EXTRACTING AND ANALYZING ATOMIZATION PHYSICS FROM HIGH-FIDELITY SIMULATIONS

by

Brendan Vilhelm Christensen

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mechanical Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

May, 2025

ACKNOWLEDGEMENTS

I am incredibly lucky to have encountered and interacted with so many amazing individuals in my time as a graduate student at Montana State. First, I want to thank my committee, Dr. Mark Owkes, Dr. Erick Johnson, Dr. Joe Seymour, and Dr. Yaofa Li. Additionally, this work was funded by the National Science Foundation under Grant Nos. 2152737 and 1749779.

When I started graduate school, I had an Earth science degree and knew nothing about coding and next to nothing about fluids. So, when Dr. Mark Owkes took me on as a student to study computational fluid dynamics, he was definitely taking a risk. I am so grateful for his patient and steady presence, which has fostered me into a proper scientist and researcher.

My labmates made the Barnard basement a fun and productive environment. Julian for nerding-out with me about cooking and fishing. Venkat for guiding me through the first years of my program, for getting me out of the lab and into the mountains, and some of the best food I've ever had. Jacob, Allison, Patrick, and Benjamin introduced me to the Bozeman running community. The simple act of inviting me out to run has changed my life. In the last semester of my Ph.D. I got married to my incredible wife, McKaelee. She is the most incredible human that I know and her unwaivering support has provided me with endless possibilities. I live in paradise and am pursuing my passions because of her. Thank you for everything you do, McKaelee, I am so excited about all of our future adventures.

Finally, my parents have given me every opportunity to succeed. They set an example of integrity and dedication, showing me that trying is always worth it.

TABLE OF CONTENTS

1. INTRODUCTION 1

Motivation 1
Background - Experimental Work 2
Background - Numerical Work 8
Background - Reduced-order Models 13
Contributions 16
Organization of this document 17

2. EFFICIENT EXTRACTION OF ATOMIZATION PROCESSES FROM HIGH-FIDELITY SIMULATIONS 19

Contribution of Authors and Co-Authors 19
Manuscript Information 20
Abstract 21
Introduction 21
Methods 25
Computational Platform 25
Breakup and Coalescence Event Identification 26
Identification Numbers 26
Coalescence Identification 28
Breakup Identification 29
Addressing Fictitious Events 29
Data Extraction 32
Graphical Database 33
Application of Extraction Tool on Diesel Jet 38
Simulation Setup 38
Addressing Fictitious Events Confirmation 38
Secondary Atomization Analysis 40
Coalescence Analysis 41
Local Flow Field Statistics 42
Atomization Evolution 44
Droplet Shape and Size Evolution 44
Time Evolution 44
Computational Cost 46
Conclusions 48
Acknowledgements 49

TABLE OF CONTENTS – CONTINUED

3. NUMERICAL STUDY ON NEWTONIAN AND NON-NEWTONIAN DROPLET AEROBREAKUP MECHANISMS 50

Contribution of Authors and Co-Authors 50
Manuscript Information 51
Abstract 52
Introduction 52
Background 55
Statement of Purpose 56
Methodology 57
Numerical Methods 57
Atomization Simulation Statistics Extraction Tool 59
Shear-Thinning Viscosity 60
Gas Domain 60
Simulation Configurations 61
Resolution Considerations 62
Characteristic Time 63
Comparison to Experimental Studies 64
Breakup Morphologies 65
Large-Scale Deformation 66
Early Instability Formation, Wavelengths, and Stripping 67
Instability Wavelength Evolution 72
Late-stage Primary Breakup Mechanisms 77
Secondary Atomization 78
Breakup Dynamics from ASSET 80
Droplet Genealogy Analysis 81
Primary Breakup Statistics 82
Secondary Breakup Cascade Analysis 83
Effect of Shear-Thinning Viscosity 88
Some Comments on Grid Effects and the Practical Application of DNS 92
Conclusions 95

4. A COMPUTATIONAL FRAMEWORK FOR SIMULATING AND ANALYZING ATOMIZATION IN ELECTRICALLY CHARGED DIESEL-TYPE JETS 98

Contribution of Authors and Co-Authors 98
Manuscript Information 99
Abstract 100
Introduction 101

TABLE OF CONTENTS – CONTINUED

Mathematical Formulation 105
Governing Equations 105
Interface Boundary Conditions 107
Numerical Approach 109
Data Extraction 110
Computational Domain and Simulations 112
Results 113
Uncharged and Charged Simulation Comparison 116
Charged Simulations Comparison 117
Primary Atomization Analysis 118
Electric Charge Density 121
Conclusions 123

5. A GEOMETRIC AND TOPOLOGICAL DATA ANALYSIS OF ATOMIZATION FOR MODEL DEVELOPMENT 126

Contribution of Authors and Co-Authors 126
Manuscript Information 127
Abstract 128
Introduction 128
Background - Topological Data Analysis 131
Background - Atomization Modeling 133
Project Outline and Scope 135
Methods 136
Computational Platform 136
ASSET Methodology 137
Event Identification 138
Geometry Extraction 139
Velocity Calculation 140
The Directionally Integrated Weighted Euler Characteristic Transform (DIWECT) 141
Results 148
Simulation Setup 148
Droplet Matching 150
Clustering Results 150
Conclusions 159

6. LOCAL WEBER NUMBER CHARACTERIZATION 162

TABLE OF CONTENTS – CONTINUED

7. CONCLUSIONS AND OUTLOOK 166

Summary 166
Developing ASSET 166
Applying ASSET 167
Extracting Data for Model Development 168
Future Work 169

REFERENCES CITED 172

APPENDICES 191

APPENDIX A : Liquid Identification and Coalescence Code 192
APPENDIX B : Neo4j Data Input 195
APPENDIX C : Python Neo4j Driver Script 198

LIST OF TABLES

Table Page

1. Table 1 Non-dimensional parameters used in the diesel jet simulation. 39
2. Table 2 The average number of coalescence events between primary breakup and the breakup stage is listed in the second column. The third column displays the ratio of coalescence events to breakup events along breakup paths. 42
3. Table 3 The cost of each process within the tool. Displayed as the average percent of time per timestep. Sampled from 175 timesteps at the end of the simulation. 47
4. Table 4 Non-dimensional parameters for the simulations. CAD is cells across diameter, simulations are identified as W for water and ST for shear-thinning, 1,2,3 for progressively finer resolution, and L is the larger diameter case. NN means non-Newtonian, thus variable Reynolds numbers. 62
5. Table 5 Analysis of droplet production from primary ligaments and primary shed droplets in W1 and R1 simulations. $\langle d_{primary} \rangle$ is the mean diameter as a function of original drop diameter (D_0). 90
6. Table 6 Non-Dimensional simulation parameters. 112
7. Table 7 Droplet statistics and spray cone angle approximations for select simulations. 119
8. Table 8 Non-dimensional parameters for the simulation. D is initial droplet diameter, dx is grid spacing, ρ_g is gas density, U_g is the bulk gas velocity, μ_l is liquid viscosity, σ is the surface tension coefficient 149

LIST OF FIGURES

Figure Page

1. Figure 1 An example of two droplets fluxing into the same cell. Shaded regions are semi-lagrangian flux regions [138], which transport liquid and $\mathcal{L}$'s into the center cell due to the velocity shown with vectors. The $\mathcal{L}$'s of the two droplets are stored in the list $\mathcal{L}_{\text{Flux}}$. 28

2. Figure 2 An example of two droplets undergoing breakup and coalescence. Two droplets start on the left, and the top droplet breaks into two droplets. The two droplets are identified and the split event is processed. The bottom two droplets coalesce, the merge is identified and the merge event is processed. Note that the $\mathcal{S}$ can vary between timesteps, but the $\mathcal{L}$ is persistent until a split or merge causes the value to change..... 30

3. Figure 3 The diagram on the left shows the issue with fictitious merging and splitting events in the original algorithm [159]. The diagram to the right shows how the split and merge identifier works in the present work. Notice that at $t = 2$ in the original algorithm, the two structures occupy neighboring cells and fictitiously merge back together. The updated code avoids the fictitious merge and the structures do not merge back together. This is because the criterion for coalescence was changed to require the structures to exist within the same cell, not adjacent cells and the breakup identification uses $\mathcal{L}$ in the logical statement for identifying structures. 31

4. Figure 4 Example of data points in Neo4j graphical database. Colors represent the number of breakups a structure has undergone (blue = primary, yellow = secondary, green = tertiary, red = quaternary). Lines connect related droplets, i.e., droplets which split from each other. The liquid core is present at the center of the image. Each node contains relevant statistics from the breakup event..... 35

5. Figure 5 Rendering of the liquid jet run in simulation. $\mathcal{L}$ values at final timestep represented by colors with $\mathcal{L}$ ranging from 1 to $76,608$ for the number of breakups that occurred. Rendered using VisIt [22]..... 39

LIST OF FIGURES – CONTINUED

Figure Page

6. Figure 6 Plot displaying the distribution of breakup stages throughout the entire simulation (blue, circle nodes), e.g. about $7,600$ droplets underwent a 5^{th} breakup throughout this system. The yellow, triangle nodes represent the distribution of breakup stages at the final simulation time. 41

7. Figure 7 A probability density function of the local Weber Number associated with every breakup in the simulation. 43

8. Figure 8 A probability density function of droplet diameter as a function of the number of breakups 45

9. Figure 9 A probability density function displaying the change in diameter between breakup events. 45

10. Figure 10 The cumulative number of primary, secondary, and total droplets. Note: secondary droplets in this image represent the classical definition of secondary droplets (i.e. all breakup following primary). 46

11. Figure 11 Neo4j graph database from an diesel jet atomization simulation [24]. Dots represent droplets, with relationships between them representing split events connecting child and parent droplets. The color represents how many breakup events led to that droplet (i.e. red = four breakup events, green = three, yellow = two, blue = one). 59

12. Figure 12 Simulation setup images. 63

13. Figure 13 Mass loss of initial droplet with images at the same non-dimensional times T. Experimental parameters: $We = 780$, $Oh = 2.4 \times 10^{-3}$, $Re_g = 2.2 \times 10^4$. Experimental images reproduced with permission. 65

14. Figure 14 Early stages of of droplet aerobreakup, largely dominated by stripping (SIE) regime, from (a) Water (W1 simulation), and (b) Shear-thinning (ST1 simulation). 68

15. Figure 15 Bag breakup observed in the W1 Simulation. 69

16. Figure 16 Instabilities early in water droplet development at three different resolutions at $T = 0.11$..... 70

LIST OF FIGURES – CONTINUED

Figure Page

17. Figure 17 Rendering of a water drop (W3) simulation showing initial wave development, transverse modulation instabilities, and ligament stretching from the modulations. 71

18. Figure 18 Rendering from water drop (W3) showing the early stages of instability formation and primary breakup ($T = 0.11$). 72

19. Figure 19 Tree summarizing observed primary droplet production pathways in the stripping regime. 73

20. Figure 20 Asymmetric bridge breakup, creating whipping ligament in R2 simulation. 74

21. Figure 21 Bridge breakup via double pinch-off points. 74

22. Figure 22 Instances of bridge development, stretching, and ultimately detachment from water drop (W2 simulation). 74

23. Figure 23 Examples of wave evolution with time for a water drop (W1 Simulation). 75

24. Figure 24 Measurements of drop diameter as it flattens and RT Wavelength from water drop (W1 simulation). 77

25. Figure 25 Late stage breakup evolution of 25a W1 and 25b R1 79

26. Figure 26 Dense cloud of droplets (orange) breaking off of the initial liquid drop (blue) from the WL simulation. 80

27. Figure 27 A ligament from the (orange) undergoes Plateau-Rayleigh breakup in the WL simulation. 80

28. Figure 28 Breakup stage of final droplets in the W1 simulation. 82

29. Figure 29 Average primary droplet size as a function of T. 83

30. Figure 30 Aspect Ratio (AR) and renderings of various droplets obtained from the water drop simulation. 85

31. Figure 31 Statistics describing the effect of aspect ratio on secondary breakup mechanisms of W1 and ST1 simulations. 86

32. Figure 32 Primary and final droplet diameter distributions from the water drop (W1) simulation. 87

LIST OF FIGURES – CONTINUED

Figure Page

33. Figure 33 Probability density function of final water droplet diameters (d_{eq}) as a function of their original primary droplet parent diameter (d_0). Gamma functions fit with parameters $\alpha = [15.86, 10.27, 6.62]$ and $\beta = [0.05, 0.06, 0.08]$ for resolutions W1, W2, and W3, respectively. 88

34. Figure 34 Effective viscosity of breakup events from the shear-thinning simulations. Water viscosity plotted for reference. 89

35. Figure 35 Bulk atomization rate for water and shear-thinning drops. 90

36. Figure 36 Primary drop size plots, showing overall distribution, and time evolution. 94

37. Figure 37 Simulation domain (left); Navier-Stokes mesh in blue and e-Mesh in black (right) showing relative scale of each domain. 112

38. Figure 38 The Lo_00 (uncharged) and Lo_05 (low electric charge) simulations, showing minimal liquid breakup. 113

39. Figure 39 Renderings of experiment reported by Yule and Shrimpton [220] (reproduced with permission) and Lo_10 simulation (a, images not the same scale), Hi_00 (b), and Hi_10 (c) simulations showing contiguous liquid core (blue) and dislodged structures (orange)); Bloom plots visualizing breakup events in simulations using ASSET (middle row, (d)-(f)); and axial view of simulations (bottom row, (g)-(i)). Full simulation parameters detailed in Table 6. 114

40. Figure 40 Charged simulations statistics for $N_{\text{ei}} = 0.1$. Lo_10 (left column) and Hi_10 (right column). 118

41. Figure 41 A magnified segment of the Lo_10 simulation, displaying ligaments stretching off of the liquid core (blue) and bulging due to Plateau-Rayleigh instability, ultimately forming large droplets and small satellite droplets (both orange). The VOF resolution for simulations is 60 $D_{\text{jet}}/\Delta x$. 119

42. Figure 42 Axial planes x-y colored by charge density. 122

LIST OF FIGURES – CONTINUED

Figure Page

43. Figure 43 Charge density distributions for droplets in the Lo_10 (a) and Hi_10 simulations (b); scatter plot of droplet charge distribution in Lo_10 (orange) and Hi_10 (blue) simulations (c). Red dashed line shows initial value for L_Q of the injected liquid volume. 122

44. Figure 44 Neo4j graph database representing data extracted from a numerical simulation of a diesel-type jet from [24]. Dots represent droplets with lines connecting related droplets (i.e parent and child droplets). Colors represent the number of breakups taken to get to a given droplet. 140

45. Figure 45 Comparison of the volume conserving PLIC interface and the topologically sound, watertight marching cubes interface. 141

46. Figure 46 Height filtration in direction d. Simplices become bold when they are below the plane and contribute to the Euler characteristic for that height t. 144

47. Figure 47 Simulation domain setup. 149

48. Figure 48 Droplet 1 and the four closest matches based on DIWECT 151

49. Figure 49 Droplet 2 and the four closest matches based on DIWECT 152

50. Figure 50 Droplet 3 and the four closest matches based on DIWECT 153

51. Figure 51 Droplet 4 and the four closest matches based on DIWECT 154

52. Figure 52 Total displayed variance of data with varying PCA dimensions. 155

53. Figure 53 Visualization of the droplets in K-means clusters for $K = 4$. 156

54. Figure 54 Statistics for clusters created by k-means clustering. 157

55. Figure 55 Statistics for clusters created by k-means clustering. 158

56. Figure 56 Renderings of samples of droplets from clusters 1,9, and 14 in $K = 16$ clustering output. 159

LIST OF ALGORITHMS

Algorithm Page

1. Algorithm 1 Methodology Summary 147
2. Algorithm 2 $\mathcal{L}$-transport and Coalescence Pseudocode 194

NOMENCLATURE

a	Maximum distance from center of gravity to liquid interface
b	Minimum distance from center of gravity to liquid interface
α	Liquid volume fraction
ϕ	Phase indicator
$\mathcal{L}$	Liquid identification number
$\mathcal{S}$	Structure identification number
L	Characteristic length
T	Non-dimensional breakup time
$\boldsymbol{n}$	Normal vector
ε	Strain-rate tensor
$\boldsymbol{\omega}$	Vorticity
R_{eq}	Radius of a sphere with equivalent volume to a droplet
σ	Surface tension coefficient
$\mathbf{u}$	Velocity vector
$\mathbf{g}$	Gravitational acceleration
ρ	Density
μ	Dynamic viscosity
ν	Kinematic viscosity
Oh	Ohnesorge number
We	Weber number
Re	Reynolds number
ζ	Uniformity parameter
γ	Aspect ratio
ι	Irregularity parameter
ψ	Symmetric Difference Shape (SDS) Parameter
η	Kolmogorov length scale
η_H	Hinze scale
Δ	Computational cell side length
ϵ	Energy dissipation rate

ABSTRACT

Liquids break apart into droplets in countless natural and industrial processes, yet predicting when and how this breakup occurs remains a fundamental challenge. Understanding these mechanisms can lead to more efficient spray systems and deeper insights into environmental flows. However, studying atomization at a fundamental level requires analyzing immense amounts of simulation data, often reaching tens to hundreds of terabytes. To overcome this barrier, this research develops a method for extracting key information from large-scale atomization simulations, significantly reducing data storage and processing requirements while capturing the physics of liquid breakup.

This method was applied to two practical cases: the breakup of water and shear-thinning liquids in high-speed gas flows, and the behavior of electrically charged liquid jets. In both cases, the tool enabled the identification of previously unknown breakup mechanisms. Additionally, a technique was developed to extract and vectorize liquid shapes and surrounding flow fields, allowing for more advanced data analysis. This approach will enable fast comparisons of liquid shapes and flow fields, providing a foundation for a reduced-order model. Ultimately, this research establishes a new way to access data from high-fidelity simulations with the intent that the data be used to improving predictive models for engineering and scientific applications. By providing access to data from expensive, large-scale simulations and uncovering insights from the underlying physics, these findings contribute to the advancement of multiphase flow research and the design of next-generation spray systems.

1

INTRODUCTION

Motivation

Atomization—the disintegration of coherent liquid into a spray—is a fundamental process in numerous industrial and environmental applications, directly influencing efficiency, performance, and overall effectiveness. Wildfires burn an average of 2.8% of the global land area each year, with some regions, such as Africa and Australia, experiencing over 8% of their total land area burned in a single year [55]. Aerial firefighting is a critical method for combating these fires and mitigating risks to human life and the environment. It relies on large-scale atomization, as airtankers deploy thousands of liters of water or fire-retardant to suppress or contain wildfires [101]. The effectiveness of these drops depends on atomization dynamics, where droplet size and distribution directly impact fire suppression and ground coverage. In internal combustion engines, fuel is injected as a spray to optimize combustion efficiency [15, 52, 109]. Maximizing fuel surface area, ensuring uniform air-fuel mixing, and reducing droplet size minimize unburned hydrocarbons, leading to improved efficiency and lower emissions. Similarly, breaking waves on the ocean surface drive air-sea exchange processes [217], producing sea spray aerosols that act as cloud condensation nuclei and influence tropical storm intensity and global climate dynamics [8, 14, 86]. Other critical applications, such as inhalers [16] and irrigation systems [105, 129], also depend on precise droplet formation to optimize drug delivery and water distribution. Given its broad impact, understanding atomization across these diverse systems is essential for enhancing performance, efficiency, and environmental outcomes.

Despite centuries of experimental research and the recent advancement in numerical simulations, fundamental questions about atomization remain unanswered–*when and how do*

liquids break up?. These are deceptively simple questions, yet resolving them would reveal the underlying physics of atomization. However, the complexity of the process has made it difficult to fully characterize, hindering the development of physics-based models essential for predicting atomization behavior in both engineering and environmental systems. This dissertation seeks to bridge that gap by implementing a novel methodology into high-fidelity numerical simulations to extract detailed insights into the breakup process, providing a necessary foundation for future model development. By identifying key mechanisms and conditions that govern droplet formation, this work aims to advance our fundamental understanding of gas-liquid interactions. These insights will ultimately lead to faster and more accurate engineering of spray-based technologies—from fire suppression to fuel injection—while also improving our ability to describe and predict natural processes, which influence global climate systems.

The following chapters contain self-contained manuscripts, and while each provide a more focused background relevant to its specific study, the following sections will more broadly describe the current state of atomization research. This background will outline the current capabilities and limitations of both experimental and numerical approaches in studying atomization, emphasizing what can be extracted from each. Additionally, it will discuss the present state of atomization modeling.

Background - Experimental Work

Because of its wide applicability, there is a considerable body of atomization research that spans multiple centuries [152, 162]. And a key concept that unifies these studies is the role of dimensionless parameters—especially the Reynolds (Re) and Weber (We) numbers—in describing atomization behavior. The Reynolds number compares inertial to viscous forces

as,

$$Re = \frac{\rho U L}{\mu}, \tag{1}$$

where ρ is density, U is the velocity in the fluid, L is a characteristic length (often jet or droplet diameter), and μ is the dynamic viscosity. Depending on the regime of interest, Re may be defined in the gas or the liquid phase to describe turbulence within that phase. The Weber number [214] compares disruptive inertial forces to the restoring effects of surface tension as,

$$We = \frac{\rho_g U_{slip}^2 L}{\sigma}. \tag{2}$$

The Weber number is generally used to describe the inertial forces from aerodynamic disturbances, thus density is the gas density ρ_g and the velocity is defined as the difference between the gas and liquid velocities (slip). In this context, $\rho_g U_{slip}^2$ represents the inertial (aerodynamic) pressure, while σ/L represents the Laplace pressure from surface tension (σ). The balance of these forces determines whether—and how—liquids break up in an atomizing flow.

To explore how Re and We influence atomization, several canonical flow configurations have been extensively studied:

- Round liquid jet in quiescent gas: A single cylindrical jet issuing into still gas provides a straightforward system for investigating the onset and growth of interfacial instabilities. Early work by Rayleigh [152] focused on this configuration, and fuel injection in internal combustion engines is a prominent practical application.

- Air-blast atomizers: Here, a round liquid jet is sheared by a high-velocity annular gas flow. By varying the gas–liquid slip velocity, researchers can easily adjust both Re and We. A well-known study in this context is by Marmottant and Villermaux [119]. Air-blast atomizers appear in many large-scale applications, including rocket engines.

- Planar liquid sheets: Instead of a cylindrical jet, a thin planar sheet of liquid is ejected into a gas environment (which may be quiescent or air-blasted). Though breakup mechanisms in sheets and round jets share common features, the planar geometry simplifies visualization and allows clearer observation of ligament formation and droplet breakup [117].

- Droplet aerobreakup: This scenario directly matches the original definition of the Weber number, which assumes a spherical droplet. Studies on isolated droplets subject to gas flow offer insights into secondary atomization processes relevant to both engineering systems and natural phenomena such as raindrops. Theofanous has a comprehensive series of papers on this subject [191–193, 195].

These cases represent the most commonly studied atomizing systems. They allow for controlled systems with which to study the effect of Reynolds and Weber numbers. Throughout the rest of the present work, several terms will be used to describe the process of atomization. These terms are derived from the body of research describing the systems outlined above.

- Core: The initial coherent liquid structure (e.g., the jet or large parent droplet).

- Primary: The first generation of liquid structures that detach from the core.

- Secondary: Any subsequent generation of structures that form from previously detached ligaments or droplets.

- Spray cloud: A dense collection of droplets forming downstream in high-We flows.

Most research to date has been experimental, relying on visual analysis and measurements of physical sprays. Studying atomization experimentally presents several challenges. One limitation is that cameras, while capable of focusing on multiple planes, primarily

capture information within a specific depth of field. As a result, dynamics occurring outside the focal region may be under-resolved, and the available data is often constrained to a two-dimensional representation. Efforts to address this limitation have included the use of multiple cameras to reconstruct three-dimensional structures [89, 143, 144, 180]. Most practical applications exhibit high Reynolds and Weber numbers, which exacerbate experimental challenges. First, these regimes are highly scale-variable. For example, in fuel-injection, centimeter-scale flow structures determine spray angle and penetration length, yet droplets can form down to the micrometer scale, drastically affecting combustion dynamics [215]. Other applications, such as aerial firefighting, exacerbate this disparity by spanning from meters to micrometers [101]. This variability means that cameras must be able to capture small-scale physics with high-resolutions *and* maintain a large field-of-view, which often proves impractical, causing them to constrict their field of view. In addition, visualizing inside the spray cloud produced in high-We atomization systems is a significant challenge. This cloud forms as large quantities of small droplets are created, making the region effectively opaque. Consequently, it obscures both the primary breakup mechanisms—where the initial liquid core disintegrates—and subsequent secondary breakup processes occurring beyond the visible outer region of the spray.

Due to these difficulties in resolving small-scale features, many experimental studies have instead focused on bulk spray characteristics, such as spray angle, penetration length, breakup length, and instability wavelengths (a thorough summary can be found in *Atomization and Sprays* [100]). However, in order to fully describe the process of atomization, from initial liquid volume to the final spray, it is necessary to characterize and track liquid structures and quantify parameters relevant to breakup. To address this, modern experimental researchers have developed a suite of diagnostic tools to extract quantitative data from atomizing sprays.

A category of these quantitative methods are laser-based diagnostics. The most well-

known among these techniques is particle image velocimetry (PIV), which is traditionally used to track tracer particles and provide a velocity field in (generally single-phase) flows. However, recent advancement has adapted this technology for use to track atomized droplets (e.g. Basu et al. [9]) using successive images of a region of the flow with post-processing to identify the same droplet from one time to the next. Another common method, which attempts to characterize droplets in dense spray regions is Phase Doppler Particle Analysis (PDPA), which utilizes the Doppler shift in scattered light to measure drop sizes and their velocities [13, 59]. Laser diffraction caused by interactions with droplets has also been utilized to study sprays. Instruments such as diffraction granulometers measure the diffraction pattern of a laser beam interacting with a spray, analyzing the intensity of scattered light over a range of angles to determine droplet size distributions (e.g. Smallwood et al. [182]). This analysis is based on Fraunhofer or Mie scattering theories. While effective in many applications, diffraction-based techniques assume spherical droplets, which can introduce errors in certain atomization regimes. Laser-induced fluorescence (LIF) can be used to observe the interior of dense spray regions or the interior of liquids themselves by doping a liquid with fluorescent markers which become excited by certain wavelengths of light, emitting an excited fluorescent proton. LIF has been used both to observe internal droplet breakup mechanisms [195] and, in combination with other techniques, to study dense spray regions [125]. These techniques still suffer to fully characterize dense spray regions in addition to challenges with constrained fields of view, out of focus droplets, and many of these techniques rely on a spherical droplet assumption, which we will see in the results in Chapters 3 and 4 that this can be an unreasonable assumption in certain sprays.

Recent advances in ballistic and x-ray imaging of atomizing flows have attempted to visualize through a dense cloud of droplets. X-ray technology is well suited for imaging the dense regions of a spray, where aspects of primary atomization can be observed [63]. Additionally, X-ray computed tomography (CT) can provide 3D reconstructions of sprays,

offering detailed spatial distributions of liquid mass. However, CT scans require stationary objects, meaning they can only provide time-averaged images of internal spray features [77]. Other X-ray techniques, such as high-speed radiography, allow for single-shot or time-resolved imaging but still suffer from limitations such as high cost, two-dimensional projection effects, and limited field of view for high-resolution techniques. Ballistic imaging is an optical technique that improves visibility into dense sprays by filtering out scattered light [113, 168, 181]. It works by using a fast optical gate to capture only the first photons that pass straight through the spray, while ignoring those that scatter multiple times and blur the image. This allows researchers to see clearer details of the liquid core, including ligament formation and primary breakup. However, ballistic imaging requires high-power lasers, has a limited field of view, and is sensitive to alignment, making it challenging to apply in all situations.

Several experimental studies have had significant contributions to our understanding of atomization physics. One of the most important studies is the seminal work of Marmottant and Villermaux [119], who captured detailed images of atomizing jets at lower Reynolds and Weber numbers compared to those in practical fuel injection systems. At these lower conditions, the breakup process follows a structured progression of hydrodynamic instabilities, where shear at the liquid-gas interface generates Kelvin-Helmholtz waves which ultimately grow and lead to Rayleigh-Taylor instabilities. As these instabilities amplify, ligaments form and fragment under Rayleigh-Plateau instabilities, breaking into droplets. Their study linked observed breakup patterns to linear stability analysis, showing reasonable agreement between measured instability length scales and theoretical predictions.

