. In *3rd ASME/JSME Joint Fluids Engineering Conference* (1999), vol. 208, p. 616.

[195] Schmidt, David P, Gopalakrishnan, S, and Jasak, Hrvoje. Multi-dimensional simulation of thermal non-equilibrium channel flow. *International journal of multiphase flow 36*, 4 (2010), 284–292.

[196] Senecal, PK, and Leach, Felix. Diversity in transportation: Why a mix of propulsion technologies is the way forward for the future fleet. *Results in Engineering 4* (2019), 100060.

[197] Senecal, PK, Pomraning, E, Richards, KJ, and Som, S. Grid-convergent spray models for internal combustion engine CFD simulations. In *Internal Combustion Engine Division Fall Technical Conference* (2012), vol. 55096, American Society of Mechanical Engineers, pp. 697–710.

[198] Senecal, PK, Pomraning, E, Richards, KJ, and Som, S. Grid-convergent spray models for internal combustion engine computational fluid dynamics simulations. *Journal of Energy Resources Technology 136*, 1 (2014).

[199] Senecal, PK, Richards, KJ, Pomraning, E, Yang, T, Dai, MZ, McDavid, RM, Patterson, MA, Hou, S, and Shethaji, T. 2007-01-0159 a new parallel cut-cell cartesian CFD code for rapid grid generation applied to in-cylinder diesel engine simulations. *SAE SP 2125* (2007), 95.

[200] Sforzo, Brandon A, Matusik, Katarzyna E, Powell, Christopher F, Kastengren, Alan L, Daly, Shane, Skeen, Scott A, Cenker, Emre, Pickett, Lyle M, Crua, Cyril, and Manin, Julien. Fuel nozzle geometry effects on cavitation and spray behavior at diesel engine conditions. Tech. rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2018.

[201] Shahangian, Navid, Sharifian, Leila, Miyagawa, Jun, Bergamini, Stefano, Uehara, Kazuhiro, Noguchi, Yasushi, Marti-Aldaravi, Pedro, Martínez, María, and Payri, Raul. Nozzle flow and spray development one-way coupling methodology for a multi-hole GDI injector. Tech. rep., SAE Technical Paper, 2019.

[202] Shankar, Varun, Portwood, Gavin, Mohan, Arvind, Mitra, Peetak, Viswanathan, Venkat, and Schmidt, David. Rapid spatiotemporal turbulence modeling with convolutional neural ODEs. *Bulletin of the American Physical Society* (2020).

[203] Siebers, Dennis, Higgins, Brian, and Pickett, Lyle. Flame lift-off on direct-injection diesel fuel jets: oxygen concentration effects. *Sae Transactions* (2002), 1490–1509.

[204] Siebers, Dennis L. Liquid-phase fuel penetration in diesel sprays. *SAE transactions* (1998), 1205–1227.

[205] Siebers, Dennis L. Scaling liquid-phase fuel penetration in diesel sprays based on mixing-limited vaporization. *SAE transactions* (1999), 703–728.

[206] Sirignano, Justin, MacArt, Jonathan F, and Freund, Jonathan B. Dpm: A deep learning PDE augmentation method with application to large-eddy simulation. *Journal of Computational Physics 423* (2020), 109811.

[207] Skeen, Scott, Manin, Julien, and Pickett, Lyle M. Visualization of ignition processes in high-pressure sprays with multiple injections of n-dodecane. *SAE International Journal of Engines 8*, 2 (2015), 696–715.

[208] Smagorinsky, Joseph. General circulation experiments with the primitive equations: I. the basic experiment. *Monthly weather review 91*, 3 (1963), 99–164.

[209] Snoek, Jasper, Larochelle, Hugo, and Adams, Ryan P. Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems* (2012), pp. 2951–2959.

[210] Som, Sibendu, Ramirez, Anita I, Longman, Douglas E, and Aggarwal, Suresh K. Effect of nozzle orifice geometry on spray, combustion, and emission characteristics under diesel engine conditions. *Fuel 90*, 3 (2011), 1267–1276.

[211] Sommerfeld, Arnold. *Mechanics of deformable bodies: Lectures on theoretical physics, Vol. 2*, vol. 2. Elsevier, 2016.

[212] Soteriou, Celia, Andrews, Richard, and Smith, Mark. Direct injection diesel sprays and the effect of cavitation and hydraulic flip on atomization. *SAE transactions* (1995), 128–153.