While these findings provide valuable insight into fundamental atomization mechanisms, real-world systems introduce additional complexity. High-energy atomization regimes, such as those found in many engineering and environmental applications, are dominated by turbulence, making the structured instability growth observed in simplified jets much more

difficult to identify. Linear stability analysis, while useful at lower Reynolds and Weber numbers, fails to fully describe the breakup process under turbulent conditions, where chaotic interactions drive fragmentation.

Further advances in understanding droplet breakup have been made by Theofanous and collaborators [191–193, 195], who investigated droplet aerobreakup—where an isolated droplet deforms and fragments due to aerodynamic forces. Their research mapped several regimes that describe breakup dynamics based on the relative contribution of Kelvin-Helmholtz or Rayleigh-Taylor instibilities. These studies have greatly improved our understanding of the interplay between aerodynamic, viscous, and surface tension forces. A more detailed discussion of these works can be found in Chapter 3.

These studies represent some of the most influential experimental works, displaying the role of experimentation in advancing our understanding of atomization. While these investigations have provided valuable insight into breakup mechanisms and instability growth, they often rely on idealized conditions that do not fully capture the complexity of real-world atomization. Challenges largely associated with optical limitations, and the difficulties of characterizing high-energy flows prevent a comprehensive understanding of the process. Despite continuous advancements in experimental methodology, gaps remain in our ability to resolve the complete dynamics of atomization, especially in dense spray regions.

Background - Numerical Work

To address limitations associated with experimental methods, numerical simulations have emerged as a promising complement to experiments, offering the ability to analyze three-dimensional, multi-scale flow features even within dense spray regions. These simulations, often called direct numerical (DNS) or high-fidelity simulations, rely on the solution of the Navier-Stokes without any turbulence modeling, ideally capturing all relevant spatial and temporal scales. DNS was first applied to gas-liquid flows about twenty years ago [35,

56, 122]. Since then, a majority of research has been focused on improving the numerical methods and pushing the limits of continuously improving computational efficiency.

Despite the tremendous promise of DNS in advancing atomization research, capturing its underlying physics remains exceptionally challenging for several interrelated reasons. First and foremost is the vast range of spatial and temporal scales that must be resolved. Capturing all of the physics necessitates computational grids of billions of cells [110]. Simulations of this magnitude are computationally prohibitive for most practical systems.

A second major difficulty arises from solving the Navier-Stokes equations over the phase interface, where sharp discontinuities exist in material properties and in the pressure from surface tension. Methods such as the ghost-fluid approach [46] handle these discontinuities without smearing values over multiple grid cells, but require precise knowledge of interface location at every time step. A variety of methods have emerged for interface transport—broadly classified as interface capturing [160] or interface tracking [37]. The latter has emerged as the more promising path, continuing to dominate the literature on numerical methods. Two leading examples are the level-set (LS) method [135, 169] and the volume-of-fluid (VOF) method [69].

LS tracks an iso-surface of a smooth function, making the interface location straightforward to compute, but struggles with strict mass conservation. Although conservative LS improvements [137] reduce mass losses, they still introduce approximations. In contrast, VOF tracks the volume fraction of fluid in each cell, providing direct physical significance and mass conservation. VOF methodologies can be categorized as algebraic or geometric. Algebraic VOF methods compute fluxes based on volume fractions without explicitly reconstructing a sharply defined interface, which can introduce numerical diffusion. By contrast, geometric schemes (e.g., piecewise linear interface construction, PLIC [218]) reconstruct the interface within each cell, allowing sharper resolution of the interface at the cost of more complex calculations. The simulations conducted in this work use an un-split, semi-Lagrangian

transport scheme [138, 139], which geometrically transports planar interfaces in a momentum and mass conserving manner (this methodology is also used in the transport of identification scalars, described in Chapter 2). While this method is the state-of-the-art right now, it relies on planar approximations of curved interfaces, which introduces numerical error into the calculations. The field is still rapidly advancing and future extensions of these methods involve using parabolic reconstructions of the interface [41], offering even greater accuracy with smoother curvature approximations.

Most DNS research on atomization has focused on improving accuracy, often requiring increasingly complex calculations, while also developing cost-saving strategies to efficiently solve these equations and address resolution challenges. Given these intertwined challenges, much of the field's progress has appropriately prioritized numerical method development. However, another significant obstacle remains: the difficulty of extracting meaningful physical insights from simulations. Most numerical method improvements require validation through test-case simulations, yet once validated, the resulting datasets are rarely analyzed for atomization physics. Instead, they are often discarded, despite containing valuable quantitative data. This is largely due to their immense size—ranging from tens to hundreds of terabytes—making it nearly impossible to isolate useful information such as droplet formation pathways, primary droplet formation, or model-relevant breakup statistics. The sheer volume of data overwhelms standard post-processing tools, limiting efforts to quantify and interpret the fundamental mechanisms of atomization.behind atomization.

Despite these obstacles, several landmark numerical studies have advanced our understanding of atomization physics. Much of this work has focused on qualitatively describing breakup pathways using three-dimensional rendering of jets and developing regime maps to describe dominant breakup mechanisms based on common non-dimensional parameters like Weber, Reynolds, and Ohnesorge numbers. Some more recent works have been able to extract more quantitative data from DNS, however these are generally large-scale or bulk

statistics rather than the small-scale mechanisms extracted in the present work.

Early works studied fast moving jets injected into quiescent air, analyzing the effect of increasing Weber and Reynolds numbers on primary breakup mechanisms. Desjardins and Pitsch [32] analyzed a planar jet and found that turbulent eddies carrying enough kinetic energy to overcome surface tension forces led to early interface deformation, then stretching of these instabilities led to ligament formation and ultimately breakup due to Plateau-Rayleigh instabilities or aerodynamic forces. Shinjo and Umemura [174] analyzed a round jet in quiescient air, describing ligament formation as the leading tip "rolls up". They found that collisions of droplets from the leading edge with the upstream core led to further instability formation and ultimately further droplet production. They also observed that increasing Weber number led to a decrease in ligament and droplet diameters.

Following these early works, a series of papers from Sirignano, students, and collaborators studied DNS of various atomization systems and made several impactful analyses. Their first two works [81, 82] simulated round jets focused on vortex dynamics and their interaction with the phase interface to create Kelvin-Helmholtz waves and ultimately ligaments. Their post-processing allowed for visualization of near-field vorticies that would prove very difficult, if not impossible to obtain by experimental methods. These works also confirmed that droplet size and spray angle are influenced by viscosity and surface tension. The second two works [221, 222] simulated planar jets. These works focused largely on describing and mapping instability formation and primary atomization pathways and how they differ with different Reynolds and Weber numbers. They also further the discussion on vorticity generation and interaction with the interface. These works largely seek to describe breakup phenomena qualitatively and assign mappings based on large-scale or bulk parameters. The studies described here are the most notable and are representative of the most common type of numerical study on atomization, although the list is not exhaustive.

Descriptions of atomization breakup pathways and regime mappings are vital to our

overall understanding of atomization. However, statistically robust descriptions of small-scale breakup processes and physics are sparse. Some recent works seek to extract more of this small-scale quantitative information in a manner more similar to the present work. For example, Pairreti et al. [140] studied the effects of grid resolution on droplet sizes. They were able to gather the number of droplets and their sizes at various times, showing the evolution of droplet size distributions as a function of time. They did not comment on their methodology to extract droplet sizes, although their calculation of the Sauter Mean Diameter indicates that they were able to calculate droplet surface area and volume. However, their analysis did not capture the exact moments of breakup or droplet ancestry, which limits understanding of the underlying cascade process.

The most similar work to that described in this dissertation is that of Trautner et al. [200], which was published in the same month as the manuscript in Chapter 2. The authors developed a method to identify primary droplets when they detach from the liquid core, providing statistics from that point in the simulation. Their method shows several differences from the methodology of ASSET (Chapter 2). They use the same connected-component labeling (CCL) algorithm from Herrmann [66] to identify liquid core and separated droplets, but instead of assigning identification numbers they define four different subsets of the volume fraction (α). This is done such that the total volume fraction at at the start (s) of a timestep (t) in cell (i, j, k) is

$$\alpha^{t_s}(i,j,k) = \alpha^{t_s}_{core} + \alpha^{t_s}_{drop} + \alpha^{t_s}_{untag} + \alpha^{t_s}_{mix}, \tag{3}$$

where *core* is volume that is part of the core, *drop* is liquid entrained in droplets, *untag* is liquid from spurious liquid structures (smeared interfaces or flotsam), and *mix* is liquid from both *core* and *droplet*, giving it an unknown origin. Tracking these separate volume fractions over the course of a timestep, they again use the CCL algorithm to identify the liquid core and droplets and any volume fraction subsets that do not match the newly defined positions

are relabeled (e.g. α_{core} that is identified as a droplet is relabeled α_{drop}). This methodology provides information about when primary structures form, however, it does not provide identification numbers to allow for tracking of individual structures like ASSET, therefore it is limited to only primary droplet analyses. Additionally, this paper analyzes the local Weber number around droplets, which is another goal of this dissertation, however, their methodology is limited compared to that described in this work. Further discussion of this appears in Chapter 6.

In summary, numerical simulations of atomization have made significant strides in recent decades, driven by the need to overcome experimental challenges. Although current numerical methods can capture large-scale dynamics and provide valuable insights, challenges remain in extracting and analyzing small-scale, statistically relevant quantitative descriptions of breakup from the massive datasets produced by DNS. The methods and results described in the following chapters build on this foundation, offering an innovative approach to analyze high-fidelity atomization simulations.

Background - Reduced-order Models

While DNS offers a promising future for studying atomization, the associated computational costs of running these simulations make them impractical for most applications. This has led many researchers and engineers to utilize reduced-order atomization models. As such, atomization modeling has been a subject of extensive research, leading to the development of numerous models, each built on a set of simplifying assumptions tailored to specific flow systems. Because of this, every model is limited in its applicability and will fail to accurately account for all aspects of the physics.

The vast majority of atomization models predict single-droplet breakup, which is essentially secondary atomization. While some primary atomization models exist [75, 165], they are relatively sparse compared to drop breakup models and are typically based on

linear stability theory, describing the growth of unstable waves at the liquid-gas interface. A common approach for modeling an entire atomization system is "blob injection", where large liquid structures, often sized to match the nozzle diameter, are introduced into the computational domain [25]. Secondary atomization models are then applied to predict how these structures fragment into smaller droplets. The data extracted from ASSET are particularly well suited for refining secondary atomization models, as they provide detailed insights into droplet breakup based on shape, size, and local flow conditions. Consequently, the discussion here will focus primarily on secondary atomization models.

Three of the most widely used models have remained foundational in the field for nearly 40 years: the Taylor Analogy Breakup (TAB) model [134], the Pilch-Erdman (PE) model [145], and the Reitz-Diwakar (RD) model [154]. These models represent secondary atomization models and in modern studies are often used in conjunction with other, more advanced techniques like Eulerian-Lagrangian models. The following is a summmary of these three methods:

- **Taylor Analogy Breakup (TAB) Model:** Treats the droplet as a spring-mass-damper system, where aerodynamic drag acts as the driving force, surface tension as the spring (restoring force), and viscosity as the damper. The model solves a second-order equation for droplet deformation and declares breakup once a threshold displacement is reached. The Sauter mean diameter of the resulting droplets is estimated using an energy balance, and mass conservation dictates how many droplets form at that size.

- **Pilch-Erdman (PE) Model:** Relies on a critical Weber number (We_{cr}), above which breakup occurs. The flow regime determines the breakup "mode" (e.g., bag, stripping, catastrophic) via a piecewise function of the Weber number. A breakup time is defined from the mode. Then, the largest stable droplet diameter is estimated using empirical correlations and mass conservation used to compute final droplet sizes and quantities.

- **Reitz-Diwakar (RD) Model:** Similar to TAB, adopts a fixed critical Weber number of 12 to trigger breakup. Separates fragmentation into either bag or stripping modes, each associated with different empirical correlations. A breakup time is determined based on Weber number and stable droplet radius for each mode. A linear rate of droplet size decay is applied until the stable radius is reached, with droplet counts found via mass conservation.

The above models were all published in 1987 and much research has been conducted to improve them since then. Enhancements to the TAB model include E-TAB [186] and MTAB [184]. The RD model has been improved and is often referred to now as the Kelvin-Helmholtz-Rayleigh-Taylor (KH-RT) model [142]. Additional models include the Droplet Deformation Breakup (DDB) Model [76], the Eulerian-Lagrangian spray and atomization (ELSA) model [211], the bag type breakup (BTB) model [209], and the multi-mode breakup (MMB) model [208].

The hybrid Volume-of-Fluid Lagrangian-Eulerian Model (VoFLE) is identified as one of the most promising recent models. It combines an DNS solution to the Navier-Stokes equations for the coherent liquid core and a transition to Lagrangian droplet tracking for separated droplets. Comprehensive detail on this method is given in the introduction of Chapter 5.

The most prominent secondary atomization models depend on empirical data to define their assumptions and tune parameters. For example, the three foundational models [134, 145, 154] each rely on an estimation of the critical Weber number [132] to predict droplet breakup. However, the classical definition of the Weber number assumes a one-dimensional flow field and a perfectly spherical droplet, which oversimplifies the three-dimensional and multi-scale nature of droplet breakup. In reality, droplet deformation, local velocity fluctuations, and surface instabilities all contribute to breakup dynamics in ways that are not fully captured by existing models. Chapter 6 introduces an improved local Weber

number formulation derived from numerical simulations, addressing these limitations.

Beyond breakup onset, another major limitation of current models is their handling of resultant droplet distributions. The currently used approaches might be functional for specific applications, but they fail to capture the true complexity of atomization, where droplet formation depends on local flow conditions, complex liquid shapes, and the interaction of numerous instabilities. Because experimental methods struggle to track these factors in a statistically significant manner, model development has been constrained by limited data availability.

ASSET (Chapter 2) directly addresses these gaps by providing a systematic way to extract detailed, high-fidelity breakup statistics from large-scale numerical simulations. Unlike traditional experimental methods, ASSET allows for the identification and tracking of individual liquid structures throughout the entire atomization cascade, offering a complete picture of the breakup process. This includes precise measurements of droplet shapes, sizes, local velocity fields, and fragmentation events—information that is otherwise unavailable in statistically relevant quantities. With access to these data, more robust reduced-order models can be developed, improving accuracy while maintaining computational efficiency.

Contributions

This dissertation will contribute in a variety of ways to the study of gas-liquid flows, atomization, and spray development, specifically focused numerical simulations of this complex process. These contributions include:

1. Development and Implementation of the Atomization Simulation Statistics Extraction Tool (ASSET). ASSET is integrated into large-scale numerical fluid solvers to extract detailed, localized data from atomization processes during simulation run-time. This methodology enables direct access to relevant physics that would otherwise be obscured

within tens to hundreds of terabytes of simulation data.

2. High-fidelity numerical simulations of droplet aerobreakup are conducted to elucidate the physics of atomization in water and shear-thinning liquids at high density ratios and Weber numbers. Descriptions of instability growth, primary, and secondary atomization mechanisms are described in detail and novel data about secondary atomization cascades are extracted with ASSET.

3. High-fidelity numerical simulations of electrically charged liquid jets are run. The results highlight the interplay of increasing Reynolds and Weber numbers and electric charge and their effect on atomization.

4. A novel method is introduced for extracting droplet geometries and flow fields from atomization simulations and converting them into vectorized representations. The outputs are processed using geometric and topological data analysis. The result is a vectorized database able to faithfully capture physical data, while reducing data storage overhead and providing a format usable by advanced data-analysis and machine learning applications.

5. A method to extract a local Weber number from numerical simulations is developed. This work improves on relatively naive approaches currently used, taking into account local flow variations and avoids boundary layer interference to capture an unperturbed gas velocity.

Organization of this document

Chapters 2, 3, 4, and 5 are copies of published and submitted journal manuscripts. Each chapter is self-contained and can be read separately. These chapters describe the development of ASSET, the droplet aerobreakup simulations, electrohydrodynamics simulations, and the

geometric and topological data analysis methodology, respectively. Chapter 6 outlines the new method to extract the local Weber number. This method builds on the work described in Chapter 2 and a journal article describing this work is currently under preparation.

19

EFFICIENT EXTRACTION OF ATOMIZATION PROCESSES FROM HIGH-FIDELITY SIMULATIONS

Contribution of Authors and Co-Authors

Manuscript in following chapter

Author: Brendan Christensen

Contributions: Methodology, Software, Data Curation, Writing - Original Draft, Visualization

Author: Mark Owkes

Contributions: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding Acquisition

Manuscript Information

Brendan Christensen and Mark Owkes

Computers and Fluids

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
_X_ Published in a peer-reviewed journal

Elsevier Ltd.
30 March 2023
Computers & Fluids, 254, 105808
https://doi.org/10.1016/j.compfluid.2023.105808

Abstract

Understanding the process of primary and secondary atomization in liquid jets is crucial in describing spray distribution and droplet geometry for industrial applications and is essential in the development of physics-based low-fidelity atomization models that can quickly predict these sprays. Significant advances in numerical modeling and computational resources allow research groups to conduct detailed numerical simulations and accurately predict the physics of atomization. These simulations can produce hundreds of terabytes of data. The substantial size of these data sets limits researchers' ability to analyze them. Consequently, the process of a coherent liquid core breaking into droplets has not been analyzed in simulation results even though a complete description of the jet dynamics exists. The present work applies a droplet physics extraction technique to high-fidelity simulations to track breakup events as they occur and extract data associated with the local flow. The data on the atomization process are stored in a Neo4j graphical database providing an easily accessible format. Results provide a robust, quantitative description of the process of atomization and the details on the local flow field will be useful in the development of low-fidelity atomization models.

Introduction

Atomizing sprays have a wide range of applications in industrial and environmental fields (e.g. fuel injection [223], agricultural sprays [99], pharmaceutical sprays [118], and air-sea interaction [50]). Consequently, the body of research on this topic is vast and spans multiple centuries [152, 162]. Researchers have made significant progress in understanding many aspects of the atomization process, however, the mechanisms that drive instabilities in coherent liquid structures and ultimately cause breakup are still widely unresolved. This is, in part, due to limitations in experimental data collection from atomizing flows. Visual analysis

of atomizing jets is difficult because of the scale and speed at which these systems develop. Furthermore, the development of droplets creates an opaque cloud, which blocks the view of the liquid core and severely limits the ability to study primary instability and breakup mechanics. Recent advances in ballistic [113, 168, 181] and x-ray [63, 108, 210] imaging of atomizing flows have attempted to remedy visualization issues. These methods provide snapshots of the atomization process and lack important temporal information. Because of challenges with current experimental methodology, numerical simulations, and particularly high-fidelity simulations that resolve the relevant time and length scales, have been developed and provide an alternative method to study the physics of atomization.

Advances in numerical methods and computational efficiency within the past decade have greatly expanded the capabilities of numerical simulations. Researchers are now able to simulate multiphase flows with exceptional accuracy. Numerically simulating multiple phases is not a trivial process, however. These systems involve a wide range of topological scales in fluid and turbulent structures in addition to jumps in physical qualities at the phase interfaces. Significant effort has been and is still invested in improving numerical simulation methodology to more accurately and efficiently simulate atomizing systems. Accordingly, most high-fidelity simulation studies on atomization to date are focused on the development of numerical methods and testing against experimental results rather than using the simulations to produce novel atomization data. But, with the pending maturity of the field, several research groups have been able to gather some insightful statistics from high-fidelity atomization simulations.

Following Gorokhovski and Herrmann's review of atomization simulations [56], the field developed rapidly. Among the first high-fidelity simulations of atomizing liquid were Shinjo and Umemura [174] and Desjardins and Pitsch [32] who simulated round and planar liquid jets, respectively. These works provided a foundation for future simulation studies by identifying numerous breakup mechanisms and providing qualitative descriptions of

previously unclassified processes. Following these pioneering studies, various other works were published that utilized high-fidelity simulations to expand upon existing descriptions and identify novel breakup mechanisms. Several of these groups incorporated vorticity measurements into their post-processing to strengthen descriptions of breakup by integrating this additional dimension to their analysis [81, 82, 106, 175, 213, 221, 222]. Additionally, many researchers focus their efforts on comparing linear-stability theory [11, 154] and high-fidelity simulations [2, 31, 51, 95, 175]. These studies process renderings of the atomizing jet to analyze and track wavelengths as they become more unstable, ultimately leading to breakup. Several recent studies have obtained turbulence statistics throughout the simulation using constrained spatial and/or temporal domains to limit data-processing requirements [60, 161, 198].

The numerical studies mentioned above have greatly improved our understanding of the atomization process. They developed robust qualitative descriptions of breakup events and utilized bulk quantities to elucidate some of the underlying physics of liquid breakup. However, despite the recent advances in atomization simulations, there remain significant gaps in our understanding of the process. This is primarily due to a lack of available methodology for obtaining many relevant statistics from high-fidelity simulations. In particular, current research methods have not been able to extract local, temporally continuous, quantitative data on liquid breakup. The data include information on droplet shape and size characteristics, local flow field data in both the liquid and gas phases, and information on how these values evolve temporally and spatially throughout a spray system. An appreciable challenge in obtaining this data is derived from the massive size of resultant data sets; these can be hundreds of terabytes or even larger. Parsing such data sets for relevant information is not practical and is often impossible.

While numerical simulations are becoming more efficient, it remains computationally expensive and requires the use of high-performance computing. Reduced-order atomization

models aim to provide a viable alternative to high-fidelity simulations at a fraction of the computational cost. Some prominent examples of atomization models include the Taylor analogy breakup (TAB) model [134], the Pilch-Erdman (PE) model [145], the Kelvin-Helmholtz-Rayleigh-Taylor (KH-RT) model [142, 154], the Eulerian-Lagrangian spray and atomization (ELSA) model [211], the bag type breakup (BTB) model [209], the multi-mode breakup (MMB) model [208], and the modified Taylor analogy breakup (MTAB) model [184]. These atomization models try to predict how and when a liquid structure breaks apart without fully resolving the physics. As mentioned above, however, there is little information on the mechanisms that control breakup events and even less information on the process of atomization. Model developers would greatly benefit from local data sampled from breakup events throughout the simulation.

The present work aims to address the above obstacles through the improvement of a framework, first introduced in Rubel and Owkes [159], that extracts relevant data from breakup and coalescence events in atomization simulations. The tool operates concurrently with the simulation, extracting information from every breakup or coalescence event. This eliminates the need for extensive post-processing of massive simulation datasets. Improvements to the tool include a major revision of the breakup and coalescence identification algorithm, implementation of a greater temporal sampling range, and the addition of several new data types to be extracted. The values extracted are not entirely novel, however, the ability to extract these values from around every breakup event in an atomizing system is novel. Previous studies rely on time-consuming post-processing methods and/or must constrain the data they extract either temporally or spatially, which limits the statistical relevance of the dataset. The present work provides a method to gather significant quantities of these data without constraining the location or time within the simulation. Through the use of this tool, many researchers will gain access to extensive and easily accessible data sets from atomization simulations.

Methods

The droplet physics extraction tool, originally proposed in Rubel and Owkes [159], is improved and applied in this work. The tool is implemented into high-fidelity atomization simulations to gather data from discrete events within these simulations coinciding with liquid breakup and coalescence. This section will outline the methodology for this process in detail.

Computational Platform

The proposed extraction tool can be applied to any high-fidelity Navier-Stokes solver, given the solver incorporates two identification numbers used to determine breakup and coalescence events, which will be defined in the following sections. This work employs the NGA computational platform [33, 35, 138, 139]. NGA solves the two-phase formulations of the Navier-Stokes mass and momentum conservation equations, defined as

$$\frac{\partial \rho_\phi}{\partial t} + \nabla \cdot (\rho_\phi \boldsymbol{u}_\phi) = 0 \tag{4}$$

and

$$\frac{\partial \rho_\phi \boldsymbol{u}_\phi}{\partial t} + \nabla \cdot (\rho_\phi \boldsymbol{u}_\phi \otimes \boldsymbol{u}_\phi) = -\nabla p_\phi + \nabla \cdot (\mu_\phi [\nabla \boldsymbol{u}_\phi + \nabla \boldsymbol{u}_\phi^T]) + \rho_\phi \boldsymbol{g} \tag{5}$$

respectively. In these equations, ρ_ϕ is density, $\boldsymbol{u}_\phi$ is the velocity vector, p_ϕ is the pressure, μ_ϕ is the dynamic viscosity, t is time, $\boldsymbol{g}$ is the gravitational acceleration vector and ϕ is the phase indicator with either $\phi = g$ (gas) or $\phi = l$ (liquid). These equations are solved on a staggered Cartesian mesh, in which scalar values, such as pressure, are stored at cell centers and velocities at the cell faces. Time is discretized with an iterative Crank-Nicolson formulation and a semi-implicit correction is applied on each sub-iteration [23]. Away from the phase interface, mass, momentum, and any other scalars are transported with conservative, high-

order finite difference operators [33]. Near the interface, a geometric volume-of-fluid (VOF) method is utilized to ensure conservative transport of mass and momentum [138, 139]. The VOF scheme utilizes semi-Lagrangian transport to predict the volume of liquid that fluxes through each cell face during a timestep. This is accomplished through the calculation of geometric flux regions from the velocity field. In Fig. 1, the geometric flux regions are displayed as the shaded regions. These are signed volume regions that calculate fluxes into or out of cells in a given timestep. They are constructed from velocities at the cell vertices. This formulation of geometric transport is un-split, and the use of cell vertex velocities prevents flux region overlap and an accumulation of errors. The interface is reconstructed using a piece-wise linear interface reconstruction (PLIC) [218] with interface normal vectors computed using the efficient least-squares VOF interface reconstruction algorithm (ELVIRA) [146]. The pressure Poisson equation is solved utilizing the ghost fluid method [46] and a black box solver [30]. Interface curvature is computed with the adjustable curvature evaluation scale (ACES) method [136]. NGA is fully parallelized with a message passing interface (MPI) protocol and scales well to tens of thousands of cores [32].

Breakup and Coalescence Event Identification

The coalescence and breakup identification processes operate through the implementation of two identification numbers, which are integers unique to every independent liquid structure within a simulation. These values are referred to as the structure identification number $\mathcal{S}$ and the liquid identification number $\mathcal{L}$. Every liquid structure in the simulation domain is tagged with each of these values. The process in which $\mathcal{S}$ and $\mathcal{L}$ are assigned, transported, and reassigned is the basis for identifying liquid splitting and merging events within a simulation.

Identification Numbers Liquid identification numbers ($\mathcal{L}$) are transported with the liquid. and provide a time history of where the liquid moves. The un-split, semi-Lagrangian

VOF scheme, described in Section 2, predicts the volume of liquid that fluxes through a cell face. $\mathcal{L}$'s are assigned to every region where the volume fraction of liquid is greater than zero. So, when liquid is fluxed through a cell face, the $\mathcal{L}$ that inhabits that liquid will also be fluxed. $\mathcal{L}$ fluxes are assigned a sign (+) or (-) to track the direction they move through the face using the same sign convention as the VOF transport. $\mathcal{L}$ transport is similar to the transport of liquid, however, when multiple $\mathcal{L}$'s move into the same cell, they must be treated separately instead of being summed. A list is created for each cell, which tracks all the unique values of $\mathcal{L}$ that inhabit or move into the cell within that timestep. Further discussion of how the list is used to identify coalescence events is provided in Section 2. Because these values are intended to provide a history of liquid movement, $\mathcal{L}$'s are not reassigned at every timestep. They are persistent through time and are only reassigned after a split event creates a new liquid structure or a coalescence event destroys a liquid structure.

Structure identification numbers ($\mathcal{S}$) are assigned at every timestep following liquid transport. The values of $\mathcal{S}$ are not consequential, provided that every structure is assigned a unique value. Independent structures are identified and tagged using a band-growth algorithm first described by Herrmann [65]. In the present application, the algorithm operates as follows. To begin, $\mathcal{S}$ is assigned on the entire domain to be zero. Then, the domain is looped through to find the first untagged liquid node, meaning the liquid volume fraction is greater than zero and $\mathcal{S} = 0$. When an untagged liquid node is found, all adjacent cells are compared, using a key logical statement to create continuous liquid structures. This statement has been modified in the present work and drastically improves the accuracy of breakup and coalescence identification. The logical statement requires that: 1) there is liquid in that cell, 2) it is untagged ($\mathcal{S} = 0$), and 3) the $\mathcal{L}$ of the adjacent cell must equal that of the current cell. The third requirement is the modification and ensures that structures, which were not together at the previous timestep (meaning different $\mathcal{L}$ values), do not coalesce in adjacent cells. Additionally, the algorithm does not prematurely identify

breakup. In order for structures to break up, there must be a full cell between them. This process prevents structures from breaking up and re-merging immediately following breakup as structures advance and naturally inhabit adjacent cells. Herrmann's work also described the parallelization of this process, which is utilized in the extraction tool. This requires that special consideration be given to communicating $\mathcal{S}$ between processors. Many other methods exist to assign the structure identification number. This type of tagging problem is known as Connected Component Labeling, e.g. [64, 72].

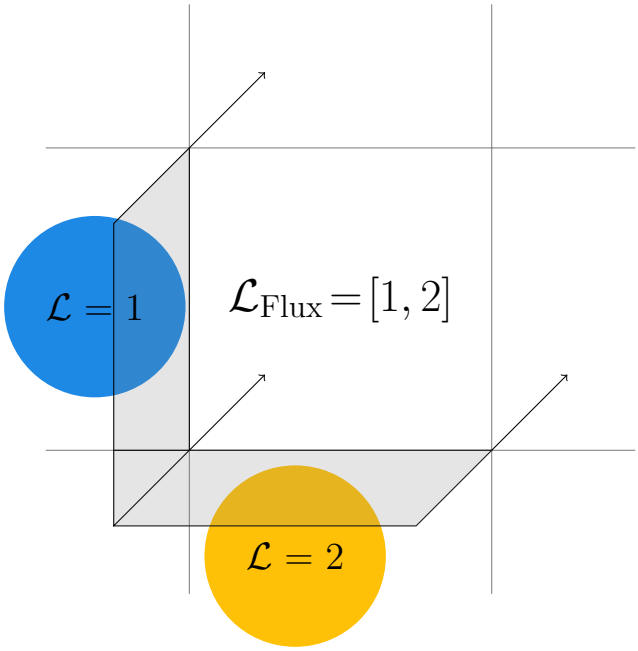


Figure 1: An example of two droplets fluxing into the same cell. Shaded regions are semi-lagrangian flux regions [138], which transport liquid and $\mathcal{L}$'s into the center cell due to the velocity shown with vectors. The $\mathcal{L}$'s of the two droplets are stored in the list $\mathcal{L}_{\text{Flux}}$.

Coalescence Identification Coalescence of liquid structures is identified by locating liquid structures with more than one unique associated $\mathcal{L}$. The $\mathcal{L}_{\text{Flux}}$ lists, described in Section 2, store all the $\mathcal{L}$ that flux into a given cell during a timestep. These lists are looped through and any containing multiple unique $\mathcal{L}$'s indicate coalescence. This represents a significant change in the algorithm. Previously, structures that were in adjacent cells

could be identified as coalesced, which led to repeated non-physical breakup and coalescence identifications (see Section 2). Fig. 1 displays a visual representation of this process. See A for the $\mathcal{L}$ transport and coalescence identification algorithm. When coalescence is identified, data are extracted from the event, and the new merged structure takes the smaller $\mathcal{L}$ of the two parent structures.

Breakup Identification Liquid structure splitting or breakup is identified when two different structures (two unique $\mathcal{S}$'s) have the same $\mathcal{L}$. Two structures become independent of each other when they are not in adjacent cells i.e., there is a full cell separating them. This is when the $\mathcal{S}$ assignment portion of the code assigns two structures with unique $\mathcal{S}$'s. If the two structures have the same $\mathcal{L}$, this indicates that a breakup occurred. Since $\mathcal{L}$'s are persistent through time, the same value in neighboring cells indicates that these two droplets were part of the same structure at the previous timestep. When a breakup is identified, data are extracted from the event, and a new $\mathcal{L}$ is assigned to the smaller structure(s) created by the breakup event. See $t = 0$ and $t = 1$ in Fig 2 for a visual representation of this process.

Addressing Fictitious Events An issue from the previous work, described in Rubel and Owkes [159] was the occurrence of numerous fictitious merge and split events. These occurred when the tool identified a structure that broke into two structures and then the two structures coalesced. The algorithm identified this process and repeated many times, when in fact the liquid structures never re-merged following breakup. The authors of the original algorithm opted to run the breakup identification portion of the code every 10 timesteps to avoid an accumulation of non-physical results. This severely limited the potential usefulness of the tool.

The error stemmed from the original $\mathcal{S}$ assignment algorithm. Droplets in adjacent cells were automatically treated as the same structure, which resulted in recently split liquid structures coalescing into their parent droplet when they were in neighboring cells. The present work introduces two major updates to address this issue. First, the algorithm

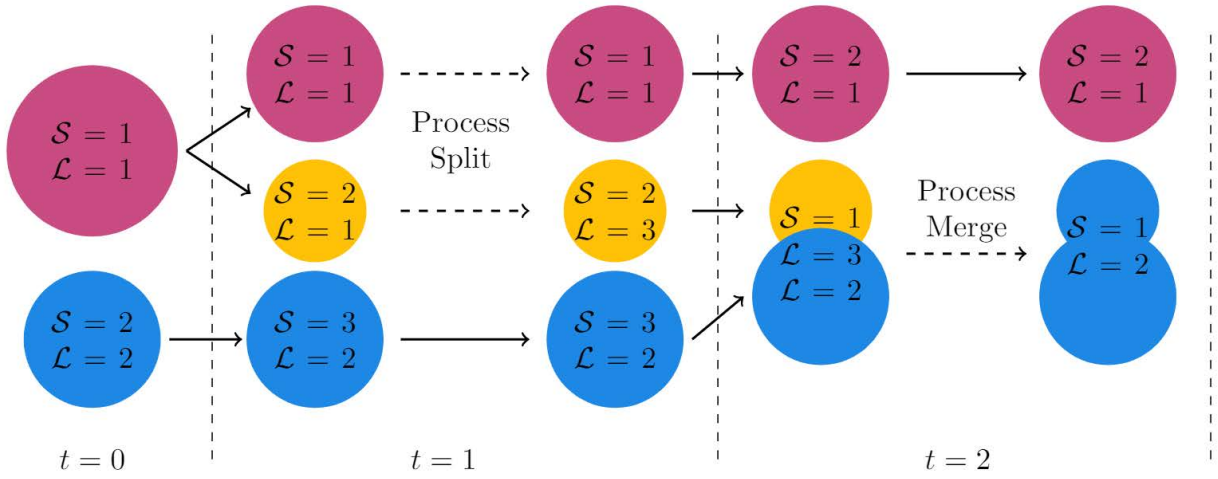


Figure 2: An example of two droplets undergoing breakup and coalescence. Two droplets start on the left, and the top droplet breaks into two droplets. The two droplets are identified and the split event is processed. The bottom two droplets coalesce, the merge is identified and the merge event is processed. Note that the $\mathcal{S}$ can vary between timesteps, but the $\mathcal{L}$ is persistent until a split or merge causes the value to change.

which identifies continuous liquid structures was improved by introducing a logical statement. The statement requires that for neighboring cells to be labeled as existing within the same structure, the cells must have the same $\mathcal{L}$. Second, the algorithm which identifies coalescence was rewritten to create $\mathcal{L}_{Flux}$ lists for each cell. This was implemented to only identify coalescence when multiple $\mathcal{L}$'s occupy the same cell, indicating two liquids, which were separate at the previous timestep have merged together. Together these changes require droplets to be separated by at least one grid cell for breakup to be identified, contrarily coalescence requires droplets to occupy the same grid cell. This difference in length scales for breakup and coalescence ensures that droplets do not breakup and then fictitiously coalesce. For more information on these updates, see Section 2. The updated algorithms drastically boost the accuracy of the tool and allow it to be run at every timestep without erroneous event identifications. This provides a high-resolution sampling of data as the jet breaks up further and breakup and coalescence happen more frequently.

Fig. 3 illustrates this updated $\mathcal{S}$ assignment process. At $t = 0$ a droplet is breaking up

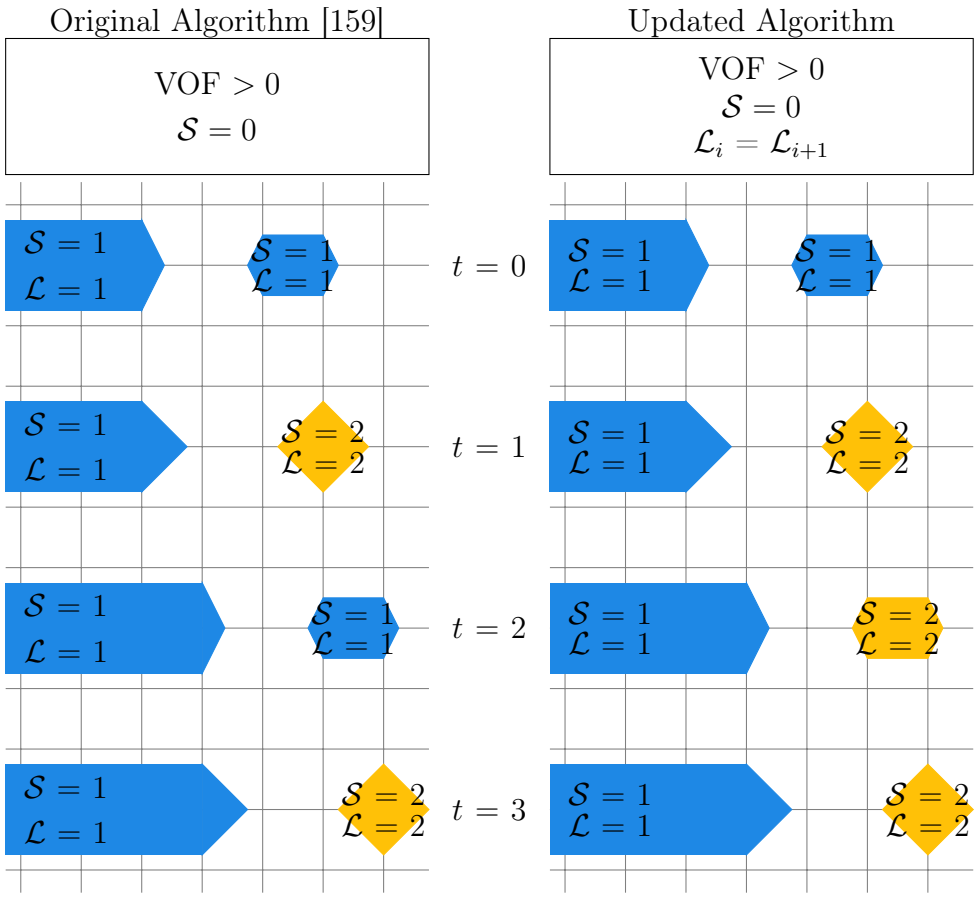


Figure 3: The diagram on the left shows the issue with fictitious merging and splitting events in the original algorithm [159]. The diagram to the right shows how the split and merge identifier works in the present work. Notice that at $t = 2$ in the original algorithm, the two structures occupy neighboring cells and fictitiously merge back together. The updated code avoids the fictitious merge and the structures do not merge back together. This is because the criterion for coalescence was changed to require the structures to exist within the same cell, not adjacent cells and the breakup identification uses $\mathcal{L}$ in the logical statement for identifying structures.

so that there is liquid in adjacent cells. Since a split has not been identified yet, the $\mathcal{L}$ of both cells is the same, resulting in a continuous structure. Then, at $t = 1$ the structure on the right moves, leaving a full cell between it and the other structure. The cells adjacent to the blue structure are void of liquid, so the $\mathcal{S}$ value is only assigned to that structure. Since the droplet on the right becomes independent, it is assigned a new $\mathcal{S}$, a split is processed and it is also assigned a new $\mathcal{L}$. Finally, at $t = 2$, previously, the two droplets would have

merged back together as they occupy neighboring cells. But the structures remain separate with the new algorithm because the $\mathcal{L}$ of the adjacent droplet does not equal the $\mathcal{L}$ of the blue droplet, they remain separate structures and will not coalesce unless they enter the same cell.

Data Extraction

The updated algorithm used to identify breakup and coalescence is run every timestep and thus identifies these events when they occur throughout the entire simulation. When an event is identified, important statistics including the location, volume, local flow field, and shape characteristics are computed for each liquid structure and associated with that event. These statistics are computed for each structure present in the simulation at a timestep and stored in structure arrays. The structure arrays are linked lists of the statistics and identification numbers that are used to tie the statistics to the breakup and coalescence events.

The original work identified breakup events and extracted data every 10 timesteps causing the extracted data to be temporally removed from the actual event. The updated algorithm runs every timestep, thus identifies breakup when it occurs, furthermore, the temporal sampling range was extended by extracting data from the timestep immediately preceding breakup in addition to the timestep following breakup. We accomplish this by saving the structure arrays from the previous timestep, which contain the statistics and $\mathcal{L}$'s of the structures at that time. Then, when breakup occurs, the data extraction algorithm searches the old structure arrays for the structure which has an $\mathcal{L}$ equal to the old $\mathcal{L}$ of the newly split structures. This improvement provides insight into the conditions which lead to breakup and the resultant structures from the breakup.

To highlight some potential uses of the tool, the present work extracted the locations of the events, the gas and liquid velocities, droplet volume, and droplet shape information.

We calculate structure volume simply as

$$\mathcal{V}_{\text{struct}} = \sum_{i=1}^{N_s} \mathcal{V}_{\text{cell},i}\ \alpha_i \tag{6}$$

where N_s is the total number of cells encompassing the structure, $\mathcal{V}_{\text{cell},i}$ is the volume of the i^{th} cell, and α_i is the volume fraction of liquid within the i^{th} cell. For the purposes of this study, liquid and gas velocities are calculated as volume averaged velocities,

$$\boldsymbol{u}_{\text{liquid}} = \frac{\sum_{i=1}^{N_s} \boldsymbol{u}_i \mathcal{V}_{\text{cell},i}\ \alpha_i}{\sum_{i=1}^{N_s} \mathcal{V}_{\text{cell},i}\ \alpha_i} \tag{7}$$

$$\boldsymbol{u}_{\text{gas}} = \frac{\sum_{i=1}^{N_s} \boldsymbol{u}_i \mathcal{V}_{\text{cell},i}\ (1-\alpha_i)}{\sum_{i=1}^{N_s} \mathcal{V}_{\text{cell},i}\ (1-\alpha_i)} \tag{8}$$

respectively. These values are calculated by considering all cells that contain a liquid structure and all cells adjacent to those. Where $\boldsymbol{u}_i$ is the velocity vector in the cell. This method for calculating the velocities is preliminary and does not fully capture the local flow field dynamics. Future work focused on extracting topological data from the flow field will work to better quantify these values.

Graphical Database

The primary goal of the tool is to make data from atomization simulations more accessible. Special consideration was given to ensuring that extracted data was stored in an efficient and easily queriable format. We opted to store the atomization data in a Neo4j graphical database. Graph databases are commonly used in the corporate sector for companies to create connections between users and products through paths, which can reveal patterns in group dynamics and buying trends. The same principles can be applied to atomization simulations through the construction of paths that connect droplets through breakup and coalescence events. Storing atomization data in a graph database format allows

researchers to analyze the evolution of liquid as it moves from the liquid core to small droplets in an atomizing flow. In this work, the Neo4j database is used to store the atomization data. Data from the simulation in this work is initially written to a CSV file. Then, rather than simply storing the data as a disconnected list of events, the CSV file is uploaded to Neo4j. The graph format presents an easy way to connect related events and develop a droplet ancestry. Additionally, storing data in graph databases provide a unique way to visualize atomization data and a novel method for studying the atomization process.

Fig. 4 displays how data are stored in a Neo4j database. This image represents some of the data from the simulation described in Section 3. The nodes are droplets produced from breakup events. The breakup events are represented by the lines connecting the nodes. The node colors represent the number of times they have broken up. The liquid core is in the center of the image and the red droplets represent the fourth breakup event. This image is made up of all the droplets which broke up 4 times during the first 2.3 μs of the simulation, and all the intermediate droplets between the coherent liquid core and these fourth breakup droplets. Within each node is stored the extracted data from the point and time in the simulation in which that droplet broke up and became independent. So, this database is saturated with data, but organized in a logical and accessible format.

The graph database can be efficiently edited, reorganized, and parsed with the Cypher Query Language. It is syntactically similar to SQL but designed to specifically query nodes and relationships and the paths formed with these components. As mentioned above, data is imported to Neo4j via a CSV file, organized in such a way that each row is an independent liquid breakup or coalescence event. These rows are imported into Neo4j as droplet nodes and breakup and coalescence relationships are created. See B for a detailed description of the process to import data into Neo4j. Following data import, Cypher can be used to further organize the data and/or query the data to analyze the atomization process. Below are some examples of the capabilities of Cypher in the present application.

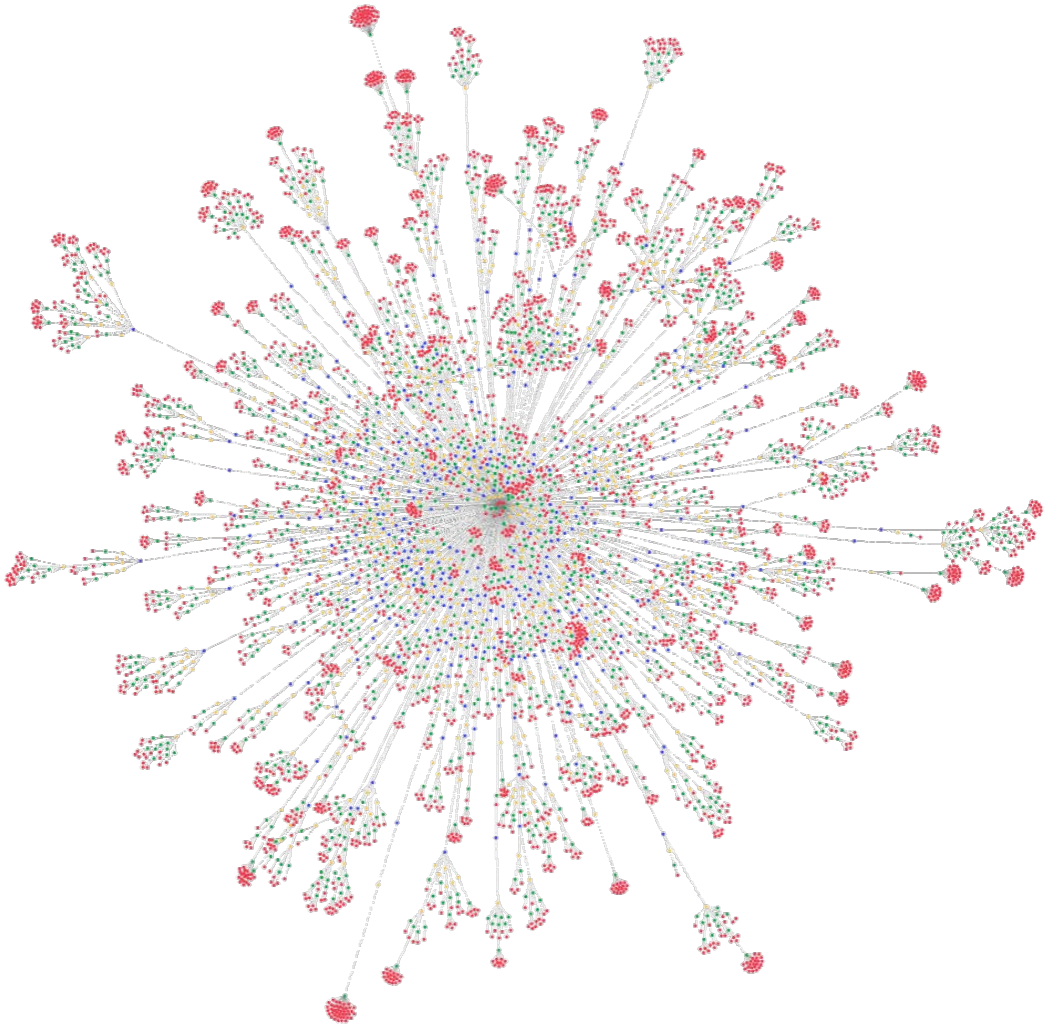


Figure 4: Example of data points in Neo4j graphical database. Colors represent the number of breakups a structure has undergone (blue = primary, yellow = secondary, green = tertiary, red = quaternary). Lines connect related droplets, i.e., droplets which split from each other. The liquid core is present at the center of the image. Each node contains relevant statistics from the breakup event.

```
// Rename the liquid core
MATCH(n:droplet)      // Find nodes with the "droplet" label
WHERE n.Event='None' // Node with Event = none is the liquid core
CALL apoc.refactor.rename.label("droplet","core",n)  // Rename
    node
YIELD committedOperations
RETURN committedOperations

// Identify and rename primary droplets
MATCH (n:droplet)
WHERE n.OldLID = 1    // Find droplets which broke off of the
    liquid core
WITH collect(n) as p // Compile a list of nodes matching criteria
CALL apoc.refactor.rename.label("droplet","primary",p)
YIELD committedOperations
RETURN committedOperations

// Rename secondary,tertiary, etc. (repeat until no "droplet"
    nodes remain)
MATCH (n:droplet),(d:primary)
WHERE n.OldLID = d.NewLID // Find "droplet" nodes that split from
    "primary"
WITH collect(n) as p
CALL apoc.refactor.rename.label("droplet","secondary",p)
YIELD committedOperations
RETURN committedOperations
```

This portion of code first identifies the core by finding the only node which has an event not equal to either breakup or coalescence. Then, it renames this droplet "core". Note: the "Awesome Procedures on Cypher" (APOC) library must be enabled to access the renaming features. Next, primary droplets are identified by finding all the droplet nodes which previously had a $\mathcal{L}$ (LID) equal to one (the liquid core has $\mathcal{L} = 1$). These droplets are renamed "primary". After primary droplets are identified, secondary droplets can be identified using a similar process. The algorithm loops through all droplets and looks for nodes that previously had $\mathcal{L}$ equal to the $\mathcal{L}$ of a primary droplet. This process can then be repeated for each subsequent breakup event until no more "droplet" nodes remain. This is very useful in order to analyze the evolution of droplets as they break up further. This will be analyzed in more detail in the Results section.

```
// Collect all breakup paths of droplets which break up six times
MATCH (n:sixth), (c:core), p = shortestPath((c)-[:Split*]->(n))
RETURN [d in nodes(p)| d.Volume]
```

The "shortestPath" function is used to extract the paths between specific nodes. In the above example, the shortest path from the liquid core to droplets that broke up six times is queried. Then, the volume of each droplet in that path is returned. This feature provides a simple method to extract useful relational data from the breakup process.

In addition to the Neo4j graphical database system and the Cypher Query Language, we utilized Python for analysis of results. Both programs have benefits and drawbacks, which led to the utilization of both. Neo4j allowed us to build paths very easily between droplets which breakup from each other or coalesce together. From this information, we can easily output a CSV that displays how statistics evolve with breakup events or coalescence events. Following the output of these CSV files, it is simple to analyze the data using Python's data science libraries. Additionally, multiple options exist to query or create graph databases

through Python. These include the official Neo4j driver for Python and the py2neo Python library. These prove to be very powerful tools because they allow users to combine python loops, if statements, and data science libraries with the unique organizational system of Neo4j. The py2neo library was used in the present work to build the time-series plots in section 2. See appendix C for an example python script querying a Neo4j database.

Application of Extraction Tool on Diesel Jet

Simulation Setup

To test the utility of the updated tool we ran a simulation inspired by a diesel injector. The simulation is the same as that described in Rubel and Owkes [159] and consists of a turbulent liquid injection into acquiescent air. The turbulence was computed from a preliminary turbulent pipe-flow simulation and the velocity field was stored and used as the inlet boundary condition for the liquid jet. Table 1 provides the parameters of the simulated jet. The simulation was run on 160 processors on the Hyalite High-Performance Computing Cluster at Montana State University. The dimensions of the computational mesh are $\mathcal{N}_x \times \mathcal{N}_y \times \mathcal{N}_z = 1024 \times 128 \times 128$. Note that the resolution of this simulation is not fine enough to accurately capture small-scale interface features or the smallest scales of turbulence. However, this simulation is sufficient to demonstrate the efficacy of the proposed tool. Future work will focus on applying the tool to a higher resolution simulation to more accurately extract information on physical phenomena.

Addressing Fictitious Events Confirmation

Section 2 discusses an update in the coalescence and breakup identification algorithm, which is intended to prevent the occurrence of fictitious events. In the original work [159], the fictitious events coincided with every breakup within the simulation, which forced the authors to institute a delay in the breakup identification portion of the code. This meant they only

Number	Definition	Value
Bulk Reynolds number	$\rho_l U_{\text{jet}} D_{\text{jet}} / \mu_l$	25,000
Bulk Weber number	$\rho_l U_{\text{jet}}^2 D_{\text{jet}} / \sigma_l$	10,000
Density ratio	ρ_l / ρ_g	40
Viscosity ratio	μ_l / μ_g	1.67
Domain length	L_x / D_{jet}	60
Domain widths	$L_{y,z} / D_{\text{jet}}$	7.5
Cells across diameter	$D_{\text{jet}} / \Delta x$	17.06
CFL number	$\lvert u \rvert_{\max} \Delta t / \Delta x$	0.4

Table 1: Non-dimensional parameters used in the diesel jet simulation.

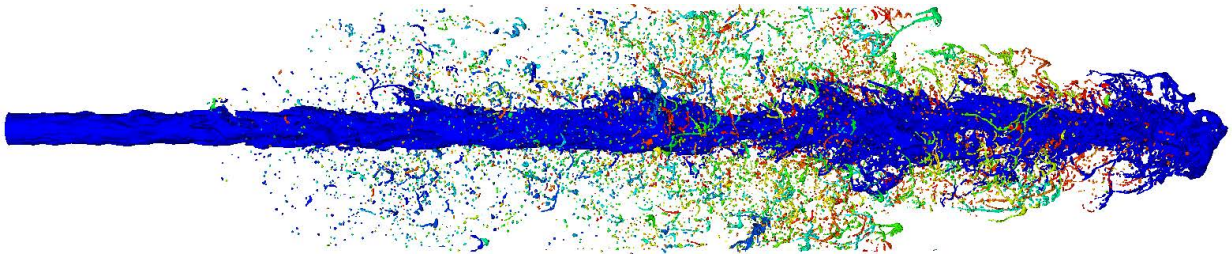


Figure 5: Rendering of the liquid jet run in simulation. $\mathcal{L}$ values at final timestep represented by colors with $\mathcal{L}$ ranging from 1 to 76,608 for the number of breakups that occurred. Rendered using VisIt [22].

identified breakup events every 10 timesteps. In the present work, breakup identification is performed on every timestep. It is reasonable in complex flow systems that droplets may breakup and then re-merge with their parent droplet, so the algorithm is not intended to prevent these phenomena from occurring altogether.

To confirm that the updated method is producing reasonable results, we utilized Neo4j to identify all the droplets which split and re-merged with their parent droplet. Then, the difference in time between the coalescence event and the initial breakup event was calculated. 17.93% of all droplets re-merged with their parent droplet, with the majority of those being primary droplets coalescing with the liquid core. This is a well-documented mechanism [174] and is expected in a round jet injected into quiescent gas. Analyzing only secondary droplets, we found that 9.31% of secondary droplets re-merged with their parent droplets. The fictitious events were documented to occur within one or two timesteps of breakup, but in the present simulation, the average time between coalescence and breakup is 61.7 μs which translates to about 300 timesteps. Given these statistics, we conclude that the updates made to the tool in the present work remedied the fictitious breakup and coalescence issue addressed in the original work and the identified events in the present work are consistent with physical events in the spray.