[213] Sphicas, Panos, Pickett, Lyle M, Skeen, Scott, Frank, Jonathan, Lucchini, Tommaso, Sinoir, David, D'Errico, Gianluca, Saha, Kaushik, and Som, Sibendu. A comparison of experimental and modeled velocity in gasoline direct-injection sprays with plume interaction and collapse. *SAE International Journal of Fuels and Lubricants 10*, 1 (2017), 184–201.

[214] Spiegel, Edward A, and Veronis, G. On the boussinesq approximation for a compressible fluid. *The Astrophysical Journal 131* (1960), 442.

[215] Srinivas, Niranjan, Krause, Andreas, Kakade, Sham M, and Seeger, Matthias. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995* (2009).

[216] Srivastava, Nitish, Hinton, Geoffrey, Krizhevsky, Alex, Sutskever, Ilya, and Salakhutdinov, Ruslan. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research 15*, 1 (2014), 1929–1958.

[217] Stevens, Rick, Taylor, Valerie, Nichols, Jeff, Maccabe, Arthur Barney, Yelick, Katherine, and Brown, David. Ai for science. Tech. rep., Argonne National Lab.(ANL), Argonne, IL (United States), 2020.

[218] Strek, Piotr, Duke, Daniel, Swantek, Andrew, Kastengren, Alan, Powell, Christopher F, and Schmidt, David P. X-ray radiography and CFD studies of the spray G injector. Tech. rep., SAE Technical Paper, 2016.

[219] Strotos, George, Koukouvinis, Phoevos, Theodorakakos, Andreas, Gavaises, Manolis, and Bergeles, George. Transient heating effects in high pressure diesel injector nozzles. *International Journal of Heat and Fluid Flow 51* (2015), 257–267.

[220] Su, TF, Patterson, MA, Reitz, Rolf D, and Farrell, PV. Experimental and numerical studies of high pressure multiple injection sprays. *SAE transactions* (1996), 1281–1292.

[221] Sutskever, Ilya, Martens, James, Dahl, George, and Hinton, Geoffrey. On the importance of initialization and momentum in deep learning. In *International conference on machine learning* (2013), pp. 1139–1147.

[222] Taira, Kunihiko, Hemati, Maziar S, Brunton, Steven L, Sun, Yiyang, Duraisamy, Karthik, Bagheri, Shervin, Dawson, Scott TM, and Yeh, Chi-An. Modal analysis of fluid flows: Applications and outlook. *AIAA journal 58*, 3 (2020), 998–1022.

[223] Tanaka, Hidenori, Nayebi, Aran, Maheswaranathan, Niru, McIntosh, Lane, Baccus, Stephen A, and Ganguli, Surya. From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction. *arXiv preprint arXiv:1912.06207* (2019).

[224] Tekawade, A, Mitra, P, Sforzo, BA, Matusik, KE, Kastengren, AL, Schmidt, DP, and Powell, CF. A comparison between CFD and 3D x-ray diagnostics of internal flow in a cavitatingdiesel injector nozzle. In *ILASS-Americas 30th Annual Conference on Liquid Atomization and Spray Systems, Tempe, AZ* (2019).

[225] Tekawade, A, Sforzo, BA, Matusik, KE, Kastengren, AL, and Powell, CF. 3d imaging of cavitating flow in a diesel injector at practical conditions using x-ray micro-ct. *30th ILASS Americas* (2019).

[226] Tekawade, Aniket, Sforzo, Brandon A, Matusik, Katarzyna E, Fezzaa, Kamel, Kastengren, Alan L, and Powell, Christopher F. Time-resolved 3d imaging of two-phase fluid flow inside a steel fuel injector using synchrotron x-ray tomography. *Scientific Reports 10*, 1 (2020), 1–9.

[227] Tennekes, H. Eulerian and lagrangian time microscales in isotropic turbulence. *Journal of Fluid Mechanics 67*, 3 (1975), 561–567.

[228] Tieleman, T, and Hinton, G. Divide the gradient by a running average of its recent magnitude. coursera: Neural networks for machine learning. *Technical Report.* (2017).

[229] Tryggvason, Grétar, Bunner, Bernard, Esmaeeli, Asghar, Juric, Damir, Al-Rawahi, N, Tauber, W, Han, J, Nas, S, and Jan, Y-J. A front-tracking method for the computations of multiphase flow. *Journal of computational physics 169*, 2 (2001), 708–759.

[230] Um, Kiwon, Brand, Robert, Holl, Philipp, Thuerey, Nils, et al. Solver-in-the-loop: Learning from differentiable physics to interact with iterative PDE-solvers. *arXiv preprint arXiv:2007.00016* (2020).