Secondary Atomization Analysis

Common descriptions of breakup regimes within an atomizing system involve discussion of primary atomization, i.e. droplets splitting from the liquid core, and secondary atomization, i.e., all further breakup. As seen in Fig. 6, many droplets are created after multiple breakup events. Further analysis shows that 53.6% of the total number of droplets formed in the test diesel jet broke up three or more times, with 17.0% percent breaking up at least five times. At later times in the simulation, this percentage becomes more pronounced. The yellow curve with triangle markers in Fig. 6 is the distribution of droplets at the final

timestep. 67.7% of drops present at this timestep broke up 3 or more times. Thus, a majority of the breakup in the simulation occurred after secondary breakup. This indicates that the final spray field is heavily dependent on mechanisms within these later atomization regimes, which are all generally grouped in secondary atomization.

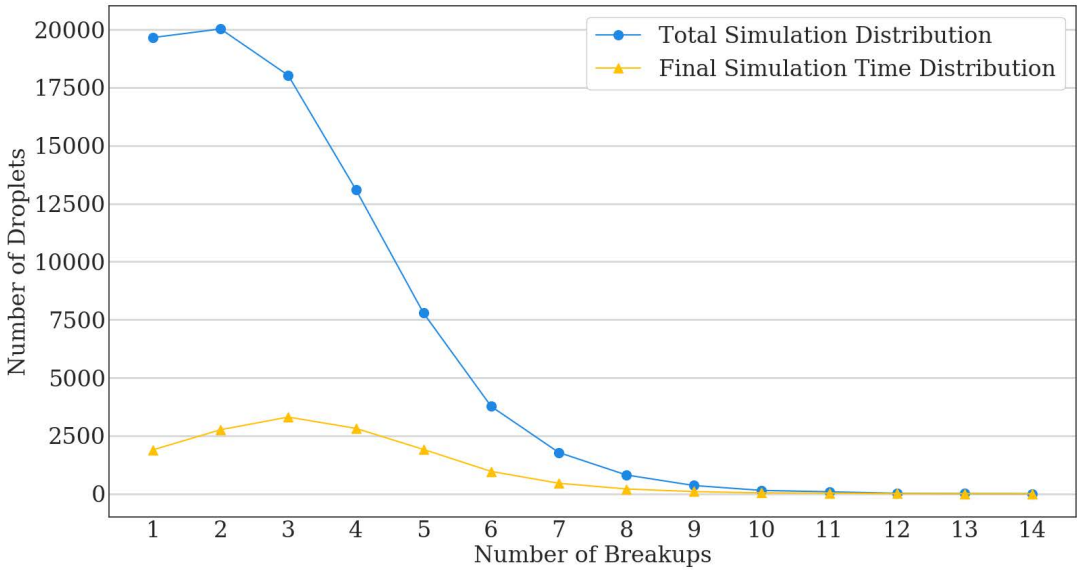


Figure 6: Plot displaying the distribution of breakup stages throughout the entire simulation (blue, circle nodes), e.g. about $7,600$ droplets underwent a 5^{th} breakup throughout this system. The yellow, triangle nodes represent the distribution of breakup stages at the final simulation time.

Coalescence Analysis

Coalescence is a vital process in the study of atomization. Coalescing droplets will result in larger droplets, and in turn, lead to more breakup. It is well understood that coalescence occurs in atomization, however, without access to efficient extraction methods, researchers have not been able to fully quantify or analyze how it affects a full atomizing system. A recent study by Prakash et al. [150] was able to identify coalescence event locations on a limited time domain within their jet in cross-flow simulation. Our tool provides the ability to analyze the

prevalence of coalescence events throughout entire atomizing systems. Furthermore, from the paths created by Neo4j, we can track the relationship between breakup and coalescence from the evolution of every liquid structure within a system. This provides statistically relevant insights into a little-understood process. The present simulation produced 76,608 breakup events and 57,908 coalescence events. Table 2 displays the average number of coalescence events for a given number of breakup events. These data were calculated using a path-finding algorithm in Neo4j. The algorithm finds the shortest path of breakup events from droplets at the given breakup stage back to the liquid core and counts the number of coalescence events along that path. The average number of coalescence events shown in the table is the average amongst all droplets that have broken up the associated number of times. The data in Table 2 indicates that the proportion of coalescence events increases as the jet evolves and more breakup occurs. This makes intuitive sense, because as the jet evolves, a high droplet density cloud develops, which leads to more droplet collisions.

Breakup Stage	Average Number of Coalescence Events	Ratio of Coalescence to Breakup
Secondary	1.077	0.5385
Third	1.899	0.6330
Fourth	2.898	0.7245
Fifth	3.974	0.7948
Sixth	5.241	0.8735
Seventh	6.823	0.9747
Eighth	8.276	1.035

Table 2: The average number of coalescence events between primary breakup and the breakup stage is listed in the second column. The third column displays the ratio of coalescence events to breakup events along breakup paths.

Local Flow Field Statistics

Another use for this tool is to better understand how the local flow field is affecting the atomization process. This is information that is exceedingly difficult, if not impossible, to

obtain through experimental methods. Moreover, gathering a sufficient sample of statistics on these local flow mechanisms has remained nonviable for those conducting numerical simulations. This work extracted preliminary data on the local flow field surrounding breakup events and calculated the resultant local Weber number, defined as $\mathrm{We}_{\mathrm{local}} = \rho_g U_s^2 L/\sigma$, where ρ_g is the gas velocity, U_s is the slip velocity (defined as $U_s = |\boldsymbol{u}_{\mathrm{liquid}} - \boldsymbol{u}_{\mathrm{gas}}|$), L is the characteristic length (in this case, it is the equivalent spherical diameter of the droplet prior to breakup), and σ is the surface tension coefficient. A probability density function (PDF) of the logarithm of $\mathrm{We}_{\mathrm{local}}$ number is developed from every breakup within the simulation (Fig. 7). The values are centered roughly about $\mathrm{We}_{\mathrm{local}} = 0.37$. These are very small values of the Weber number, which likely indicates that aerodynamic forces are not major factors in breakup in this system. This is logical because the simulation is of liquid injected into quiescent air.

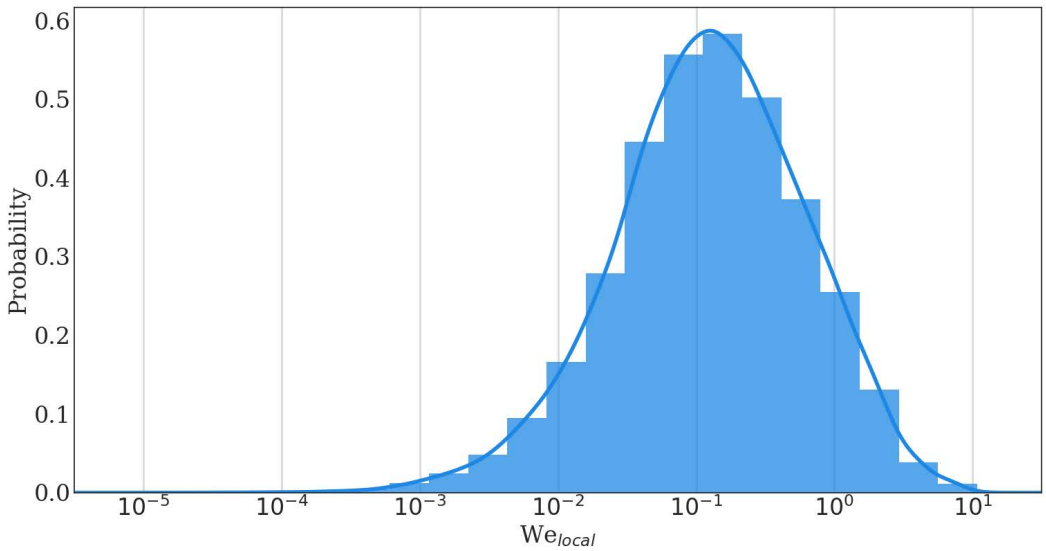


Figure 7: A probability density function of the local Weber Number associated with every breakup in the simulation.

Atomization Evolution

The evolution of local droplet characteristics and global jet development throughout the atomization process are important attributes, which can help elucidate the underlying physics of atomization. Understanding how droplets change from the first breakup off of the liquid core to a final droplet in dilute spray could provide useful information to atomization model developers seeking to not only describe the final spray formation, but also the intermediate spray development. The tool provided in this work enables researchers to extract droplet statistics throughout a simulation and analyze the evolution of these droplets and the system as a whole.

Droplet Shape and Size Evolution Fig. 8 displays a probability density function of the equivalent spherical diameter of droplets as a function of the number of breakup events. A reasonable trend toward smaller and more uniformly sized droplets is seen. Fig. 9 displays the change in diameter of droplets between breakup stages. Notice that not all values are negative. This further confirms the prevalence of coalescence events within this system. Positive values in the figure indicate that a significant portion of droplets undergo coalescence and increase in size between breakup events. Additionally, notice the tendency of droplets to change size less as more breakup events occur, with the PDFs becoming more narrow and centered around 0 μm. The coarse resolution of the present study affects the accuracy of these values related to a physical system. However, they display the ability to analyze large quantities of statistics relevant to the evolution of liquid shapes and sizes throughout an atomization simulation.

Time Evolution Because the extraction tool samples from every timestep in which breakup occurs. Thus, once the jet develops, breakup occurs on every timestep, providing high-resolution time-series statistics describing the evolution of the jet. Fig. 10 shows an

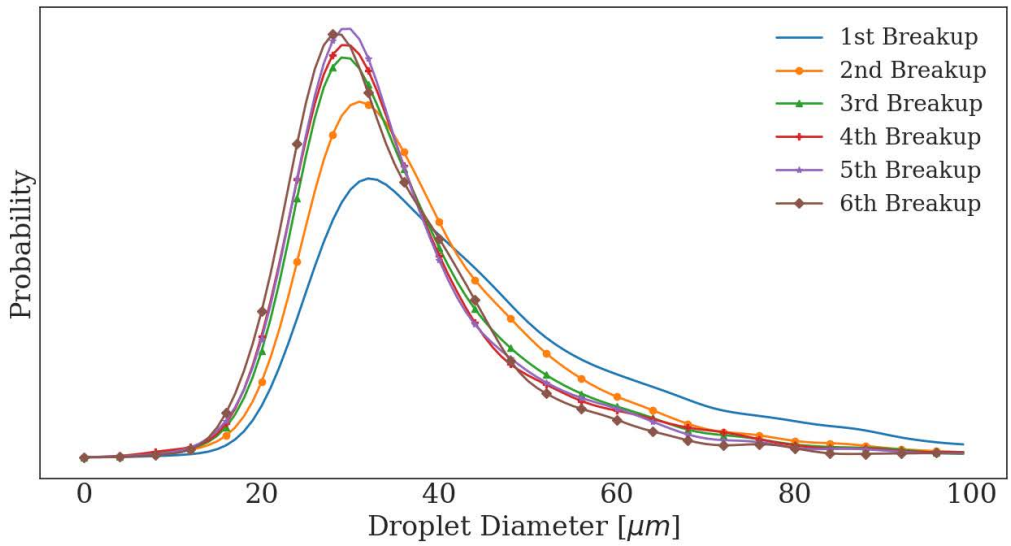


Figure 8: A probability density function of droplet diameter as a function of the number of breakups

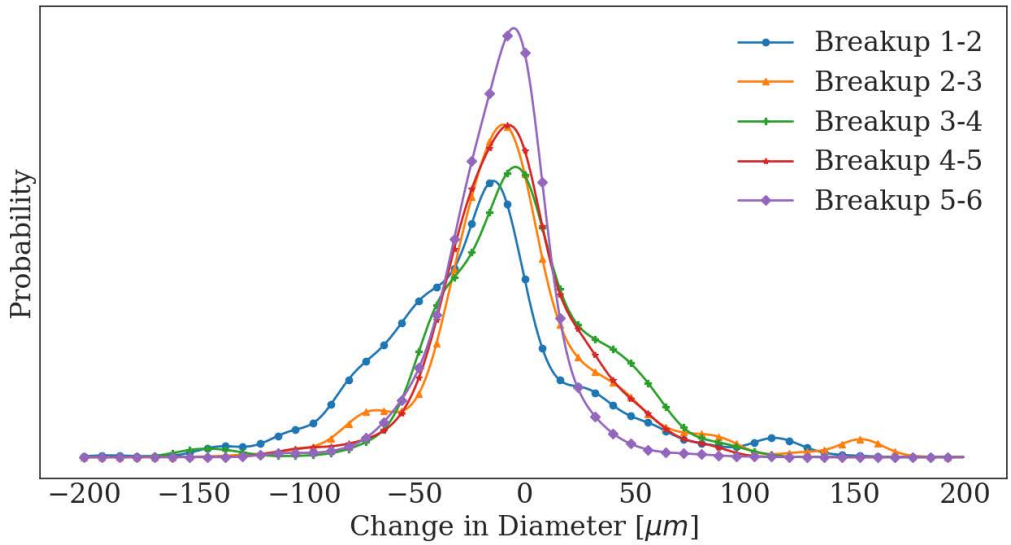


Figure 9: A probability density function displaying the change in diameter between breakup events.

example of some of the time-dependent statistics that were extracted in the present work. These plots describe the rate of droplet development in the simulation. It is clear from this plot that secondary atomization quickly dominates droplet production in this system. Future work will focus on improving the extracted quantities and further explore how atomizing systems develop through time.

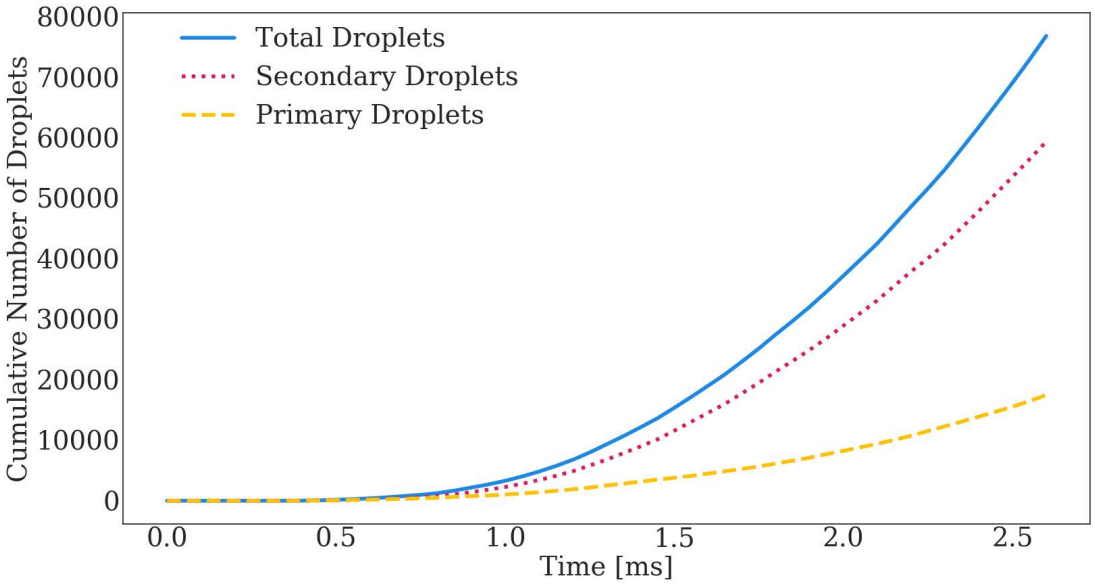


Figure 10: The cumulative number of primary, secondary, and total droplets. Note: secondary droplets in this image represent the classical definition of secondary droplets (i.e. all breakup following primary).

Computational Cost

The extraction tool is presented as an addition to atomization simulations rather than a separate post-processing method. The purpose is to 1) provide access to information that would otherwise be impossible to obtain via traditional post-processing, and 2) extract the information efficiently. Thus, the added computational cost of incorporating the tool into an already computationally intensive simulation is an important consideration. The tool assigns $\mathcal{S}$, transports $\mathcal{L}$, identifies breakup and coalescence, and outputs data. We

conducted a timing study using 175 timesteps at the end of the Diesel jet simulation. Since the tool becomes more expensive as more liquid structures are created and as more breakup and coalescence occur, this analysis represents the most expensive iterations of the tool in the present simulation. Table 3 displays the average percent of each timestep used by the four main processes of the tool. $\mathcal{L}$ transport is incorporated into the VOF scheme and the existing flux calculation. The cost of the $\mathcal{L}$ flux is negligible and only requires 0.01% of the total timestep time. Multiple $\mathcal{L}$ values can exist within a single flux, which does require a minor amount of memory. The identification of liquid structures and assignment of $\mathcal{S}$ values was not trivial and required 2.27% of each timestep. This is due to the cost of the band-growth algorithm that identifies each unique structure in the domain. At this stage in the simulation, it is identifying ∼15,000 structures per timestep. Other, more efficient, connected component labeling algorithms could be used in place of this to further reduce the computational cost. The breakup and coalescence event identifications involve simple list comparisons and only cost 0.28% of the timestep. Data extraction is not costly, only using 0.001% of each timestep. This is reasonable because it simply involves writing lines to a CSV file. The total cost for the entire tool was 2.56% of the total timestep in this simulation. This cost can be reduced by improving the $\mathcal{S}$ assignment algorithm.

Process	Percent of Timestep [%]
$\mathcal{S}$ Assignment	2.27
$\mathcal{L}$ Transport	0.01
Event Identification	0.28
Data Extraction	0.001

Table 3: The cost of each process within the tool. Displayed as the average percent of time per timestep. Sampled from 175 timesteps at the end of the simulation.

Conclusions

An extraction tool for numerical simulations of atomization was developed and tested. Improvements to the methodology originally proposed by Rubel and Owkes [159] were made. Improvements include 1) an updated algorithm to stop fictitious identification of merge and split events, 2) enabling a wider temporal sampling range, and 3) the extraction of more atomization statistics. Improvements were made through an extensive rewrite of algorithms within the tool. Two major changes were made. The first was the incorporation of a key logical statement in the portion of the code that identifies independent liquid structures. This prevents non-physical coalescence when liquid structures exist in adjacent computational cells. Additionally, the criteria for identifying coalescence were updated by improving the transport and tracking scheme for identification numbers ($\mathcal{L}$), which move with the liquid. Both of these changes together remedied the identification of fictitious breakup and coalescence events, present in the previous work. In addition, the updated tool now samples data from both the timestep preceding events as well as the timestep following them. This provides insights into the conditions which led to breakup as well as the resultant structures from breakup or coalescence.

A diesel-type jet was simulated and droplet statistics were extracted. The tool displayed utility for extracting data relevant to atomization model developers as well as providing previously inaccessible information on the underlying mechanisms of atomization. New analyses of the local flow field, the breakup evolution of liquid structures, and coalescence dynamics were introduced. The data extracted are preliminary, but show promise for the tool's utility in future high-resolution studies.

Future work will focus on the implementation of the tool into high-resolution atomization simulations to extract useful data to help better quantify atomization mechanisms on global and local scales. Special effort will be given to aiding reduced-order model developers

to develop relevant statistics for creating more accurate models. Additionally, we will focus on the improvement of sampling methods, particularly the local flow field to attempt to extract not only velocity magnitudes but also topological flow data to better understand the small-scale turbulence that affects liquid breakup.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1749779.

Computational efforts were performed on the Hyalite High Performance Computing System, operated and supported by University Information Technology Research Cyberinfrastructure at Montana State University.

NUMERICAL STUDY ON NEWTONIAN AND NON-NEWTONIAN DROPLET AEROBREAKUP MECHANISMS

Contribution of Authors and Co-Authors

Manuscript in following chapter

Author: Brendan Christensen

Contributions: Methodology, Investigation, Data Curation, Formal Analysis, Visualization, Writing - Original Draft

Author: Mark Owkes

Contributions: Methodology, Writing - Review & Editing, Supervision, Funding Acquisition

Author : Dominique Legendre

Contributions: Conceptualization, Supervision, Writing - Review & Editing

Manuscript Information

Brendan Christensen, Mark Owkes, and Dominique Legendre

Journal of Fluid Mechanics

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
_X_ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Elsevier Ltd.

Abstract

This study investigates the aerodynamic breakup of Newtonian and shear-thinning non-Newtonian droplets in high-speed gas flows using high-fidelity numerical simulations. The work utilizes the Atomization Simulation Statistics Extraction Tool (ASSET) to extract novel quantitative information from breakup and coalescence events, providing insights into the underlying physics. Numerous breakup mechanisms were identified, and an analysis of instability transitions in the shear-induced entrainment (SIE) to Rayleigh-Taylor piercing (RTP) transition regime was conducted. Results highlight the prevalence of ligament-driven breakup in these regimes, confirmed by analyzing droplets at the moment of fragmentation. Additionally, ASSET is used to examine secondary atomization cascade dynamics, providing detailed breakup statistics across multiple scales. The effect of effective viscosity in shear-thinning fluids is analyzed, revealing its role in shaping fragmentation patterns and transition behavior. These findings enhance understanding of droplet breakup in high-speed flows and provide valuable data for improving predictive models in atomization-driven applications.

Introduction

Liquid atomization due to aerodynamic breakup is a process with numerous industrial and environmental applications. The study of droplet breakup in a high-speed gas stream is particularly well studied and applies to a wide range of fields, including fuel injection in combustion engines, irrigation systems, pharmaceutical manufacturing, and naturally in rainfall and sea-spray. However, much of the existing research has focused on Newtonian fluids, leaving a gap in our understanding of how non-Newtonian, shear-thinning fluids behave under similar conditions. These fluids are used in scenarios where it is necessary for the liquid to have low viscosity during spray formation (high-shear rates) and a high viscosity on deposition (low shear-rates). Such conditions are frequently encountered in

large-scale spraying operations such as aerial firefighting, or when working with polymeric and slurry-based fluids, where controlling the rheological properties can significantly affect the final atomization pattern.

Many practical spray systems involve large-diameter droplets or nozzles and high relative gas-liquid speeds (e.g., industrial nozzles, agricultural sprayers, and environmental dispersal processes). A key example, and a motivation for this work, is the injection of a large liquid column into a high-speed crossflow, as seen in aerial firefighting systems. These systems operate in a regime that has received limited detailed study, particularly when the liquid jet fragments into large structures (referred to as meteors in [101]) that undergo further complex breakup due to their large size and high relative velocity. Capturing these secondary breakup processes is often infeasible in large-scale jet breakup studies [18] because of resolution. Most existing research on aerodynamic-force-driven liquid breakup focuses on smaller, millimeter-scale jets with moderate Weber numbers [17, 119]. At these scales and relative speeds, and given available experimental techniques, tracking high-Weber number drops beyond their initial breakup is challenging. More recent work by [157] has explored larger-diameter nozzles, where gas Weber numbers can exceed 10^4. However, understanding the full atomization process—spanning primary jet breakup, secondary drop formation, and the resulting droplet size distributions—remains difficult in these large-scale, high-speed flows.

Despite the practical importance of droplet breakup in large-scale systems, fundamental aspects of breakup morphology and fragment statistics remain poorly characterized. Key challenges include understanding the interaction of Rayleigh–Taylor and Kelvin–Helmholtz instabilities at high Weber numbers and the role of viscosity in shear-dominated regions. These complexities are further amplified in shear-thinning fluids, where local effective viscosity varies with shear rate [19, 107, 224].

A major limitation is the lack of quantitative data describing the full breakup cascade.

While experiments provide valuable insights, they struggle to capture the three-dimensional dynamics of spray formation, particularly as dense spray clouds obscure internal mechanisms. Numerical simulations offer a wealth of quantitative data and continue to improve as computational methods advance. However, they are often underutilized due to the sheer scale of data they produce, with tens to hundreds of terabytes generated per simulation. Much of this information is discarded simply because no efficient method exists to process and analyze datasets of this magnitude.

To address the complexities and gaps in understanding the aerodynamic breakup of both Newtonian and non-Newtonian fluids, this study will analyze high-fidelity numerical simulations of both systems. To efficiently utilize the simulations, this study incorporates the Atomization Simulation Statistics Extraction Tool (ASSET) [24, 159]. ASSET offers a unique capability to extract detailed, quantitative data from atomization simulations. By analyzing local statistics from each breakup event in the simulation, such as the shape, size, location, and local flow dynamics, ASSET provides access to statistics describing the underlying physics of atomization. These data are stored in a graph database, linking parent droplets to their resulting fragments and offering a detailed history of the breakup cascade. This approach not only improves our understanding of aerial firefighting related atomization stages, but also provides novel insights into aerodynamic-driven atomization in general.

The aim of this work is twofold. First, we seek to improve fundamental understanding of large-scale, high-Weber-number droplet aerobreakup, including new insights into the behavior of shear-thinning liquids. Second, we demonstrate the usefulness of high-fidelity simulations combined with advanced statistics-extraction tools, with an eye toward ultimately informing reduced-order models for spray prediction across a spectrum of industrial, agricultural, and environmental applications.

Background

To better understand the small-scale processes of large droplet atomization, it is important to review the established theories and classifications of droplet aerobreakup. The phenomenon of droplet breakup in gas streams, also referred to as aerodynamic breakup or aerobreakup is generally separated into morphologies based on the droplet-derived Weber ($We_g = \rho_g U_g^2 D_0/\sigma$) and Ohnesorge ($Oh = \mu_l/\sqrt{\rho_l \sigma D_0}$) numbers [43, 57, 68, 96, 134], where g and l subscripts are gas and liquid, respectively, ρ is density, μ is the dynamic viscosity, D_0 is the initial drop diameter, and σ is the surface tension coefficient. These morphologies, commonly categorized in order of increasing We (for low $Oh < 0.1$) are vibrational, bag, multi-mode, sheet-thinning, and catastrophic. A review by [57] notes that the transition from one regime to the next is essentially continuous, and sharp delineations have not been agreed upon. Additionally, [191] challenges the traditional understanding of the catastrophic breakup regime, finding it to be an artifact of shadowgraph visualizations. The authors instead define two criticalities to describe aerodynamic breakup. For the range $10 < We < 10^2$ breakup occurs due to Rayleigh-Taylor piercing (RTP), which encompasses bag and multimode breakup. For $We > 10^3$, breakup occurs due to shear-induced entrainment (SIE), which encompasses sheet-thinning breakup and is the terminal regime. The system being studied in this work is inspired by aerial firefighting, which involves high shear-rates as liquid is dropped into turbulent atmosphere from a subsonic aircraft. The expected We of large liquid structures in this system are near or above 10^3. Consequently, the granularity of the regimes within RTP are not necessary, and it is sufficient to describe the regimes in this system as being near the RTP-SIE transition region (Table 4).

Recent research on subsonic aerobreakup in the SIE and RTP-SIE transtion region exists largely in the form of experimental studies conducted with 2-D visualization methods [78, 170, 191, 192] and three-dimensional numerical works [39, 79, 80, 124]. These studies primarily aim to refine breakup regime maps and provide qualitative descriptions of

the mechanisms driving breakup as they transition from Rayleigh-Taylor to Kelvin-Helmholtz instabilities. However, there has been little focus on quantitatively describing the entire breakup cascade—starting from sheet or ligament formation, through the creation of large droplets, and continuing with their subsequent breakup. This gap in the research highlights the need for a more comprehensive understanding that can capture the full dynamics of droplet breakup across these regimes.

This work explores both Newtonian and non-Newtonian shear-thinning fluids. Relative to the study of Newtonian liquids, there is limited research on non-Newtonian droplet breakup. Existing literature on the subject of non-Newtonian liquids includes studies on viscoelastic fluids [85, 97, 193] and shear-thickening fluids [126, 212]. There is very limited research on the breakup dynamics of shear-thinning fluids. [224] captured images of shear-thinning coal water slurry (CWS) droplets with Weber numbers up to 10^3, producing new Weber and Ohnesorge numbers to develop a map for CWS fluids. Recent numerical studies by [19] and [107] describe new breakup regime mappings for shear-thinning fluids over a range of $13 < We < 53$ and $5 < We < 35$, respectively. These studies provide insights into the breakup mechanisms of shear thinning fluids within the RTP regime. The present work will aim to contribute to a more comprehensive understanding of shear-thinning breakup mechanisms associated with high Weber numbers ($\approx 10^3$) in the RTP-SIE transition regime, which is still largely unexplored.

Statement of Purpose

The central aim of this study is to investigate and quantify the aerodynamic breakup dynamics of both Newtonian (water) and shear-thinning (non-Newtonian) droplets in a high-speed, turbulent gas flow relevant to large-scale practical applications. We conduct fully three-dimensional, high-fidelity simulations at large density ratios, pushing computational capabilities to resolve the multi-scale instabilities that arise under high Weber-number

conditions. This research will employ the Atomization Simulation Statistics Extraction Tool (ASSET) to generate novel quantitative insights into the mechanisms driving droplet formation, breakup, and distribution in such complex environments.

Through this approach, we elucidate the interactions between shear-driven mechanisms and Rayleigh–Taylor-type instabilities, highlighting the roles that viscosity and droplet shape play in both primary and secondary atomization. The statistical insights extracted provide a granular view of fragmentation pathways—ranging from ligament and sheet formation to the eventual distribution of stable droplets. By comparing Newtonian and shear-thinning fluids within the same large-scale framework, we also clarify how non-Newtonian viscosity alters the breakup phenomenology. Ultimately, these new quantitative findings advance our fundamental understanding of droplet aerobreakup and lay the groundwork for improved reduced-order models applicable to a range of industrial, agricultural, and environmental spraying processes.

Methodology

Numerical Methods

This study employs the NGA computational platform [33, 35, 138, 139] to perform all simulations. The system being simulated is subsonic aerobreakup of water and shear-thinning liquid droplets. Thus, NGA solves the two-phase, incompressible formulations of the Navier-Stokes equations for mass and momentum conservation, expressed as

$$\frac{\partial \rho_\phi}{\partial t} + \nabla \cdot (\rho_\phi \boldsymbol{u}_\phi) = 0 \tag{9}$$

and

$$\frac{\partial \rho_\phi \boldsymbol{u}_\phi}{\partial t} + \nabla \cdot (\rho_\phi \boldsymbol{u}_\phi \otimes \boldsymbol{u}_\phi) = -\nabla p_\phi + \nabla \cdot (\mu_\phi [\nabla \boldsymbol{u}_\phi + \nabla \boldsymbol{u}_\phi^T]) + \rho_\phi \boldsymbol{g}, \tag{10}$$

where ρ_ϕ is the density, $\boldsymbol{u}_\phi$ the velocity vector, p_ϕ the pressure, μ_ϕ the dynamic viscosity, t the time, $\boldsymbol{g}$ the gravitational acceleration vector, and ϕ the phase indicator, with $\phi = g$ representing gas and $\phi = l$ representing liquid. These equations are solved on a staggered Cartesian grid, with scalar quantities like pressure stored at cell centers and velocity components at cell faces.

Time discretization is handled using an iterative Crank-Nicolson scheme, with a semi-implicit correction applied at each sub-iteration [23]. Away from the phase interface, mass, momentum, and other scalar quantities are transported using conservative, high-order finite difference operators [33]. Near the interface, a geometric volume-of-fluid (VOF) method is employed to represent the interface, momentum, and scalars. The VOF field is transported using unsplit semi-Lagrangian fluxes [138]. A refined grid near the interface ensures near-machine precision conservation of mass and momentum while also doubling the resolution of the interface [139]. The interface is reconstructed using a piece-wise linear interface reconstruction (PLIC) [218], with interface normal vectors computed using the efficient least-squares VOF interface reconstruction algorithm (ELVIRA) [146]. The pressure Poisson equation is solved using the ghost fluid method [46] and a black-box solver [30]. Interface curvature is computed using the adjustable curvature evaluation scale (ACES) method [136].

NGA is fully parallelized using a message passing interface (MPI) protocol, enabling it to scale efficiently to tens of thousands of cores [32]. This scalability, combined with its robust capabilities, has made NGA a valuable tool for investigating a wide range of atomization applications [34, 95, 138, 171]. These studies highlight NGA's ability to handle complex, real-world systems, making it well-suited for simulations aimed at uncovering the underlying physics of such processes.

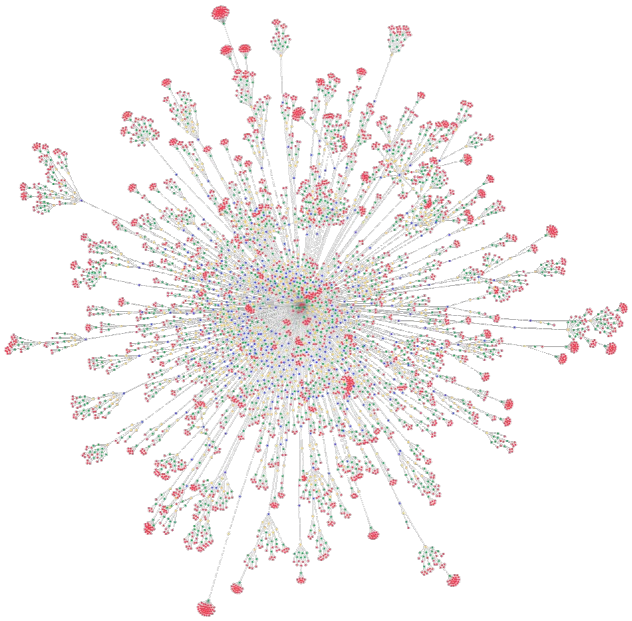


Figure 11: Neo4j graph database from an diesel jet atomization simulation [24]. Dots represent droplets, with relationships between them representing split events connecting child and parent droplets. The color represents how many breakup events led to that droplet (i.e. red = four breakup events, green = three, yellow = two, blue = one).

Atomization Simulation Statistics Extraction Tool

A key objective of this work is to gain deeper insights into the underlying physics of atomization in relevant systems. To achieve this, the Atomization Simulation Statistics Extraction Tool (ASSET), as described by [24], is applied to each simulation in this study. The primary purpose of using ASSET is to quantify the breakup process and track the evolution of liquid fragments over time and monitor varying local conditions associated with breakup. This approach provides novel analysis of both the well-studied water droplet aerobreakup system and the less understood shear-thinning liquid aerobreakup system. The outcome of each simulation is a graph database (see Figure 11) that traces every droplet back to its parent droplet(s), ultimately linking to the initial large droplet. Each entry in the database includes detailed statistics such as location, size, shape, and local Weber number.

Shear-Thinning Viscosity The shear-thinning viscosity is modeled using a power-law relation, described as

$$\mu_{\text{Le}} = kD^{n-1} \tag{11}$$

where μ_{Le} is the effective viscosity of the liquid, k is the consistency, n is the fluid index, and D is the strain-rate. For this study, the common fire-retardant, Phos-Chek XA is used to inform the parameters of the power law, according to studies done by [7]: $k = 6.78$ Pa$\cdot$s^n and $n = 0.67$. With the strain-rate field readily available in the numerical simulation, this calculation is done simply on each timestep. To prevent infinite viscosity at low shear, the viscosity reverts to k at zero shear, which is consistent with the findings of [7].

Gas Domain An additional computational challenge in similar simulations is handling compressibility of gas. Generally, these simulations are replicating shock-tube experiments, which rely on compressibility to analyze the effects of shock waves on liquid breakup [39, 124]. The application that motivated this study—aerial firefighting—entails the release of large quantities of liquid from planes flying at high speeds (generally $\approx$ 70m/s). At these velocities, the mach number in the gas remains primarily below 0.3. However, an a priori analysis of the velocity field displays that late in the simulation as the initial drop flattens, the velocity in the gas phase at the outer edges accelerates to reach velocities as high as 120 m/s, corresponding to a mach number of 0.35. While this exceeds the conventional incompressibility threshold, it result in only a $\approx 3.5\%$ density variation. Furthermore, compressibility effects would be limited to the gas phase, occurring in a relatively small about the equator of the droplet and only in the latter third of the simulation. As such, the computational cost of fully resolving gas compressibility far outweighs its minimal predicted impact on the breakup process. Thus, this work assumes the gas to be incompressible.

The gas domain is intended to represent atmospheric conditions when an aircraft drops water or fire-retardant into it at high-speeds. The domain was initialized using a

homogeneous isotropic turbulence (HIT) simulation for the initial condition as well as in the inflow on the x^- face. Common aircraft speed of 70 m/s [101] was used as the freestream velocity (U_g) added to the HIT-derived turbulence. The HIT simulation was run so that the velocities fall in the range $\pm 0.05 U_0$, this can be seen in Figure 12b, which is plotted within the range of 65 to 75 m/s.

Simulation Configurations

Five simulations were conducted to analyze the aerodynamic breakup of Newtonian and non-Newtonian droplets. Three simulations of water droplets (W1, W2, W3) were performed at increasing grid resolutions to assess the influence of grid refinement on the extracted breakup statistics. The sizes of drops considered here are 1cm in diameter. Finally, two simulations of a shear-thinning fluid (ST1 and ST2) were conducted at grid resolutions equivalent to W1 and W2, respectively, to evaluate the effects of non-Newtonian viscosity on the breakup process. In the shear-thinning cases, all physical parameters—except viscosity—are identical to those for water. See Table 4 for relevant parameters and non-dimensional numbers.

This study represents one of the highest resolution three-dimensional investigation of droplet aerobreakup mechanisms to date, with cells across diameter (CAD) values exceeding those of other similar works [39, 124]. Importantly, the simulations achieve a density ratio of $\rho_l/\rho_g = 816$, consistent with practical applications such as fuel atomization and aerial firefighting. This contrasts with many previous numerical studies, which rely on reduced density ratios for ease of computation. Additionally, the present simulations are fully three-dimensional, without imposing symmetry constraints that might oversimplify the breakup process. By resolving fine-scale features in both the liquid and gas phases, these simulations push the limits of direct numerical simulation (DNS) for droplet aerobreakup, capturing the complex process of instability formation, and fragmentation across a range of scales in a

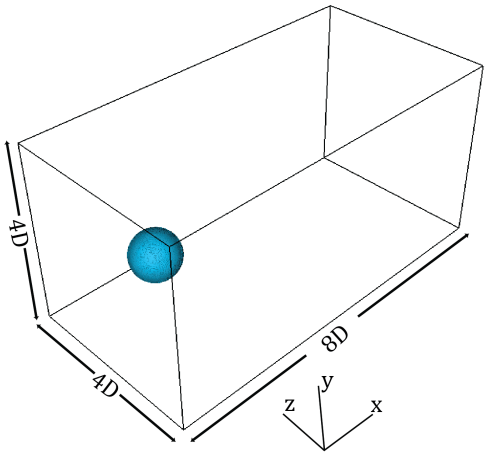
Parameters	Description	W1	W2	W3	ST1	ST2
CAD	D_0/dx	150	200	300	150	200
Bulk Reynolds	$\rho_g U_g D_0/\mu_l$	4.74×10^4	4.74×10^4	4.74×10^4	NN	NN
Bulk Weber	$\rho_g U_g^2 D_0/\sigma$	825	825	825	825	825
Bulk Ohnesorge	$\mu_l/(\rho_l \sigma D_0)^{1/2}$	1.17×10^{-3}	1.17×10^{-3}	1.17×10^{-3}	NN	NN
Grid Cell Weber	$\rho_g U_g^2 dx/\sigma$	10.98	8.251	5.504	10.98	8.251

Table 4: Non-dimensional parameters for the simulations. CAD is cells across diameter, simulations are identified as W for water and ST for shear-thinning, 1,2,3 for progressively finer resolution, and L is the larger diameter case. NN means non-Newtonian, thus variable Reynolds numbers.

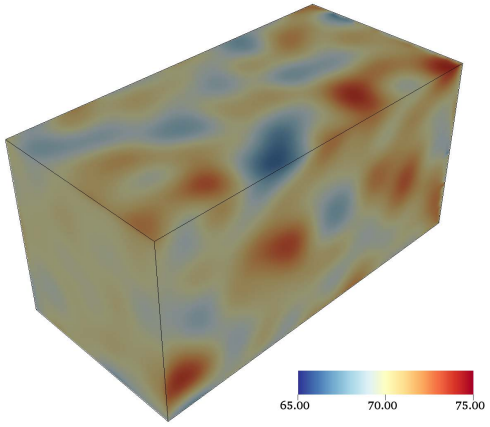
practical application.

Resolution Considerations Previous research has shown that resolving down to the smallest droplet scales, while desirable, often requires prohibitively fine mesh resolutions and significant computational resources [105, 123]. In practice, many studies justify focusing on the larger scales by noting that the early stages of breakup and large-scale liquid fragmentation, which are dominant mechanisms affecting the overall behavior of the spray [82]. Fine-scale processes such as secondary droplet breakup, although important for determining the final droplet size distribution, are often neglected or approximated due to the computational limitations of resolving them directly. This trade-off enables simulations to capture the large-scale dynamics while remaining computationally feasible.

The present work focuses on the well-resolved stages of fragmentation, which are key to understanding the complete atomization process. By utilizing ASSET to quantitatively describe these events, we extract valuable insights into the evolution of liquid fragments and the dominant instabilities, which lead to primary breakup. Even without resolving the smallest-scale physics, this approach provides a novel description of the primary breakup regime, enabling us to capture a quantitative and statistically relevant description of breakup,



(a) Simulation domain with initial droplet in place.



(b) The x-velocity plotted between 65 to 75 m/s to show HIT-derived turbulence.

Figure 12: Simulation setup images.

previously inaccessible with traditional experimental and simulation methods.

To assess the effect of resolution on the simulations, we run three different grid resolutions with cell sizes of $dx = [6.67 \times 10^{-3}, 5.00 \times 10^{-3}, 3.33 \times 10^{-3}]D_0$. In simulations W2, W3, and ST2 (see Table 4), the expected Weber number of a secondary droplet with diameter the size of a grid cell (i.e. $D_{eq} = dx$) and velocity equal to the freestream (70 m/s) is below 10. This indicates that further secondary breakup at these scales is unlikely. This justifies that the dominant mechanisms in the secondary breakup regime can be adequately captured.

Characteristic Time The results are presented relative to a characteristic deformation time

$$T = tU_g D_0^{-1} \sqrt{\rho_g/\rho_l}, \tag{12}$$

where t is physical time, D_0 is the initial drop diameter, U_g is the freestream velocity, and ρ_g/ρ_l is the gas-to-liquid density ratio. This time is used by numerous studies in literature

to describe drop aerobreakup [78, 145, 156, 191].

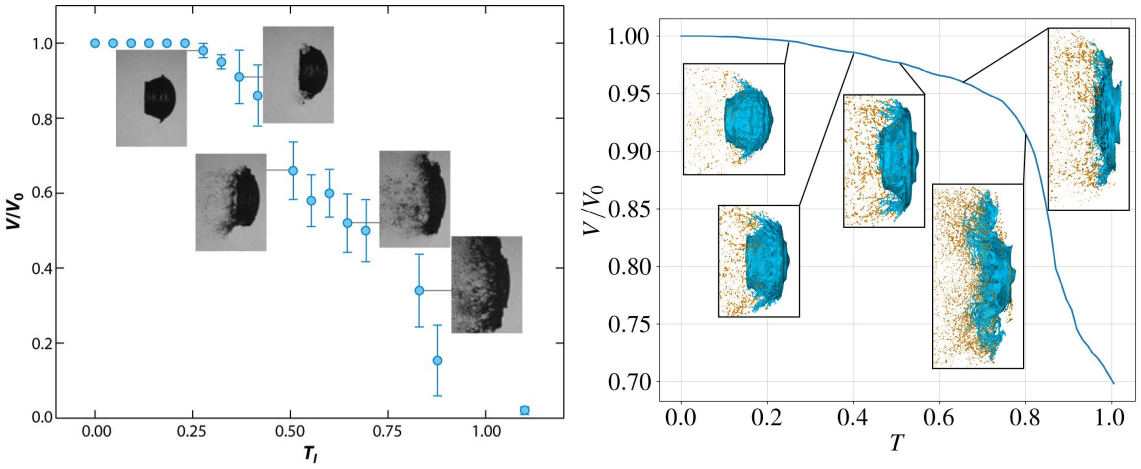
Comparison to Experimental Studies

Figure 13 displays an experimental water aerobreakup case from [193] with very similar parameters to the small-diameter water simulations run in this study ($We = [780, 825]$; $Oh = [2.4 \times 10^{-3}, 1.17 \times 10^{-3}]$; $Re_g = [2.2 \times 10^4, 4.74 \times 10^4]$, respectively). The plots illustrate V/V_0, which represents the proportion of the original droplet volume that remains intact over time. This metric provides an essential reference for comparing the progression of breakup events across both simulations and experiments.

Overall, the simulated droplet deformation aligns well with the experimental observations. Initially, a lip forms as surface waves develop; simultaneously, the downwind side of the droplet flattens and gradually folds inward, while the upwind side continues to extend with the flow. Consequently, the droplet transitions into a radially widened, flattened shape. As the droplet's rim becomes exposed to the high-speed gas flow, curling instabilities appear around its edges—visible in the final image frames of both the experimental and simulated snapshots ($T = 0.8$). In the later stages, a "backward-facing bag" emerges, a feature also reported experimentally and numerically in [39]. These parallels highlight the similarities between our simulations and established experimental work.

However, one notable difference is observed in the later stages of deformation: in our simulations, a wave and appendage develop on the upwind side of the droplet, whereas the experimental results show a smoother surface. This discrepancy is likely a minor effect of the computational grid, where deformation occurs due to the non-infinitesimal size of the stagnation point. While this effect may introduce some artifacts, it does not appear to significantly alter the overall breakup behavior or qualitative agreement between the simulations and experiments.

Additionally, while the qualitative agreement is strong, the values of V/V_0 show



(a) (b)

Figure 13: Mass loss of initial droplet with images at the same non-dimensional times T. Experimental parameters: $We = 780$, $Oh = 2.4 \times 10^{-3}$, $Re_g = 2.2 \times 10^4$. Experimental images reproduced with permission.

significant differences between the simulations and the experimental data. This discrepancy is likely due to challenges inherent in estimating droplet volume using two-dimensional images, as experimental setups are limited in their ability to capture the true three-dimensional dynamics of the process. Despite these limitations, the experimental images provided by [193] remain invaluable for visualizing the aerobreakup process and serve as an excellent reference for comparison with our simulations. The qualitative similarities observed across both studies reinforce the reliability of our numerical framework.

Breakup Morphologies

The breakup of liquid droplets in high-speed gas flows follows a complex sequence of deformation and fragmentation processes, influenced by both fluid properties and aerodynamic forces. Understanding these breakup morphologies is crucial for ultimately predicting spray formation in any application. This section examines the key stages of droplet breakup, beginning with large-scale deformation, which dictates the initial response of the

liquid structure to aerodynamic stresses. We then analyze the formation of early instabilities, their wavelength evolution, and the mechanisms that govern primary droplet shedding. Finally, we explore the transition to late-stage breakup, where competing instability modes and material properties drive the final disintegration of the droplet into secondary structures.

Large-Scale Deformation

Before exploring smaller-scale breakup mechanisms, it is important to consider the large-scale deformation of the initial drop. As the drop changes shape, it interacts with the surrounding gas flow differently (i.e. changing from SIE to RTP), which results in a temporal evolution of primary breakup mechanisms. Figure 14 shows the large-scale deformation trends for both water and shear-thinning cases. All simulated drop cases initially undergo significant large-scale deformation due to their high Weber number. While drops begin spherical, they begin to expand radially, and flatten into a "disk-like" shape due to the impact of the dynamic air pressure. The dominant breakup mechanisms change depending on the shape of the large coherent liquid drop. When initially spherical, stripping is a dominant mechanism in all cases. The outer radial edge begins to stretch downstream, creating a sheet, which extends in the direction of flow. However, as the initial drop flattens, the angle of attack of the airstream on the drop increases, leading to the formation of bags in the water drop (seen clearly in Figure 15), which dominates the breakup at this point. The non-Newtonian drop does not undergo bag breakup. Following bag breakup, the drop undergoes stripping again on the down-stream facing sheets, stretching them further in the direction of the flow. The combination of stripping and complex multi-bag breakup is indicative of the RTP-SIE transition regime, which should exhibit qualities from both Rayleigh-Taylor (RTP) and Kelvin-Helmholtz (SIE) instabilities. It should be noted that our simulations do not capture the entire droplet lifetime through the final, fully dispersed state. Computational constraints restrict the timescale over which we can track the resolved

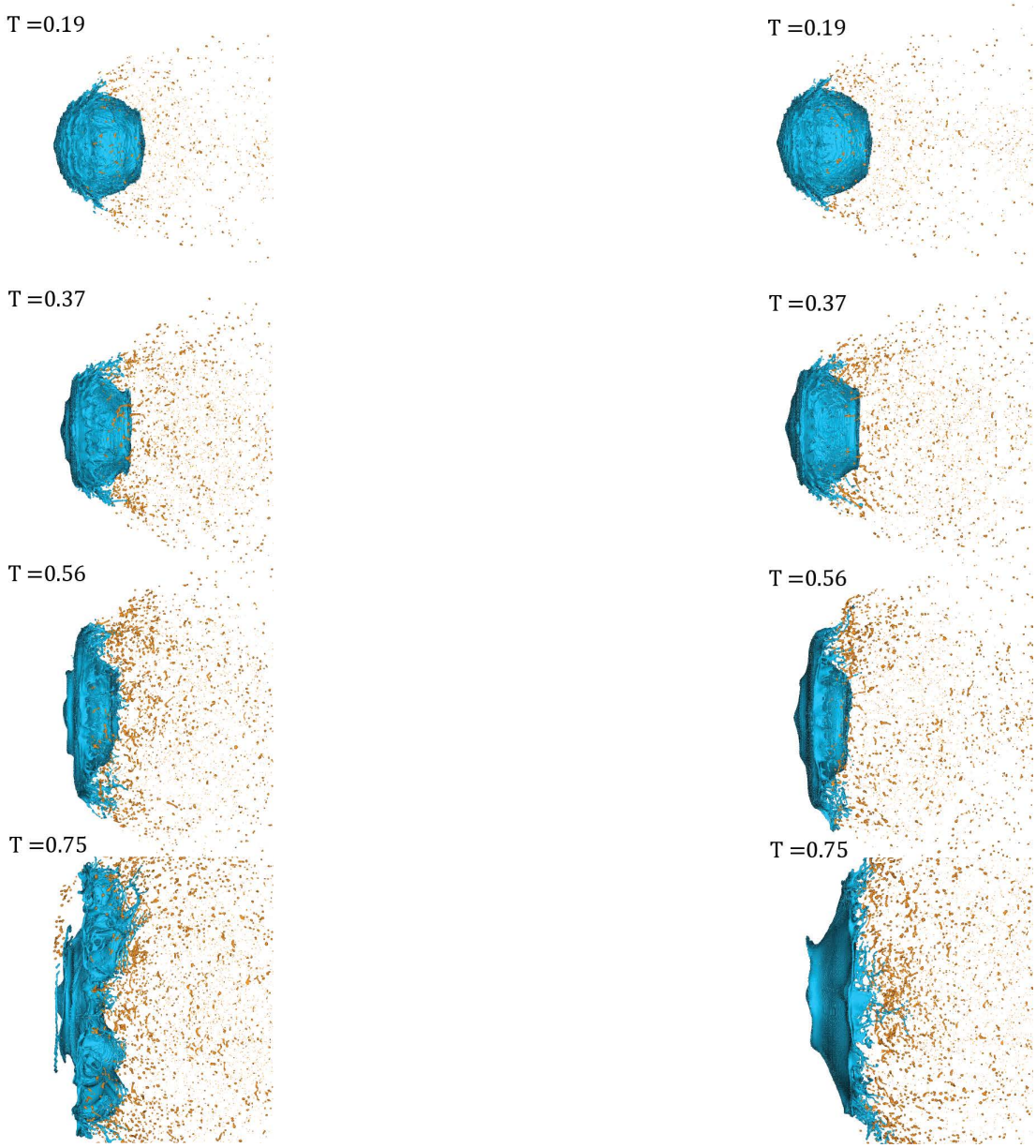
interface at high fidelity. Nevertheless, the large-scale features observed here: disk-like deformation, edge stretching, bag formation, and trailing-edge stripping—are highly relevant for liquid structures under similar aerodynamic conditions, as they dictate the primary pathways through which droplets eventually fragment. The following sections are divided into early-stage breakup and late-stage breakup, with the appearance of bags in the water simulations ($T \approx 0.7$) marking a significant shift in droplet production mechanisms.

Early Instability Formation, Wavelengths, and Stripping

As discussed in §3, the systems considered in this paper operate in the SIE–RTP transition regime, meaning they exhibit features of both Rayleigh–Taylor piercing (RTP) and shear-induced entrainment (SIE). In particular, the SIE contribution appears as a stripping mechanism in the drops' early evolution, driven by high-velocity gas shearing liquid from the windward side. This stage begins with wave instabilities on the drop surface. At this point, some numerical considerations merit discussion. Mapping a sphere onto the Cartesian grid introduces small-scale perturbations near the interface, which can seed shear-driven instabilities. However, the surrounding turbulent gas flow quickly dominates these artificial perturbations, overshadowing them with eddies at various scales.

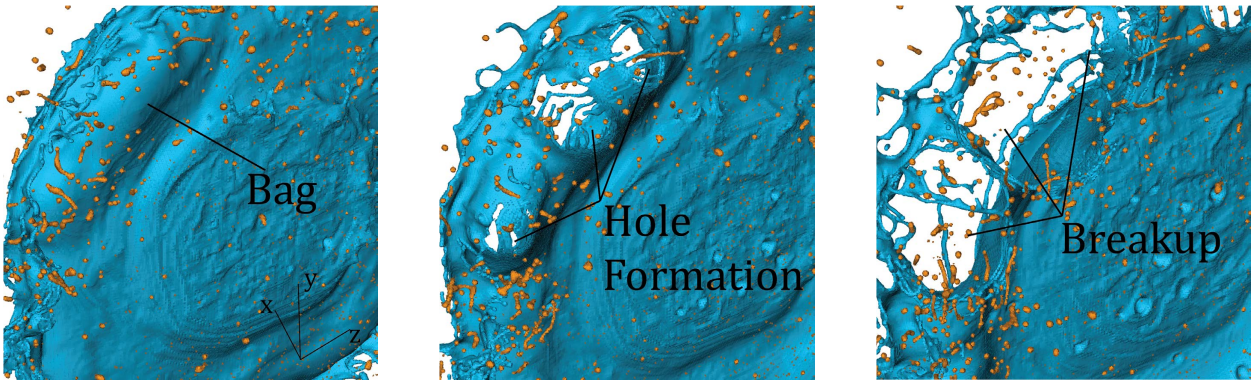
The effect of grid resolution on instability formation is illustrated in Figure 16. Notable differences are observed in the small-scale waves on the windward side of the droplets, while the large-scale instabilities near the droplet equator exhibit qualitative similarities across all cases. In each resolution, two prominent K-H waves emerge, though the smaller-scale waves vary. These results reinforce the confidence that coarser simulations can adequately capture large-scale mechanisms and still provide valuable physical insights.

Shortly after wave initiation (Figure 17a), modulations develop along the wavefront (Figure 17b). These transverse instabilities evolve into lobes (Figure 18), which can be viewed as thin sheets bounded by thicker rims. At this point, the pathways to primary



(a) (b)

Figure 14: Early stages of of droplet aerobreakup, largely dominated by stripping (SIE) regime, from (a) Water (W1 simulation), and (b) Shear-thinning (ST1 simulation).