[231] Um, Kiwon, Hu, Xiangyu, and Thuerey, Nils. Liquid splash modeling with neural networks. In *Computer Graphics Forum* (2018), vol. 37, Wiley Online Library, pp. 171–182.

[232] Vallet, Ariane, and Borghi, Roland. Modélisation eulerienne de l'atomisation d'un jet liquide. *Comptes Rendus de l'Académie des Sciences-Series IIB-Mechanics-Physics-Astronomy 327*, 10 (1999), 1015–1020.

[233] Van Der Maaten, Laurens. Barnes-hut-sne. *arXiv preprint arXiv:1301.3342* (2013).

[234] Versteeg, Henk Kaarle, and Malalasekera, Weeratunge. *An introduction to computational fluid dynamics: the finite volume method*. Pearson education, 2007.

[235] Wakuri, Yutaro, Fujii, Masaru, Amitani, Tatsuo, and Tsuneya, Reijiro. Studies on the penetration of fuel spray in a diesel engine. *Bulletin of JSME 3*, 9 (1960), 123–130.

[236] Wang, Xiangang, Huang, Zuohua, Kuti, Olawole Abiola, Zhang, Wu, and Nishida, Keiya. Experimental and analytical study on biodiesel and diesel spray characteristics under ultra-high injection pressure. *International journal of heat and fluid flow 31*, 4 (2010), 659–666.

[237] Wang, Yue, Lee, Won Geun, Reitz, Rolf D, and Diwakar, Ramachandra. Numerical simulation of diesel sprays using an eulerian-lagrangian spray and atomization (ELSA) model coupled with nozzle flow. Tech. rep., SAE Technical Paper, 2011.

[238] Wang, Zihan, Swantek, Andrew, Scarcelli, Riccardo, Duke, Daniel, Kastengren, Alan, Powell, Christopher F, Som, Sibendu, Reese, Ronald, Freeman, Kevin, and Zhu, York. LES of diesel and gasoline sprays with validation against x-ray radiography data. *SAE International Journal of Fuels and Lubricants 8*, 1 (2015), 147–159.

[239] Watt-Meyer, Oliver, Brenowitz, Noah Domino, Clark, Spencer Koncius, Henn, Brian, Kwa, Anna, McGibbon, Jeremy J, Perkins, Walter A, and Bretherton, Christopher S. Correcting weather and climate models by machine learning nudged historical simulations.

[240] Williams, Christopher, and Rasmussen, Carl. Gaussian processes for regression. *Advances in neural information processing systems 8* (1995), 514–520.

[241] Wörner, Martin. Numerical modeling of multiphase flows in microfluidics and micro process engineering: a review of methods and applications. *Microfluidics and nanofluidics 12*, 6 (2012), 841–886.

[242] Wu, Jin-Long, Xiao, Heng, and Paterson, Eric. Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework. *Physical Review Fluids 3*, 7 (2018), 074602.

[243] Wu, K-J, Reitz, RD, and Bracco, FV. Measurements of drop size at the spray edge near the nozzle in atomizing liquid jets. *The Physics of fluids 29*, 4 (1986), 941–951.

[244] Xue, Qingluan, Som, Sibendu, Senecal, Peter K, and Pomraning, E. Large eddy simulation of fuel-spray under non-reacting ic engine conditions. *Atomization and Sprays 23*, 10 (2013).

[245] Yasutomi, Koji, Hwang, Joonsik, Pickett, Lyle M, Sforzo, Brandon, Matusik, Katarzyna, and Powell, Christopher F. Transient internal nozzle flow in transparent multi-hole diesel injector. Tech. rep., SAE Technical Paper, 2020.

[246] Zainuddin, Zarita, and Pauline, Ong. Function approximation using artificial neural networks. *WSEAS Transactions on Mathematics 7*, 6 (2008), 333–338.

[247] Zhang, Hongyuan, Bavandla, Krishna C, Gao, Xiang, Gao, Jianfeng, Yi, Ping, and Yang, Suo. Optimization and uncertainty quantification of spray break-up submodel with regularized multi-task neural nets. In *AIAA Scitech 2020 Forum* (2020), p. 0909.

[248] Zhang, Ze Jia, and Duraisamy, Karthikeyan. Machine learning methods for data-driven turbulence modeling. In *22nd AIAA Computational Fluid Dynamics Conference* (2015), p. 2460.