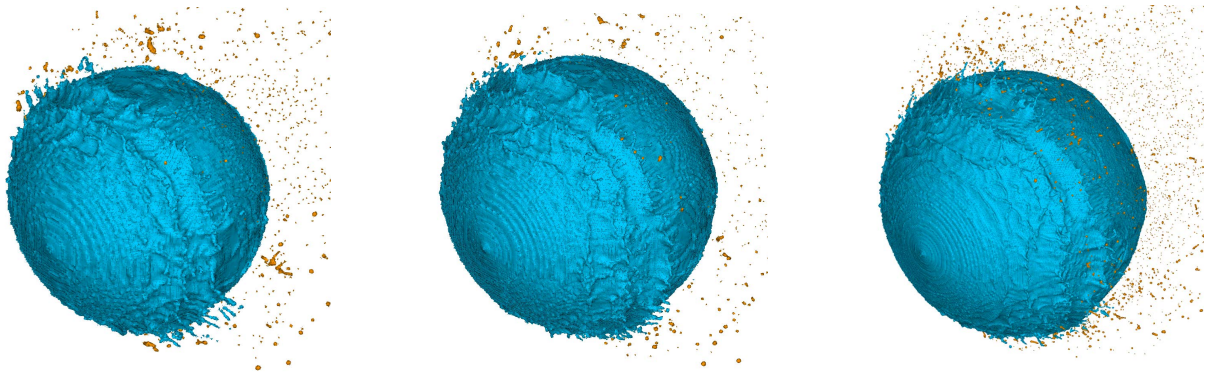
(a) $T = 0.69$ (b) $T = 0.72$ (c) $T = 0.75$

Figure 15: Bag breakup observed in the W1 Simulation.

liquid breakup branch, as summarized in the breakup tree (Figure 19). The blue-shaded boxes represent various routes to primary droplet formation.

In one pathway, lobes become narrower and stretch downstream, forming ligaments that may break away as large broken ligaments or shed smaller droplets along their length due to Plateau-Rayleigh instabilities. Alternatively, some lobes flatten further to form a thin sheet that perforates, creating many small droplets and leaving behind the thicker rim as a bridge. These bridges, in turn, continue to stretch and ultimately break. They may break anywhere along their length, leading to numerous breakup mechanisms: an asymmetric (single-side) break creates a "whip" that detaches while shedding droplets (Figure 20), a simultaneous double-end break leaves a single broken ligament oriented orthogonal to the flow (Figure 21), and a central break yields two ligaments that stretch independently (Figure 22).

Although the routes differ in detail, they all initiate from the same wave-to-lobe process and ultimately contribute to primary droplet formation. Understanding the relative frequency and dynamics of each pathway is essential for accurately predicting how much liquid mass atomizes into primary structures. These stripping modes exemplify the SIE contribution within the SIE–RTP transition regime, illuminating how high gas shear results in the continuous shedding of liquid structures from the droplet's periphery. They



(a) W1 simulation ($dx = 6.67 \times 10^{-3}$)

(b) W2 simulation ($dx = 5.00 \times 10^{-3}$)

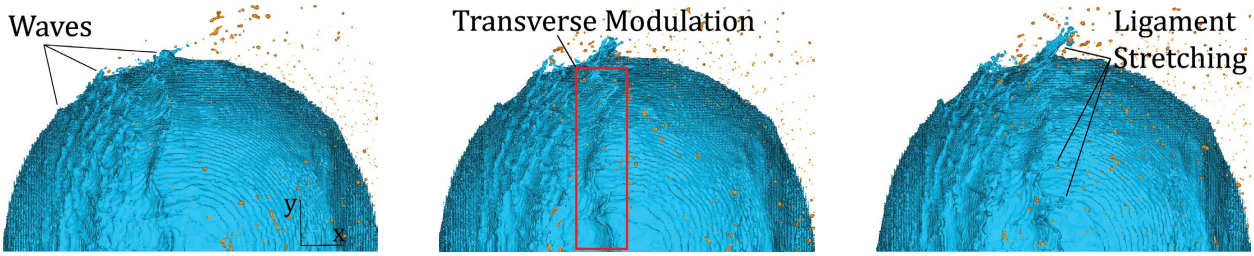
(c) W3 simulation ($dx = 3.33 \times 10^{-3}$)

Figure 16: Instabilities early in water droplet development at three different resolutions at $T = 0.11$.

bear similarities to the mechanisms described by [119] and [221, 222], though additional complexities arise in the highly turbulent, three-dimensional cases examined here.

Shortly after wave initiation (Figure 17a), modulations develop along the wavefront (Figure 17b). These transverse instabilities develop into lobes (Figure 18). These lobes can be viewed as thin sheets bounded by thicker rims. At this point the pathways to primary liquid breakup branch, as summarized in the breakup tree Figure 19, where blue-shaded boxes represent various routes to primary droplet formation.

In one pathway, lobes become narrower and stretch downstream, forming ligaments that may break away as large broken ligaments or shed smaller droplets along their length. Alternatively, some lobes flatten further to form a thin sheet that perforates and leaves behind the thicker rim as a bridge. These bridges in turn will continue to stretch and ultimately break. They may break anywhere along their length, leading to numerous breakup mechanisms: an asymmetric (single-side) break creates a "whip" that detaches while shedding droplets (Figure 20), a central break yields two ligaments that stretch independently (Figure 22), and a simultaneous double-end break leaves a single broken ligament oriented orthogonal to the flow (Figure 21).



(a) $T = 0.08$ (b) $T = 0.10$ (c) $T = 0.11$

Figure 17: Rendering of a water drop (W3) simulation showing initial wave development, transverse modulation instabilities, and ligament stretching from the modulations.

It is important to consider also how the amplitude of the initial instability wave plays a part in the breakup mechanisms and resultant droplets. When the waves first form, they have a relatively small amplitude and the resultant lobes, holes, bridges, and primary droplets formed at this stage appear to scale similarly. It is common that many small holes form along the transverse extent of a wave early in its development, like in Figure 22. These holes undergo the range of pathways described above, generally shedding smaller primary structures, ultimately leaving behind a corrugated wave. As this wave continues growing in amplitude toward the equator, it is then already seeded with many transverse instabilities from these earlier breakup mechanisms, further contributing to the complexity of the production of primary droplets.

Although the routes differ in detail, they all initiate from the same wave-to-lobe process and ultimately contribute to primary droplet formation. Understanding the relative frequency and dynamics of each pathway is essential for accurately predicting how much liquid mass atomizes into primary structures. These stripping modes exemplify the SIE contribution within the SIE-RTP transition regime, illuminating how high gas shear results in the continuous shedding of liquid structures from the droplet's periphery. These mechanisms bear similarities to those described by [119] and [221, 222], though additional complexities are observed in the highly turbulent, three-dimensional cases analyzed here.

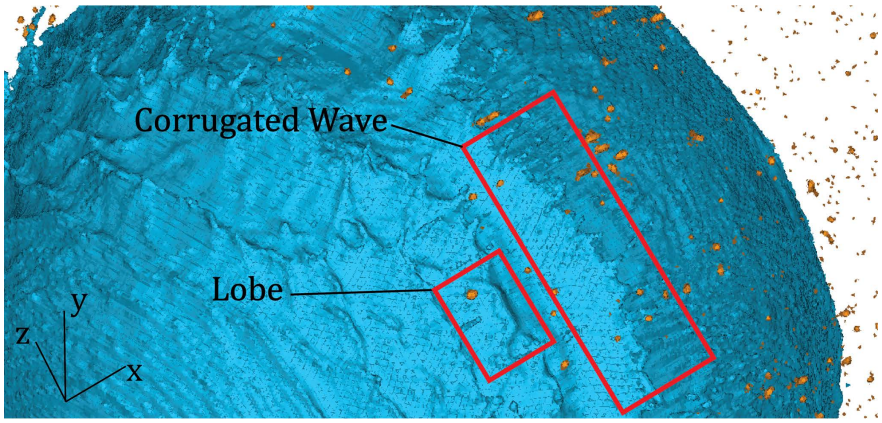


Figure 18: Rendering from water drop (W3) showing the early stages of instability formation and primary breakup ($T = 0.11$).

Instability Wavelength Evolution

A key quantitative indicator of these early breakup dynamics is the wavelength of the initial Kelvin-Helmholtz perturbations on the droplet's windward side. In our simulations, we measure wavelengths (λ_{KH}) of approximately $0.130D_0$ for W1 case and $0.151D_0$ for W3, where D_0 is the initial droplet diameter. These values align with prior high-Weber-number aerobreakup findings, namely, the wavelengths match the criteria laid out by [192] for the SIE regime to dominate breakup ($\lambda \leq 0.2D_0$). The wavelengths were determined using multiple measurements at consistent locations across the droplet interface in both simulations, based on high-resolution renderings. As these waves expand and displace liquid radially, the drop flattens, and the effect of shearing decreases, leading to wavelengths on the order of $0.2D_0$ at about $T = 0.25$ (see Figure 23b). This meets and shortly thereafter exceeds the threshold to enter RTP breakup, manifesting in the rapid flattening and the formation of bags in W1, as described in §3.

[119] similarly emphasize the role of the gas vorticity layer (δ) in setting the primary K-H wavelength with

$$\lambda_{KH} \approx \delta \sqrt{\rho_l/\rho_g}, \tag{13}$$

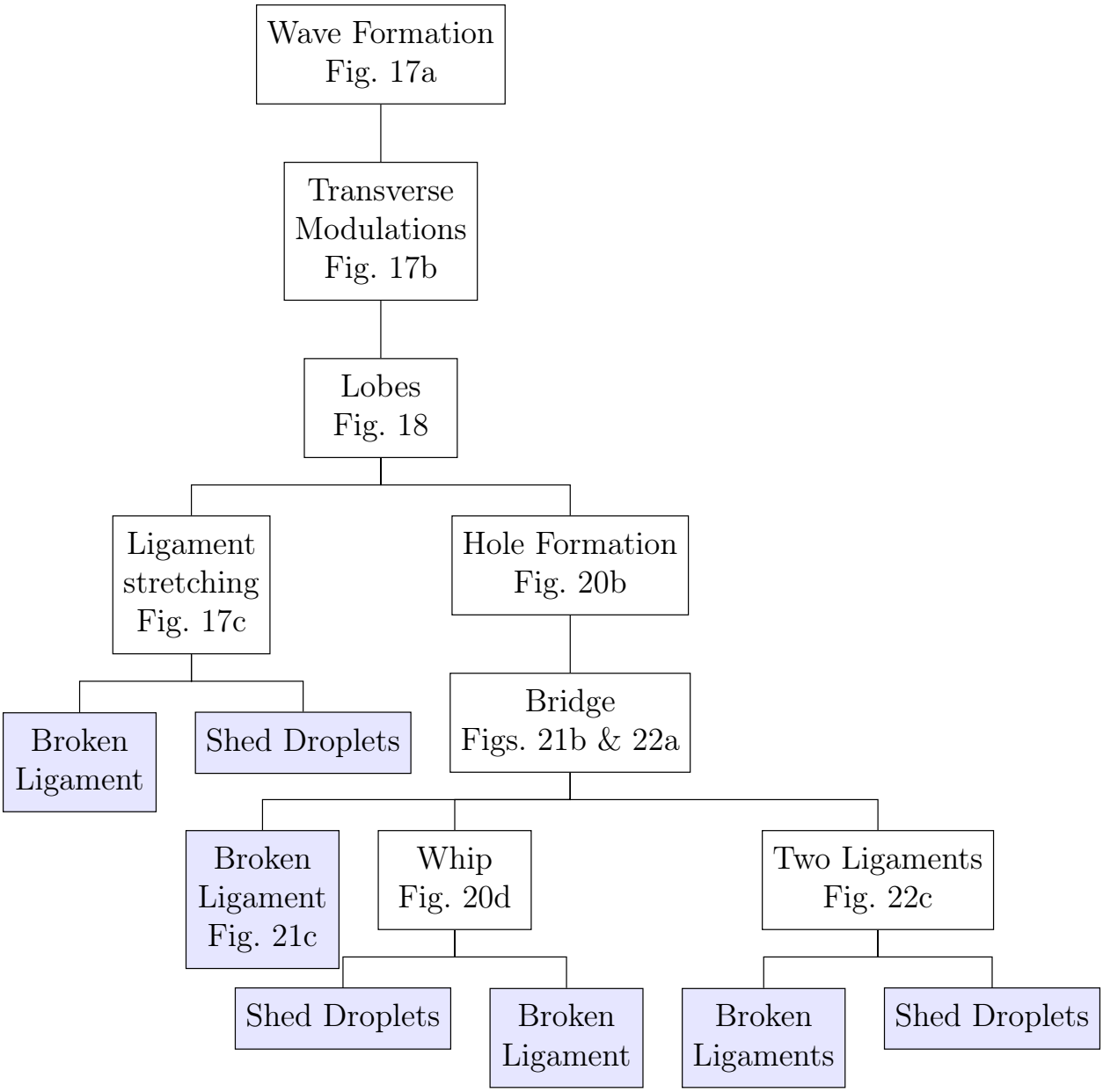
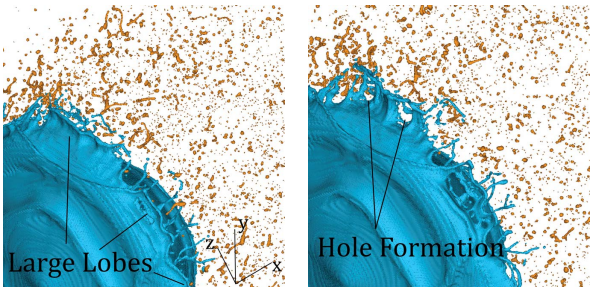
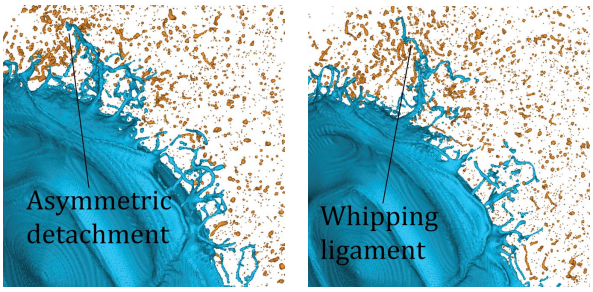


Figure 19: Tree summarizing observed primary droplet production pathways in the stripping regime.

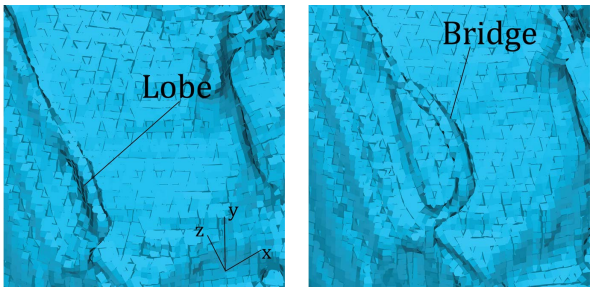


(a) $T = 0.47$ (b) $T = 0.49$

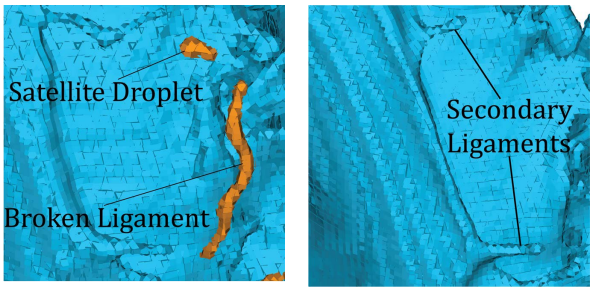


(c) $T = 0.50$ (d) $T = 0.52$

Figure 20: Asymmetric bridge breakup, creating whipping ligament in R2 simulation.

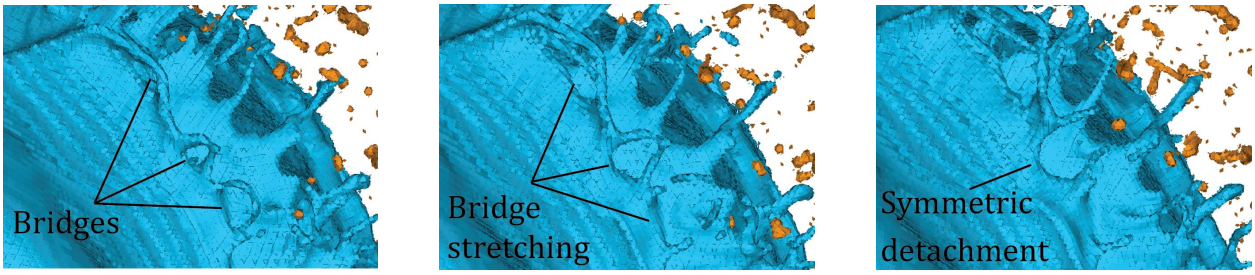


(a) $T = 0.38$ (b) $T = 0.40$



(c) $T = 0.43$ (d) $T = 0.46$

Figure 21: Bridge breakup via double pinch-off points.



(a) $T = 0.44$ (b) $T = 0.45$ (c) $T = 0.46$

Figure 22: Instances of bridge development, stretching, and ultimately detachment from water drop (W2 simulation).

(a) Early-stage K-H waves at $T = 0.15$ ($\lambda_{KH} = 0.13D_0$).

(b) Late-stage RT waves at $T = 0.46$ ($\lambda_{RT} = 0.29D_0$).

Figure 23: Examples of wave evolution with time for a water drop (W1 Simulation).

noting that thinner layers are associated with shorter waves. Their analysis focused on cylindrical jets, so the boundary layer thickness around our spherical droplets will evolve differently about the poles. Nevertheless, the same core principle applies: while the droplet remains relatively spherical, more surface area aligns or nearly aligns with the flow.

To estimate δ at the equator, where the layer is thinnest, we use the relation

$$\delta = \frac{U_{max} - U_{min}}{\mathrm{d}U/\mathrm{d}r|_{max}}. \tag{14}$$

Applying this definition to velocity profiles around the droplet interface we find that δ varies from $\delta \approx 13\ dx$ at a polar angle of 20° to $\delta \approx 3.5\ dx$ at 60°. This value of δ suggests $\lambda_{KH} = 1.2D_0 - 5D_0$ from Eq. 13. However, measurements from water droplet simulations W1 and W3 shows $\lambda_{KH} = [0.130D_0, 0.151D_0]$, respectively. This suggests that the simulation does not adhere to the classical definition of K-H waves as described for a cylindrical jet by [119]. It is likely that the spherical, rapidly deforming shape of the droplet surface leads to non-linear dynamics.

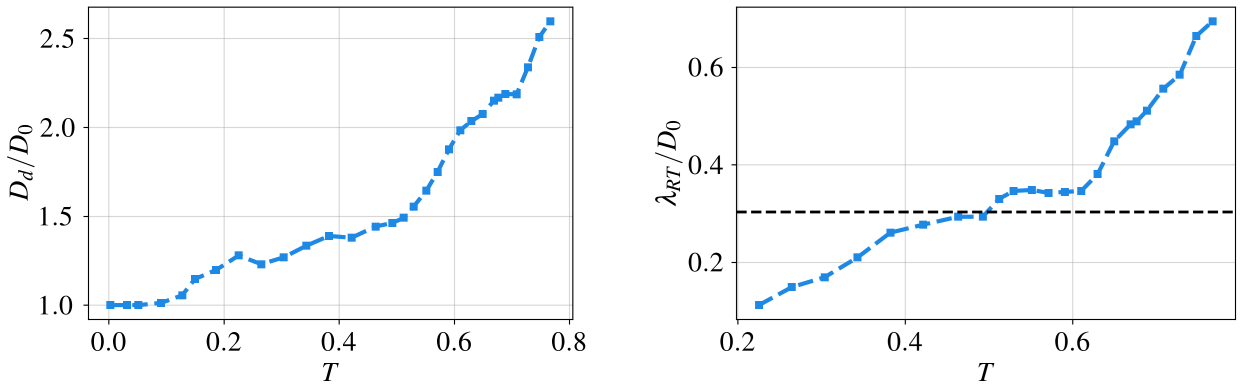
Our findings align with those of [192], who observed that the flow attaches around a polar angle of 30° and detaches near 90°, as evident in Figure 17. The relatively thin boundary layer in this region favors shorter-wavelength shear modes. As the droplet flattens and the flow impinges more perpendicularly, the boundary layer effectively thickens on the windward side of the drop, shifting the instability to longer wavelengths and thus heralding a transition from SIE to RTP-type breakup.

As the droplet continues to flatten under aerodynamic loading, the relative importance of shear-driven Kelvin–Helmholtz (KH) modes diminishes, and Rayleigh–Taylor (RT) modes begin to dominate. The point of this transition can be quantified with estimates of the most unstable RT wavelength. This depends on the acceleration of the gas-stream into the drop. Qualitatively, we observe RT waves as the drop begins to flatten, in agreement with [194]. Approximating the surface to a circular disk and balancing the drop inertia with the drag force yields acceleration

$$a = \frac{3C_D\rho_g U_g^2 D_d^2}{4\rho_l D_0^3}, \tag{15}$$

where C_D is the drag coefficient (1.2 for a round disk) and D_d is the diameter of the flattened drop. It should be noted that the D_0 in the denominator accounts for the initial drop volume, meaning this relies on the assumption that the drop has not lost mass. As can be seen in Figure 13b, the drop retains over 95% of it's mass until $T \approx 0.7$, so for this analysis, this is a reasonable assumption.

As the drop flattens and D_d increases (Figure 24a), the resultant drag force, and therefore acceleration increase. This process is inherently time-dependent and highly sensitive to the drop's evolving shape. Based on a priori knowledge of which waves become the dominant RT wave, we track their wavelength throughout the simulation until bag breakup destroys the waveforms (Figure 24b). An estimate of the most unstable RT



(a) Diameter of the drop as it flattens and expands to a disk. D_d is diameter of the disk.

(b) RT wavelengths measured over time. Black line is estimated unstable wavelength from Eq. 16

Figure 24: Measurements of drop diameter as it flattens and RT Wavelength from water drop (W1 simulation).

wavelength is given by [131] as

$$\frac{\lambda_{RT}}{D_0} = \frac{K}{\sqrt{We_g}}, \tag{16}$$

where K is an empirical parameter dependent on the drag coefficient. For $C_D \approx 1.2$, $K = 8.7$ and with $We_g = 825$, $\lambda_{RT} = 0.303 D_0$. The observed RT waves reach this size at $T \approx 0.46$, seen qualitatively in Figure 23b, where no more KH waves are seen on the upstream side of the drop. This time also corresponds to a sharp increase in diameter in Figure 24a, indicating that deformation is strongly occurring at this point. From this analysis, it is evident that after the initial growth of KH waves, RT instability quickly dominates droplet deformation, piercing the drop and driving the flow into the nonlinear sheet thinning and bag-breakup regime.

Late-stage Primary Breakup Mechanisms

The early-stage mechanisms occur in both water and non-Newtonian fluids, with few qualitative differences. However, as the initial drops begin to flatten, the primary breakup mechanisms diverge between the two cases. In both scenarios, the equator of the droplet

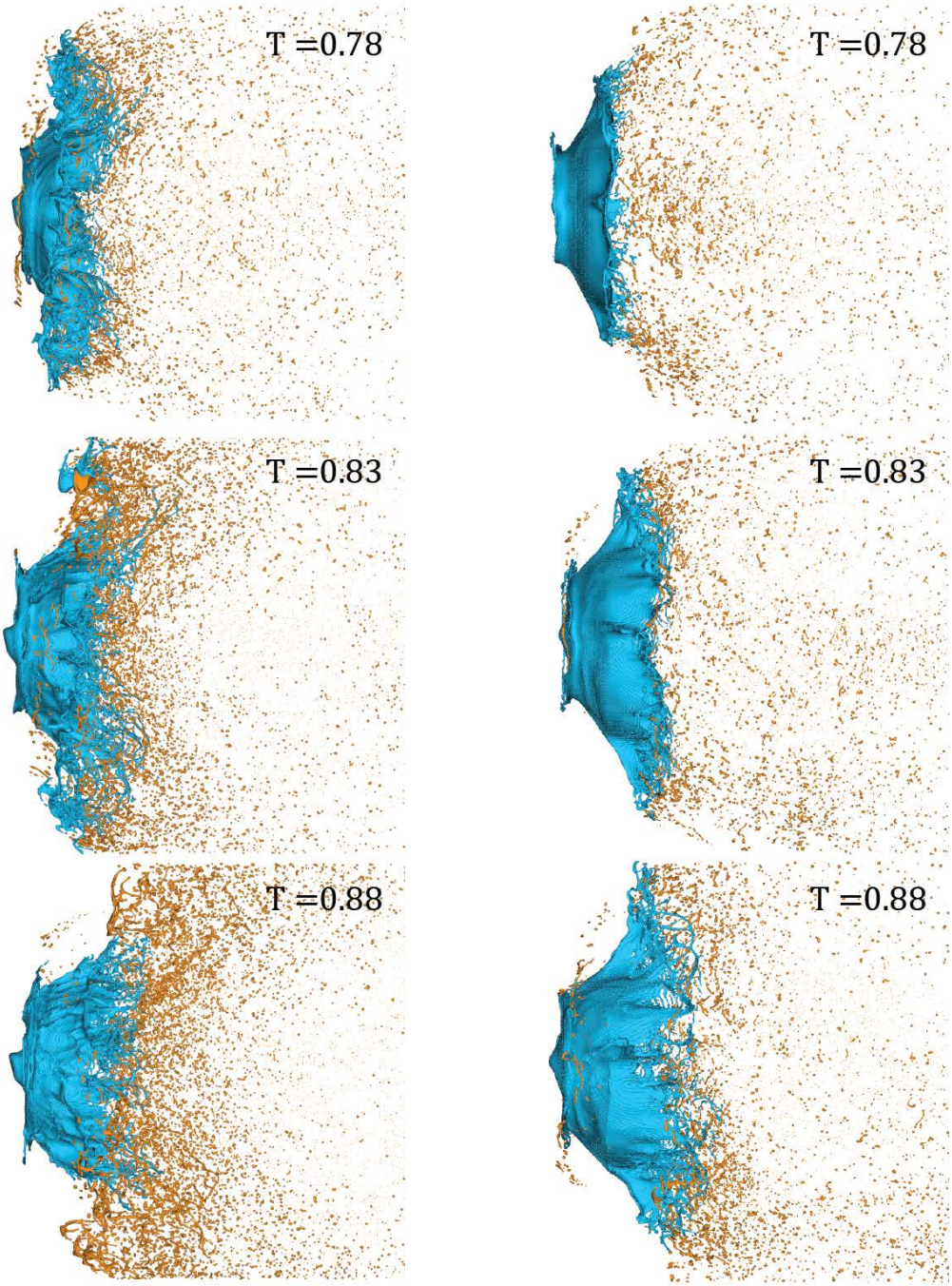
expands and elongates in the flow direction, leading to a flattened disk-like shape.

For water droplets, the flattening is followed by the formation of bags at the edges of the structure (See Figure 15). These bags inflate and eventually burst in the direction of the flow, consistent with the mechanisms observed in the SIE-RTP transition regime described in §3 and in [39]. In contrast, non-Newtonian droplets do not exhibit bag breakup. Rather, the continued flattening causes the outer edges of the disk to fold in the direction of the flow. Stripping is then continued along the sheets as they are further stretched in the direction of flow. §3 will propose a mechanism to explain the differences in behavior. Figure 25 illustrates the late-stage evolution of the breakup process. While the early stages of the breakup display many qualitative similarities, the dynamics diverge significantly as the coherent drop shape undergoes more extreme deformation.

Secondary Atomization

Qualitative analysis of secondary atomization in large-scale simulations presents challenges similar to those encountered in experimental studies. The high Weber number regime produces a dense spray cloud with a vast number of droplets to track and analyze (see Fig. 26), making manual analysis impractical without automated methods. Additionally, rendering files for each timestep range from 5 to 150 GB, depending on resolution, making large-scale automated parsing computationally prohibitive across hundreds or thousands of timesteps. These limitations restrict detailed analysis of secondary breakup mechanisms, which will be addressed in §3.

Despite these challenges, we qualitatively observe Plateau-Rayleigh breakup in ligaments that are shielded from the high-speed gas flow by larger liquid structures (see Fig. 27). These elongated ligaments are inherently unstable and undergo further disintegration due to a combination of aerodynamic forces and surface tension. Ligaments exposed to high shear flow break up aerodynamically, while those entrained within slower-moving regions primarily



(a) (b)

Figure 25: Late stage breakup evolution of 25a W1 and 25b R1

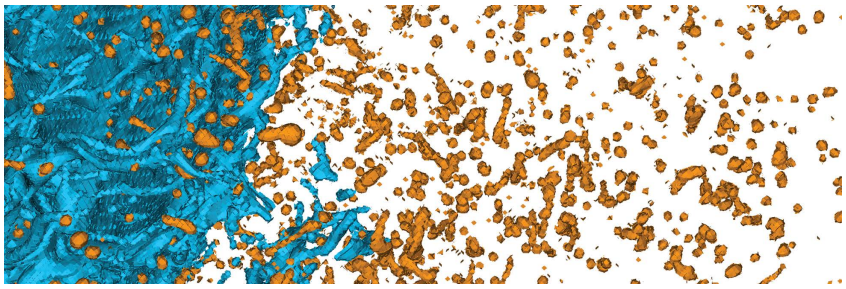
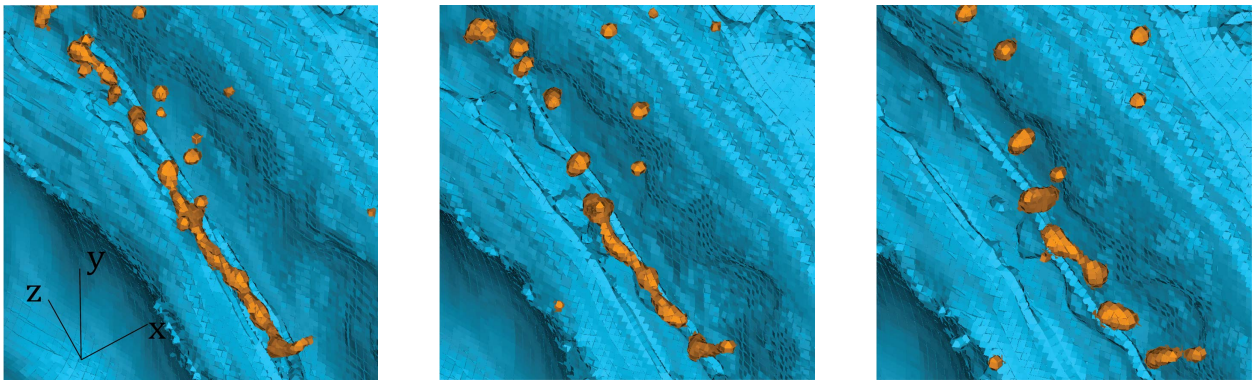


Figure 26: Dense cloud of droplets (orange) breaking off of the initial liquid drop (blue) from the WL simulation.



(a) $T = 0.141$ (b) $T = 0.145$ (c) $T = 0.148$

Figure 27: A ligament from the (orange) undergoes Plateau-Rayleigh breakup in the WL simulation.

fragment via Plateau-Rayleigh instability. This secondary atomization process plays a key role in shaping the final droplet size distribution and generating fine droplets.

Breakup Dynamics from ASSET

This section delves into the quantitative analysis of the droplet breakup process using data extracted by ASSET. By examining various droplet characteristics and breakup events, we aim to provide a statistically-driven perspective on the atomization cascade, complementing the observations discussed previously.

Droplet Genealogy Analysis

The assignment of identification numbers in ASSET enables the tracking of child-parent relationships for every droplet produced in the simulation, linking each back to the initial coherent drop. This is achieved using the Neo4j graph database management system, which facilitates a novel analysis of the breakup cascade—from the initial stages of primary breakup to the eventual form droplets take within the spray.

We define "final" droplets as those with no outgoing relationships in the graph database, meaning they are present at the end of the simulation and have not undergone further breakup. While this definition does not imply that these droplets would not break up further if the simulation were extended, it is a practical and consistent way to assess the state of droplets given the finite runtime of the simulations. Importantly, during the simulation, droplets had sufficient opportunity to continue breaking up if conditions favored it, and many did not. This justifies treating these droplets as representative of stable, terminal structures under the simulated conditions.

This approach allows for a detailed understanding of the breakup cascade. For example, it highlights that while many primary droplets remain intact and do not undergo further breakup, other structures continue fragmenting into smaller droplets. As shown in Figure 28, a significant number of "final" droplets were products of numerous breakup events, highlighting the dynamic and iterative nature of secondary atomization. While secondary atomization is important, a significant proportion of droplets do not undergo further breakup, these are largely small, nearly spherical droplets created via stripping or shedding from attached ligaments (Figure 17c). Droplets that undergo significant numbers of breakup events are likely large structures, which exhibit numerous smaller breakup events from stripping or Plateau-Rayleigh breakup (e.g. the ligament in Figure 27 which exhibits 5+ events in the sequence of images). The distribution of droplet counts across breakup stages exhibits a trend consistent with exponential decay.

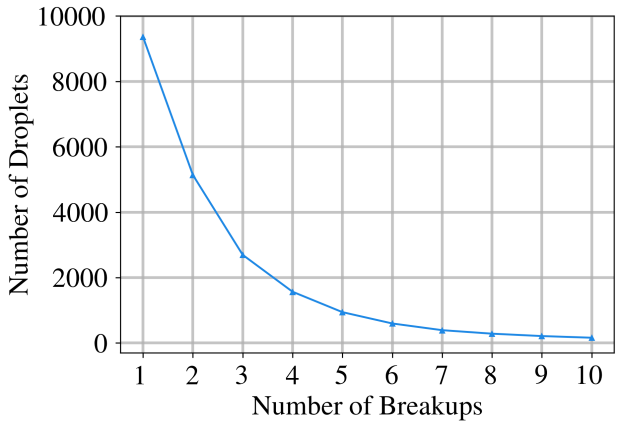


Figure 28: Breakup stage of final droplets in the W1 simulation.

In the following sections, we analyze the cascade process of breakup using the droplet genealogy discussed here. Many analyses focus on the breakup stage, which is the number of fragmentation events a droplet has undergone to reach its observed state. Using the relationship data within the Neo4j graph, we classify droplets based on their breakup histories, providing insight into the progression of the breakup cascade. This analysis highlights trends in how droplets evolve through repeated fragmentation and helps identify the conditions under which droplets stabilize or continue to break apart. Understanding these dynamics is key to characterizing the mechanisms behind both primary and secondary atomization.

Primary Breakup Statistics

Figure 29 plots the time evolution of the average size of primary droplets for the water (W1) and shear-thinning (ST1) simulations, revealing an overall increase over time. A rolling average is computed by averaging droplet sizes over a fixed window to smooth out short-term fluctuations and observe the underlying trends. Periodic spikes mark significant breakup events with larger-than-average droplets being produced. Notably, both curves exhibit pronounced spikes around that appear at similar times until after $T \approx 0.6$, where the dynamics diverge due to the presence of bag breakups in the water case, which do not occur

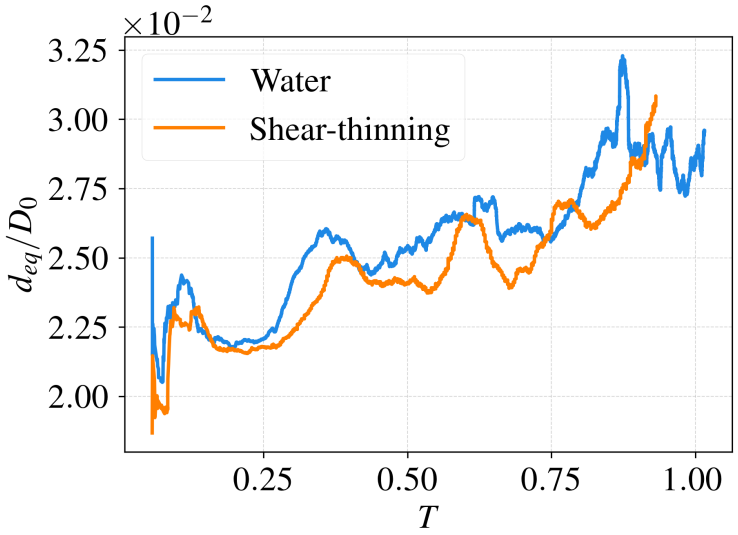


Figure 29: Average primary droplet size as a function of T.

in the shear-thinning case. These correspond closely to moments when the interfacial waves grow to their largest amplitudes at the equator and detach, forming larger-than-average droplets. The first spike ($T \approx 0.1$) corresponds to the K-H waves reaching the equator and stripping in quick succession. The next two spikes ($T = 0.35, 0.6$) are separated by $\approx \Delta T = 0.25$ correspond to RT waves reaching the equator and shedding.

The next spike in the water case corresponds to bags bursting at $T \approx 0.8$, reaching its peak near $T = 0.87$. During this interval, the rims bounding the thin bags disintegrate, releasing a significant number of relatively large fragments into the flow. The shear-thinning drop displays two subsequent spikes at $T \approx 0.75, 0.90$, which occur only $\Delta T \approx 0.15$ after the previous, likely indicating an acceleration of the RT waves as the drop flattens.

Secondary Breakup Cascade Analysis

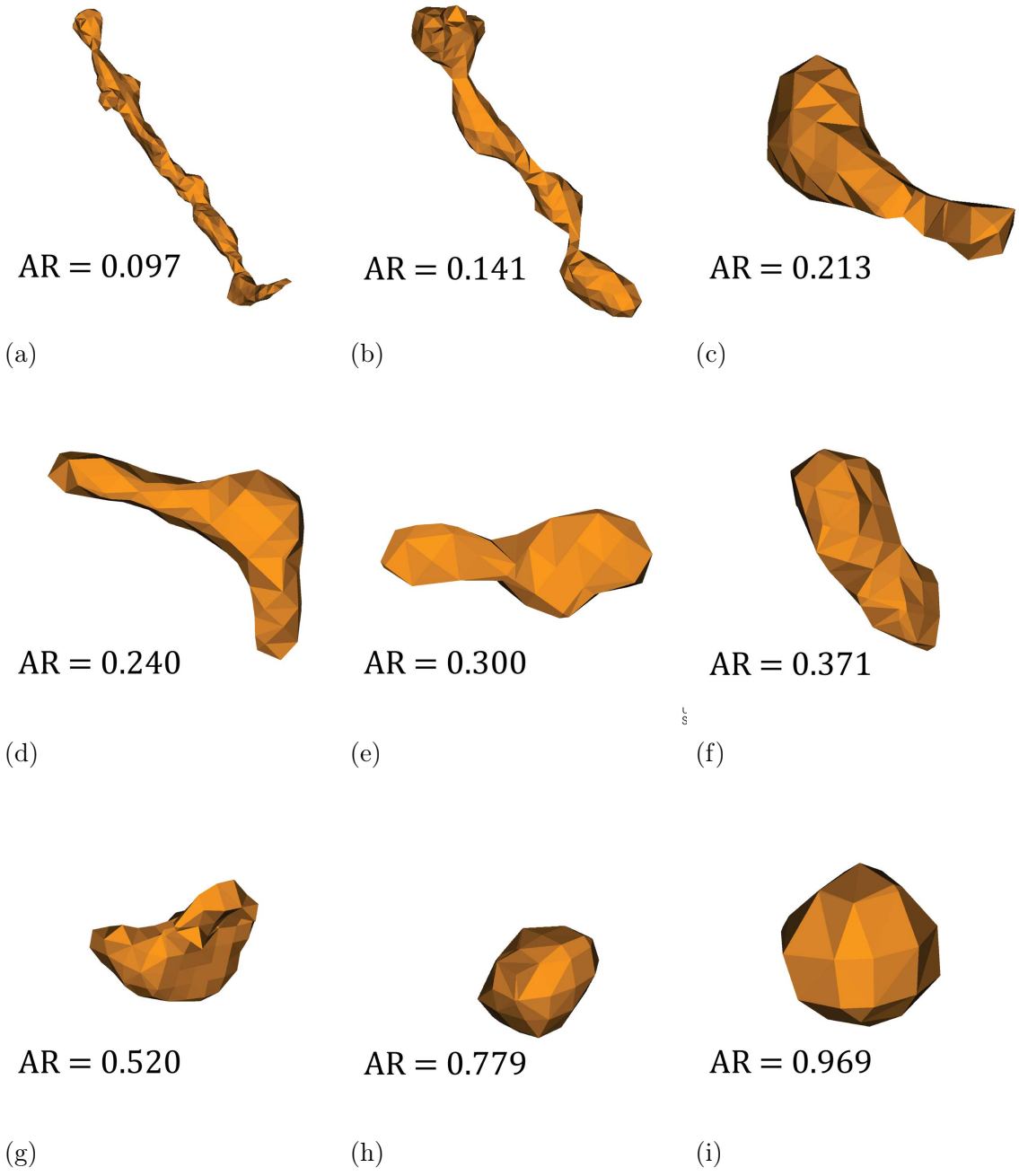
One of the key parameters influencing droplet behavior and breakup is its shape. To quantify droplet shape, we use the aspect ratio (AR), defined as the ratio of the minor axis to the major axis of the best-fit ellipsoid for each droplet. The aspect ratio ranges between 0 and 1, with a value of 1 corresponding to a perfectly spherical droplet and lower values

indicating increasing elongation. Figure 30 provides examples of droplets at various stages of the breakup process, highlighting their diverse shapes and corresponding aspect ratios. These examples demonstrate how the aspect ratio effectively characterizes different droplet morphologies. Additionally, the threshold for identifying a ligament, defined as AR = 0.3, is illustrated, along with examples of shapes slightly above and below this value to show the variations in aspect ratio. This threshold was selected based on qualitative observations, but further refinement may be necessary to establish a standardized value for ligaments.

The aspect ratio (AR) serves as a key metric for analyzing the atomization cascade. Droplets with low ARs are inherently unstable and more prone to further breakup, while higher AR droplets tend to be more spherical and stable. Using Neo4j, we can classify droplets that continue to break up versus those that remain intact by the end of the simulation. Figure 31a highlights this distinction: droplets with low ARs are highly likely to undergo additional breakup, aligning with the ligament formation mechanisms discussed in §3. In contrast, droplets with higher ARs (more spherical shapes) are much less likely to break up further. This trend is consistent across both water and non-Newtonian simulations, although the non-Newtonian case exhibits a generally lower probability of breakup.

While analyzing the number of droplets that break up is useful, it does not fully capture the underlying breakup mechanisms. A more insightful approach is to assess how much of the atomized volume can be attributed to ligament shedding and subsequent secondary atomization. By isolating simulation datasets to focus on primary structures (droplets breaking directly off the initial coherent liquid) and volume-weighting the AR distribution, we observe in Figure 31b that a significant portion of the atomized volume is concentrated in low-AR structures, or ligaments. This highlights the dominance of ligament formation in the breakup process.

To shed light on the cascade process, we next examine the characteristic sizes of both ligaments and spherical droplets. Some results below are displayed as density functions with

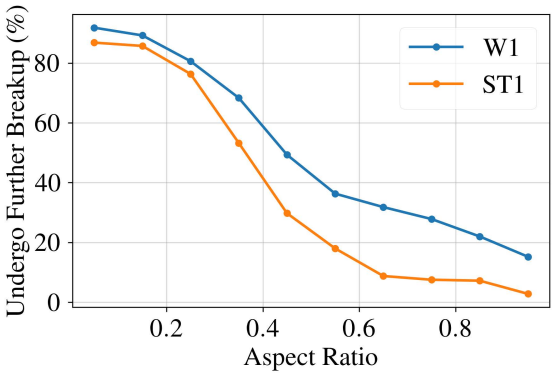


(a) (b) (c)

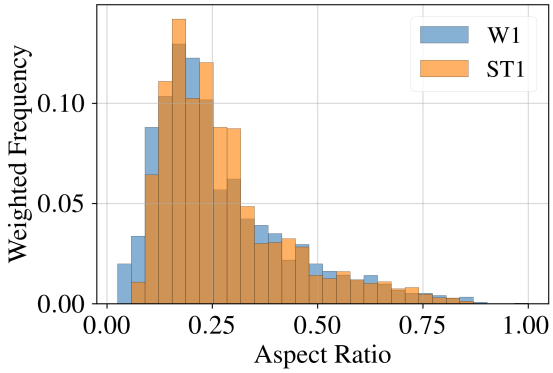
(d) (e) (f)

(g) (h) (i)

Figure 30: Aspect Ratio (AR) and renderings of various droplets obtained from the water drop simulation.



(a) Percentage of droplets undergo further breakup as a function of aspect ratio.



(b) Volume-weighted distribution of primary droplet aspect ratios for water (W1) and shear thinning (ST1) simulations.

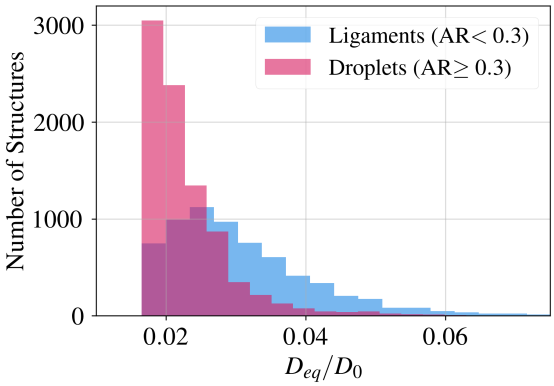
Figure 31: Statistics describing the effect of aspect ratio on secondary breakup mechanisms of W1 and ST1 simulations.

a gamma distribution

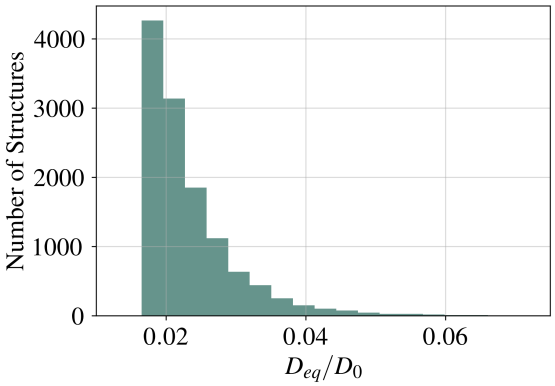
$$f(x, \alpha, \beta) = \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, \tag{17}$$

where α and β are parameters estimated from the data and Γ is the gamma function. Gamma distributions have been shown to fit well with drop size distributions [128, 205, 206] in aerodynamic shearing systems.

Larger broken ligaments often undergo multiple breakups, typically through a combination of stripping and Plateau-Rayleigh instability, progressively fragmenting into numerous smaller, more spherical droplets. As shown in Figure 32a, ligaments (AR < 0.3) consistently appear at larger volumes than spherical droplets, which aligns with the expectation that smaller droplets are dominated by surface tension, while larger droplets retain elongated shapes due to inertial effects. Notably, in Figure 32b, the final droplet size distribution closely matches that of the primary spherical droplets and lacks a long tail of larger droplets. This finding indicates that ligaments have fragmented into smaller droplets, leading to a distribution that closely matches the primary spherical droplet distribution. This displays



(a) PDF of primary structures from the W1 simulation. Gamma functions fit with parameters $\alpha = [1.74, 1.30]$ and $\beta = [0.01, 0.01]$ for ligaments and droplets, respectively.



(b) Size distribution of final structures from the W1 simulation. Gamma functions fit with parameters $\alpha = 1.32$ and $\beta = 0.01$.

Figure 32: Primary and final droplet diameter distributions from the water drop (W1) simulation.

the role of ligament breakup in shaping the final, stable population of droplets.

Turning toward secondary atomization, it is important to consider if we have a primary droplet that we predict to undergo secondary breakup, *what is the expected outcome of that breakup?* Figure 33 shows the size of a secondary (in this case any breakup stage after "primary") droplet as a function of the original size of the primary droplet it came from. The data are constrained to $T < 0.25$ so that the primary breakup mechanisms are constrained to stripping in the SIE regime. We observe that the three different resolutions yielded very similar distributions: W1 displays a peak at approximately $0.42\, d_{eq,child}/d_{eq,parent}$, with W2 and W3 showing nice agreement at $0.35 d_{eq,child}/d_{eq,parent}$. Here, sizes are presented as equivalent spherical diameters where $d_{eq,child}$ is that of the new droplet, and $d_{eq,parent}$ is that droplet's primary parent diameter.

These results have significant implications for modeling liquid breakup and atomization processes. Using ASSET to track the evolution of liquid structures in this simulation,

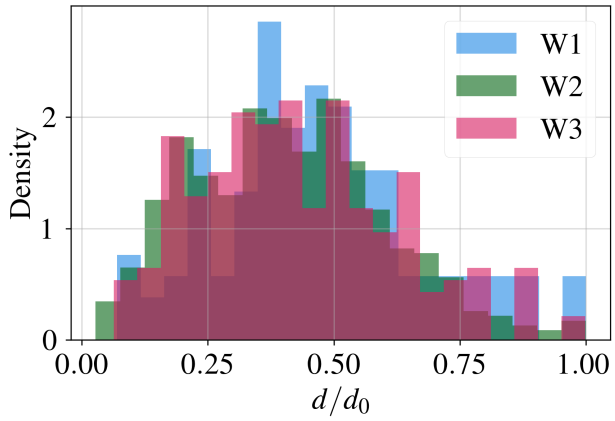


Figure 33: Probability density function of final water droplet diameters (d_{eq}) as a function of their original primary droplet parent diameter (d_0). Gamma functions fit with parameters $\alpha = [15.86, 10.27, 6.62]$ and $\beta = [0.05, 0.06, 0.08]$ for resolutions W1, W2, and W3, respectively.

we extract model-relevant data that informs predictive frameworks for droplet behavior. Specifically, if a primary droplet is identified, we can assess whether it is likely to break up and, if so, predict the expected size distribution of its resulting droplets.

Effect of Shear-Thinning Viscosity

ASSET is able to extract the effective viscosity of droplets at the moment of breakup. The value is calculated as a weighted average from all cells that contain the droplet that just broke up. This provides an analysis of the effective viscosity distribution associated with each breakup event from the simulations. As seen in Figure 14, there is a notable difference in the deformation of the two droplets. This is likely attributed to higher viscosity values in the non-Newtonian simulation. Using the power-law model for the Phos-Chek fire retardant shear thinning property results in higher viscosity than that of water, which matches experimental measurements from [7]. This can be seen in Figure 34, which shows the effective viscosity measured at every breakup event from the ST1 and ST2 simulations compared to that of water. Additionally, it can be seen that there is minimal effect of the grid on this calculation as ST1 and ST2 show similar distributions. This larger effective

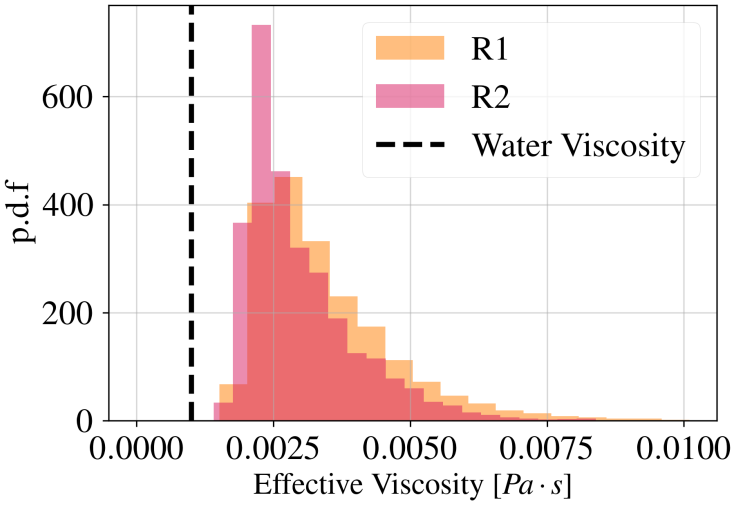


Figure 34: Effective viscosity of breakup events from the shear-thinning simulations. Water viscosity plotted for reference.

viscosity value helps explain the differences in large-scale breakup mechanisms observed in Figure 14. The higher viscosity prevents the rapid thinning of sheets as seen in the water simulation, rather leading to a smoother interface.

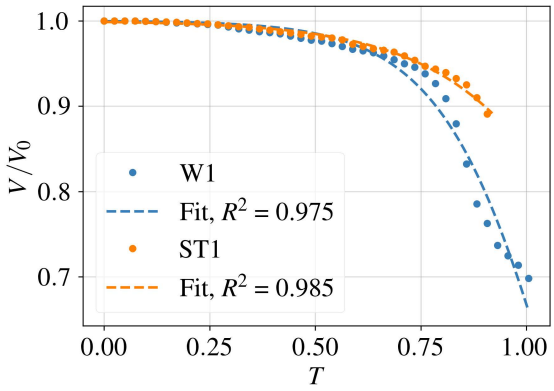
The lack of large-scale bag-breakup mechanisms leads to both fewer total droplets and a smaller average droplet diameter for primary structures developed from the non-Newtonian case (see Table 5).

Despite these differences in large-scale deformation, ligament formation and breakup remain dominant mechanisms in the shear-thinning liquid, as seen in Figure 31b. The aspect ratio distributions indicate that primary structures from both water and shear-thinning liquids predominantly exhibit low aspect ratios, signifying elongated shapes. This suggests that the breakup processes are influenced more by aerodynamic forces and ligament instability than by the fluid's rheological properties at smaller scales.

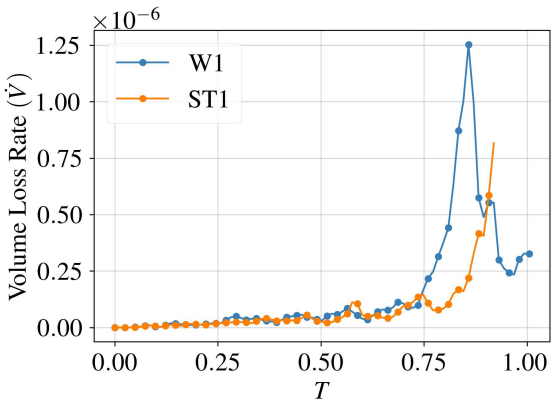
The effect of viscosity on breakup mechanics can be seen clearly in Figure 35a, where the disintegration of the initial drop is delayed in the shear-thinning simulation compared to that of the water. At $T \approx 0.9$, the shear-thinning drop has lost $\approx 10\%$ of its volume,

	Water Primary Structures		Shear-thinning Primary Structures	
	Ligaments (AR < 0.3)	Droplets (AR ≥ 0.3)	Ligaments (AR < 0.3)	Droplets (AR ≥ 0.3)
Total	5374	6995	3577	5112
% of atomized Volume	59.8	40.2	65.0	35.0
Number that break up further	1975	676	1201	247
$\langle d_{primary} \rangle$	$0.039D_0$	$0.030D_0$	$0.034D_0$	$0.025D_0$

Table 5: Analysis of droplet production from primary ligaments and primary shed droplets in W1 and R1 simulations. $\langle d_{primary} \rangle$ is the mean diameter as a function of original drop diameter (D_0).



(a) Volume of the initial water and shear-thinning drops as a function of T. Function approximated from Eq. 18.



(b) Rate of volume loss as a function of time.

Figure 35: Bulk atomization rate for water and shear-thinning drops.

whereas the water drop has lost $\approx 25\%$. The temporal evolution of the initial drop, $V(T)$, can be reasonably approximated using a modified sigmoid function,

$$\frac{V(T)}{V_0} = 1 - \frac{1}{1 + e^{-k(x-a)}}, \tag{18}$$

Where a describes the midpoint where the volume drops and k describes the steepness of the drop. These parameters were fitted to the data, with $a = [1.10, 1.35]$ and $k = [7.0, 4.9]$ for the water and shear-thinning drops, respectively. We observe an acceleration of the volume (mass) loss with time, this is plotted in Figure 35b, where the relative delay in atomization is clearly visible between the water and shear-thinning cases. Furthermore, the bag breakup region appears as a temporary spike in the water drop, after that period of rapid disintegration, the rate of volume loss drops drastically. While this provides a reasonable approximation for the regimes considered here, additional temporal data are required to capture the complete breakup process of the coherent liquid drop. Furthermore, a broader parameter sweep is needed to determine the effect of viscosity on the tuning parameters.

In Figure 25 this delay can be seen visually as the shear-thinning drop exhibits a smoothing of the interface, which holds the coherent liquid together longer, leading to long attached sheets, rather than breaking bags. This is an interesting phenomenon because the higher effective viscosity observed in Figure 34 would suggest higher Ohnesorge numbers, which would place this system more firmly in the RTP regime according to SIE-RTP maps [126, 192]. We see the same larger RTP waves ($\lambda > 0.2D_0$) appear on the shear-thinning drop, however, rather than producing bags, the radially thinning liquid stretches downstream to form sheets in the direction of flow. The shear thinning viscosity is likely responsible for this mechanism. For bag breakup to occur, part of the droplet surface must inflate into a thin film that ultimately ruptures. In the present case, as noted in §3 the drop reaches a certain flatness, which in turn thickens the boundary layer, reducing localized

shearing. Thus, the shear-thinning fluid exhibits a relatively high effective viscosity in the region that would otherwise inflate. The outer edges of the drop, which have the highest shear begin to stretch or fold along the flow direction, resulting in sheet-like structures oriented in the direction of flow.

This behavior suggests that while the large-scale breakup mechanisms are significantly altered by shear-thinning viscosity, the overall stability and evolution of the droplet are still influenced by classical aerodynamic forces. However, to fully understand the implications of shear-thinning viscosity on regime classification, a more comprehensive parameter sweep is required. Additionally, the shear-thinning liquids in this study were modeled using a simple power-law relation (Eq. 11), with parameters chosen to represent Phos-Chek XA, a widely used commercial fire retardant. Under typical flow conditions in our simulations, this formulation yields a higher-than-water viscosity, which is confirmed in experimental studies of the material [7]. However, the power-law model can become non-physical at very low shear rates, where it tends toward infinite viscosity. Many shear-thinning fluids have plateau regimes at both low- and high-shear limits–characteristics better described by more sophisticated rheological models such as the Carreau-Yasuda model [12]. This model can capture the transition from a Newtonian plateau at low shear, through a shear-thinning region, to another plateau at high shear. Adopting such models in future work may yield improved accuracy in fully transient droplet breakup simulations, particularly under the strongly varying shear conditions seen in high-Weber-number aerobreakup.

Some Comments on Grid Effects and the Practical Application of DNS

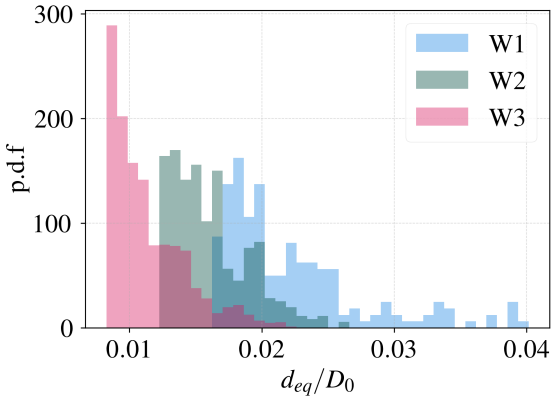
The results presented here highlight the potential of direct numerical simulation (DNS) for studying droplet aerobreakup at large density ratios and high Weber numbers. The uniform, Cartesian meshes used in these simulations exceed cells across droplet diameter (CAD) values used in previous studies [39, 124], and the density ratio ($\rho_l/\rho_g = 816$,

Reynolds ($Re_g = 47400$), and Weber ($We = 825$) match those of many practical applications. The techniques employed represent the state-of-the-art in geometric transport and interface capturing [138, 139]. In short, this methodology represents one of the most computationally advanced attempts to predict a practical application of atomization, and specifically droplet aerobreakup.

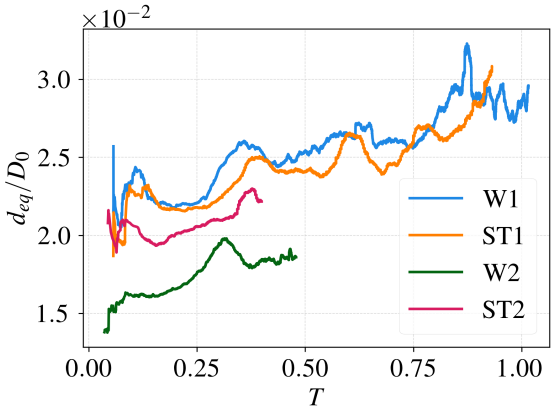
Despite these advances in computational capability and numerical methods, grid resolution remains a critical challenge [111, 124]. Although the results in §3 demonstrate qualitative agreement, certain small-scale features revealed by ASSET highlight the persistent difficulties of capturing these flows numerically. Nonetheless, several key findings remain robust, indicating promise for the approach.

A central motivation of many atomization studies is to extract droplet-size distributions for engineering and environmental modeling. Through ASSET, we obtain these distributions at every stage of the breakup cascade, offering unprecedented detail on the small-scale physics of atomization. However, the fidelity of such data is inherently limited by the underlying simulation. This work finds that the grid resolution can still influence droplet-size predictions. Figure 36a compares primary droplet-size distributions for simulations W1, W2, and W3, each exhibiting a distinct peak. The primary breakup mechanisms—ligament stripping in the shear-induced entrainment (SIE) regime and bag rupture during Rayleigh-Taylor piercing (RTP)—both involve extremely thin ligaments, films, and rapidly expanding "bags", all of which occur well below even the finest computational grids. Accordingly, accurately predicting these fine-scale processes remains a major challenge for the field. Recent progress [54] points toward improved accuracy for sub-grid scale predictions, but, for the present, drop-size distributions—particularly at the smaller end—may be underresolved, potentially skewing the resulting statistics.

Despite the differences in drop-size magnitudes, the simulations capture the dominant physical processes that govern breakup. Figure 36b shows the temporal evolution of the



(a) PDFs of water droplet size distributions for three different resolution simulations.



(b) Moving average of the size of primary droplets produced as a function of time for water and shear-thinning.

Figure 36: Primary drop size plots, showing overall distribution, and time evolution.

mean size of primary droplets for W1, W2, R1, and R2. Although the absolute values vary across resolutions, each simulation exhibits spikes at similar times. These spikes correspond to wave crests reaching maximum amplitude and detaching as larger structures (see §3), indicating that while the exact timing of ligament pinch-off or sheet rupture may differ with resolution, the overarching breakup mechanisms remain consistent. Qualitative parallels between the coarsest simulation (W1) and experimental images (see §3) further affirm this.

Another consideration is the resolution effect on secondary atomization. In Figure 33, which isolates droplets undergoing secondary breakup via SIE (i.e., $T < 0.25$), the size distributions of final droplets are nearly identical for W1, W2, and W3. This consistency suggests that while the absolute droplet sizes can vary, the cascade process and the relative size of daughter droplets remain unaffected.

ASSET offers unprecedented access to the small-scale regions of the flow, unveiling data once buried in tens to hundreds of terabytes of simulation output. Yet, by putting a magnifying glass on these fine scales, it also exposes areas that remain underresolved by current DNS capabilities. While the approach itself is robust, certain types of model-

relevant information remain limited by the inherent constraints of the numerical framework. Nevertheless, as simulations become more accurate and computational power continues to grow, we can extract data that would be otherwise inaccessible to experiments alone. This progress highlights the promise of high-fidelity simulations to drive further advances in understanding atomization.

Conclusions

In this study, we conducted direct numerical simulations (DNS) of droplet aerobreakup at high density ratios and high Weber numbers, capturing both Newtonian (water) and non-Newtonian (shear-thinning) fluids. By integrating the Atomization Simulation Statistics Extraction Tool (ASSET) into these large-scale simulations, we extracted a wealth of quantitative data describing both the primary and secondary breakup processes in unprecedented detail. The main conclusions are summarized as follows:

1. **Comprehensive description of SIE-based primary breakup and transition to RTP waves:** At the baseline Weber number of 825, the droplet breakup exhibited a combination of shear-induced entrainment (SIE) and Rayleigh–Taylor piercing (RTP) mechanisms. Early in the droplet deformation, small-scale Kelvin–Helmholtz-type instabilities and stripping (SIE) processes dominated; however, as the droplet flattened and the boundary layer on the windward side thickened, larger-scale waves formed. These waves ultimately triggered related phenomena, including bag formation and bursting in the water droplet.

2. **Novel statistical analysis of the breakup cascade:** ASSET enabled a detailed genealogy of breakup events linking individual droplets back to their parent structures. The resulting statistical data provide new insights into how ligaments form and then fragment into smaller droplets, characterizing both primary and secondary atomization

pathways. Ligament-driven breakup was shown to account for a large fraction of the atomized volume and is found to dominate secondary atomization and cloud formation.

3. **Preliminary exploration of shear-thinning fluid dynamics in the SIE–RTP transition regime:** Comparison of water (Newtonian) and a shear-thinning fluid with a power-law viscosity revealed that the higher effective viscosity of the shear-thinning fluid substantially alters late-stage breakup. Specifically, shear-thinning droplets resisted bag formation, despite a higher effective Ohnesorge number, rather retaining coherence longer than the water drop and forming elongate sheets in the direction of flow. This result suggests that shear-thinning can shift conventional regime boundaries, highlighting the need for more comprehensive parameter sweeps.

4. **DNS of large density ratios and high *Re*−*We* regimes:** By simulating droplet aerobreakup at realistic Weber numbers ($We = 825$) and liquid-to-gas density ratios ($\rho_l/\rho_g = 816$), this work extends current DNS capabilities to capture multi-scale breakup dynamics. While the simulations show a dependence on grid resolution—particularly for fine features such as ligaments, thin sheets, and bag rupture, all of which shape the final drop-size distributions—they still exhibit robust qualitative agreement with experimental trends. Crucially, the underlying mechanisms of instability formation and droplet production remain consistent across resolutions, indicating significant promise for future high-fidelity DNS as computational resources and modeling approaches continue to evolve.

5. **Future work:** A natural extension of this work is to conduct a more comprehensive parametric study on the effects of shear-thinning rheology, examining a range of consistency and flow indices to determine how these parameters shift both the large- and small-scale breakup mechanisms. Additionally, future simulations would benefit from advanced sub-grid-scale interface-tracking models that better capture the delicate

physics of ligament pinch-off and thin-sheet (or bag) rupture, which are often under-resolved at practical grid sizes. Finally, lengthening the overall simulation time would help illuminate the full droplet lifetime, extending from primary breakup and ligament formation to even later stages of secondary and tertiary fragmentation, ultimately providing a more complete picture of the atomization cascade.

Overall, these simulations advance our understanding of droplet aerobreakup in regimes that are both theoretically complex and industrially relevant. By elucidating the interplay between SIE and RTP instabilities, quantifying the cascade from ligament formation to secondary atomization, and preliminarily mapping how shear-thinning properties reshape breakup pathways, this work lays the groundwork for more accurate and robust predictive models of atomizing flows across a spectrum of engineering and environmental applications.

A COMPUTATIONAL FRAMEWORK FOR SIMULATING AND ANALYZING ATOMIZATION IN ELECTRICALLY CHARGED DIESEL-TYPE JETS

Contribution of Authors and Co-Authors

Manuscript in following chapter

Author: Brendan Christensen

Contributions: Methodology, Investigation, Formal Analysis, Data Curation, Writing - Review & Editing, Visualization

Author: Bret Van Poppel

Contributions: Conceptualization, Funding Acquisition, Writing - Original Draft

Author: Mark Owkes

Contributions: Methodology, Writing - Review & Editing, Supervision, Funding Acquisition

Manuscript Information

Brendan Christensen, Bret Van Poppel, and Mark Owkes

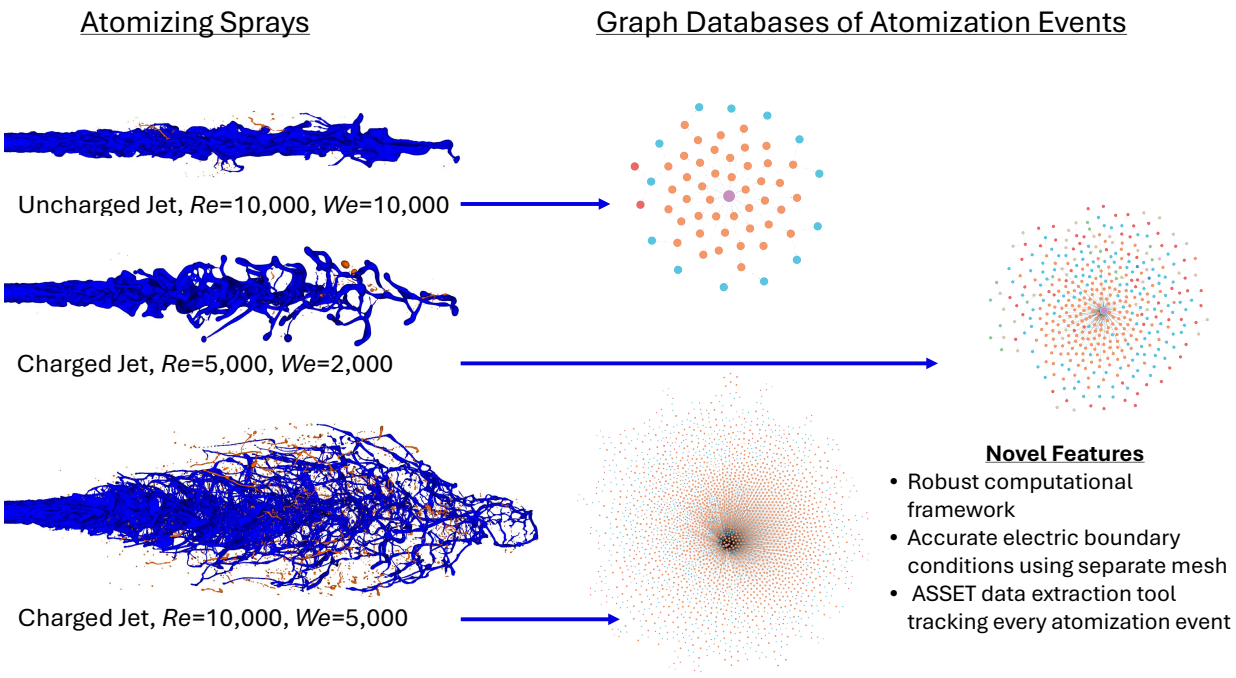
International Journal of Multiphase Flow

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
_X_ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Elsevier Ltd.

Abstract

A robust computational framework for simulating primary atomization in electrically charged liquid hydrocarbon jets is developed, tested and demonstrated in this work. First-principles-based numerical methods developed specifically for high-fidelity direct numerical simulations of electrohydrodynamic (EHD) flows are implemented within a conservative flow solver using an unsplit, geometric volume-of-fluid transport scheme that includes EHD effects. The numerical framework globally conserves mass, momentum, and the electric charge density even at the gas-liquid interface where discontinuities exist. A novel approach using a separate mesh is used to obtain accurate boundary conditions of electric potential on the computational domain. Simulations employ a recently developed physics extraction tool that tracks every breakup and coalescence event occurring during an atomizing spray. Data characterizing the atomization processes are stored in a Neo4j graphical database providing an easily accessible format to query and study the atomization process. The framework is demonstrated in five simulations focused on a region of interest relevant to primary atomization and varying fluid and EHD parameters, confirming the robustness of the methodology and its tools. Statistics for droplet size, breakup and coalescence events, onset of primary atomization, and spray cone angle are reported for all simulations revealing earlier and enhanced liquid breakup and greater spray dispersion with increasing electric charge. Some comparison with experimental work reveals the role of electric stress in destabilizing the jets and breakup consistent with the Plateau-Rayleigh mechanism.



Introduction

Electrohydrodynamics (EHD) is an interdisciplinary topic that describes the complex interaction between fluid mechanics and electric fields, explored since Lord Rayleigh revealed the competing forces and resulting dynamics of a liquid drop that is radially stressed by an applied electric field [153] and subsequently by many researchers. EHD may enable improved spray control and finer atomization so that fuel injection schemes can be inexpensively developed for the small combustion engine class. Moreover, EHD may provide efficient enhancements to hydrocarbon fuel atomization that could benefit a much broader range of engines and even other, non-combustion applications, such as pharmaceutical coating and micro- or nano-thin-film deposition [83].

The enhancement of atomization through the injection of electric charge into semi-insulating or dielectric liquids has drawn notable interest over the past several decades, progressing from pioneering work by Taylor [187–190], Melcher [120, 121], Saville [163], Kim

and Turnbull [88] and others. Kelly [87] extended the original demonstration of Kim and Turnbull. Lehr and Hiller [102] used experimental methods to confirm the observations of Kim and Turnbull, and Kelly, and predicted an injected charge density limit due to the observed phenomenon of corona discharge. Shrimpton and co-workers have studied the technique of "charge injection" through experimental [155, 179, 220] and numerical [176, 177] investigations.

Notwithstanding these efforts, the complex interactions among electrostatic charge, electric fields, and the dynamics of atomizing liquids are not well understood. Considering the challenges and expense of experiments, high-fidelity numerical simulations should be able to provide some clarity to the underlying fundamentals and dynamics of EHD atomization. While numerical simulations of realistic liquid breakup are computationally expensive to resolve small structures and require large, parallel resources to perform these simulations, recent advances in scientific computing, numerical methods, and statistical tools render simulation a viable method of investigation. Leveraging these developments, several numerical studies have been conducted over the past few decades, including early work by Baygents and co-workers [10], Collins and co-workers [27], Tomar and co-workers [196], Feng [47, 48], Sherwood [173], and Shrimpton [178]. Hu [74], Paknemant [141], Hokmabad [70], Du [40], Xie and co-workers [216], Dong and co-workers [38], Holgate [71], and Frederich [49], among many others, have explored EHD systems for a range of electrical parameters, reporting two-dimensional results for mostly static droplets and structures under an externally applied electric field. Schnitzer, Yariv and co-workers employed an electrokinetics approach to analyze EHD phenomena for conducting drops [166] and to address the Taylor-Melcher leaky dielectric model commonly used to describe the role of electric stress in poorly conducting liquids [120, 164, 167, 190]. López-Herrera and co-workers simulated Taylor cone-jets solving the electrokinetic equations [115], one of the most recent of many works exploring semi-conducting or ionic liquids. Few works published to date

report fully three-dimensional, high-fidelity multiphase simulations of electrically charged Diesel-type jets—liquids with very low conductivity and few charge carriers—to elucidate primary atomization mechanisms and processes that are comparable to classic experiments, such as [178, 179, 220].

The distribution of electric charges within an atomizing flow is governed by two main fluid characteristics: the charge advection and charge mobility timescales [29, 90]. The ratio of these timescales, defined here as the electric Reynolds number (Re_e) [183], dictates how fast the electric charges are advected by the flow relative to how fast they relax to the fluid interface. Thus, the EHD fluid properties and flow characteristics control the resulting charge distribution within atomizing flows. Van Poppel [203] employed first-principles-based numerical methods in high-fidelity direct simulations of electrically charged liquid hydrocarbon jets undergoing electrostatic-enhanced primary atomization. Subsequently, Owkes and Sheehy [172] improved upon these methods to include charge mobility and conducted three-dimensional simulations of electrically charged liquid jets to assess the effects of Re_e on atomization efficiency.

The objectives of this work are *(i)* to establish a computational framework for simulating primary atomization in electrically charged Diesel-type jets; *(ii)* to provide examples of statistics from the recent development and integration of novel tools; *(iii)* and to identify the role of electric stress and breakup mechanisms in primary atomization. The framework detailed in this work extends EHD numerical schemes and integrates two novel, recently developed tools to perform fully three dimensional direct numerical simulations of liquid jets subject to a self-precipitating electric field due to the presence of bulk charge in the liquid. The computational approach employs a state-of-the-art numerical framework that globally conserves mass, momentum, and electric charge density even at the gas-liquid interface where discontinuities exist [204]. The Navier-Stokes solver's numerical scheme sharply handles the discontinuous electric charge density, allowing robust and accurate simulations. A

conservative unsplit geometric volume-of-fluid (VOF) employing semi-Lagrangian transport is used with a piece-wise linear interface reconstruction (PLIC), leveraging the combination of VOF and PLIC to efficiently capture breakup mechanisms and small satellite droplets [149] through its sharp and accurate representation of the phase interface.

Simulation results and analysis assess the effects of EHD on atomization quality, droplet size, and primary atomization events and processes. To better address the electric boundary conditions, known to be challenging for EHD flows, simulations employ a recently developed, novel approach using a separate mesh to obtain accurate boundary conditions of electric potential on the computational domain [93, 94]. Known as *e-Mesh*, this mesh is much larger than the flow solver domain, implemented with well-defined boundary conditions for the electric potential, ϕ, and used to compute electric boundary conditions accurately and efficiently for the flow-solver domain. Another improvement is the implementation of a novel physics extraction tool that tracks every breakup and coalescence event occurring during an atomizing spray. Data characterizing the atomization processes are stored in a Neo4j graphical database providing an easily accessible format [24, 159] for analyzing primary and secondary atomization processes, although this work reports only the results of primary atomization events.

This paper is organized with §4 describing the governing equations and interface boundary conditions with relevant simplifications followed by §4 which summarizes the numerical approach employed in the flow solver and the e-Mesh tool. Data extraction employing the ASSET tools is summarized in §4 with some details provided for clarity and completeness. The simulation domain and parameters are provided in §4. Simulations are compared and primary atomization statistics and mechanisms are reported and discussed in §4.

Mathematical Formulation

This work solves the full set of governing equations to predict the dynamics of the flow (velocity and pressure), gas-liquid interface topology, and electric charge distribution within the liquid phase. The dynamics of the charge distribution are controlled by diffusion and the electric field created by the presence of space charge in the liquid domain. While assuming a constant charge density (insulating fluid) or surface charge (conducting fluid) has been done, notably reported in [27] among many others, predicting the non-trivial charge dynamics is important for accurate simulations. The formulation presented here couples the electrostatic and hydrodynamic phenomena for a poorly conducting liquid through the Maxwell stress tensor and implements gas-liquid interface boundary conditions.

Governing Equations

Conservation of mass and momentum for a variable density, low Mach number flow are given as

$$\frac{\partial \rho_i}{\partial t} + \nabla \cdot (\rho_i \boldsymbol{u}_i) = 0, \tag{19}$$

and

$$\frac{\partial \rho_i \boldsymbol{u}_i}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}_i \otimes \boldsymbol{u}_i) = -\nabla p_i + \nabla \cdot (\boldsymbol{\sigma}_i^f + \boldsymbol{\sigma}_i^e) + \rho_i \boldsymbol{g},$$

where $\boldsymbol{u}_i$ is the velocity field, p_i is the hydrodynamic pressure, ρ is the mass density, $\boldsymbol{g}$ is gravitational acceleration. The subscript i represents the phase and can take the value g for the gas phase or l for the liquid phase. The viscous stress tensor, $\boldsymbol{\sigma}_i^f$, is given by

$$\boldsymbol{\sigma}_i^f = \mu_i \left(\nabla \boldsymbol{u}_i + \nabla \boldsymbol{u}_i^t \right) - \frac{2}{3} \mu_i \nabla \cdot \boldsymbol{u}_i \mathbb{I}, \tag{20}$$

with $\mathbb{I}$ the identity tensor and μ_i is the dynamic viscosity. The Maxwell stress tensor, $\boldsymbol{\sigma}_i^e$ is used to define the electric body force, $\boldsymbol{f}_i^e = \nabla \cdot \boldsymbol{\sigma}_i^e$, and leads to,

$$\boldsymbol{f}_i^e = q_i \boldsymbol{E}_i - \frac{1}{2}\boldsymbol{E}_i^2 \nabla \varepsilon_i + \nabla \left(\frac{1}{2} \rho_i \frac{\partial \varepsilon_i}{\partial \rho_i} \boldsymbol{E}_i^2 \right), \tag{21}$$

where ε_i is the electric permittivity, q_i is a volumetric electric charge density, and $\boldsymbol{E}_i$ is the electric field vector. The three terms of $\boldsymbol{f}_i^e$ represent the Coulomb (or Lorentz) force, the dielectric force, and the electrostrictive force, respectively. Assuming incompressibility and material homogeneity within a phase, $\boldsymbol{f}_i^e$ reduces to the Coulomb force in the bulk [20, 90] as

$$\boldsymbol{f}_i^e = q_i \boldsymbol{E}_i. \tag{22}$$

The electric displacement vector, $\boldsymbol{D}$, is assumed to vary linearly with the electric field vector, as

$$\boldsymbol{D}_i = \varepsilon_i \boldsymbol{E}_i. \tag{23}$$

The electric field vector is irrotational, hence it can be expressed as the gradient of the scalar electric potential, ϕ_i, as

$$\boldsymbol{E}_i = -\nabla \phi_i. \tag{24}$$

The Gauss law can be employed for a dielectric material to relate the electric displacement vector to the volumetric charge, as

$$\nabla \cdot \boldsymbol{D}_i = \nabla \cdot \varepsilon_i \boldsymbol{E}_i \tag{25}$$

which can be simplified to the Electric Potential Poisson Equation, as

$$\nabla \cdot (\varepsilon_i \nabla \phi_i) = -q_i. \tag{26}$$

Charge conservation and current density are described by

$$\frac{\partial q_i}{\partial t} + \nabla \cdot \boldsymbol{J}_i = 0, \tag{27}$$

$$\boldsymbol{J}_i = K_i q_i \boldsymbol{E}_i - \mathcal{D}_i \nabla q_i + q_i \boldsymbol{u}_i, \tag{28}$$

where $\boldsymbol{J}_i$ is the current density, $\mathcal{D}_i$ is the molecular diffusion coefficient and K_i is the ionic mobility coefficient.

Interface Boundary Conditions

The interface boundary and jump conditions are essential to capturing the dominant mechanisms influencing the flow and atomization. While the velocity and tangential electric field components are continuous across the interface, the mass density, viscosity, and electric displacement vector experience jumps, described by

$$[\boldsymbol{u} \cdot \boldsymbol{n}]_\Gamma = 0, \tag{29}$$

$$[\boldsymbol{u} \cdot \boldsymbol{t}_j]_\Gamma = 0, \quad \text{for } j = 1, 2 \tag{30}$$

$$[\rho]_\Gamma = \rho_l - \rho_g, \tag{31}$$

$$[\mu]_\Gamma = \mu_l - \mu_g, \tag{32}$$

$$[\varepsilon]_\Gamma = \varepsilon_l - \varepsilon_g, \tag{33}$$

$$[\boldsymbol{D}]_\Gamma = \boldsymbol{n} \cdot [\varepsilon \boldsymbol{E}]_\Gamma = q_s, \tag{34}$$

$$\boldsymbol{n} \times [\boldsymbol{E}]_\Gamma = 0, \tag{35}$$

where $[(\cdot)]_\Gamma$ represents the jump of "$(\cdot)$" across the interface, Γ, and for example, $[\boldsymbol{D}]_\Gamma$ represents the jump of the electric displacement vector across the interface, $\boldsymbol{n}$ and $\boldsymbol{t}$ represent the normal and tangential vector components at the interface, subscripts l and g represent quantities in the liquid and gas phases, respectively, and q_s the surface charge. The

electric charge is modeled as volumetric space charge that migrates throughout the liquid, which implies negligible surface charge and leads to the normal component of the electric displacement vector being continuous. However, the jump in electric permittivity across the phase interface imposes a discontinuity in the normal component of the electric field, and Eq. (34) can be written as

$$\boldsymbol{n} \cdot [\boldsymbol{D}]_\Gamma = \boldsymbol{n} \cdot [\varepsilon \boldsymbol{E}]_\Gamma = [\varepsilon E_n]_\Gamma = 0. \tag{36}$$

A direct consequence of the irrotational jump condition, Eq. (35), ensures that the tangential components of the electric field, and therefore the electric potential, ϕ, are continuous,

$$[E_{t_1}]_\Gamma = 0, \quad [E_{t_2}]_\Gamma = 0, \text{ and} \tag{37}$$

$$[\phi]_\Gamma = 0. \tag{38}$$

The pressure interface jump condition includes contributions from viscous, electric and surface tension forces, and is represented as

$$-[p]_\Gamma + [\boldsymbol{n}^\mathsf{T} \cdot (\boldsymbol{\sigma}^e + \boldsymbol{\sigma}^f) \cdot \boldsymbol{n}]_\Gamma = -\gamma\kappa, \tag{39}$$

which can be simplified as

$$[p]_\Gamma - 2[\mu]_\Gamma \boldsymbol{n}^\mathsf{T} \cdot \nabla \boldsymbol{u} \cdot \boldsymbol{n} - \gamma\kappa = \left[\frac{1}{2}\varepsilon E_n^2\right]_\Gamma - \left[\frac{1}{2}\varepsilon E_{t_i}^2\right]_\Gamma, \tag{40}$$

where $[p]_\Gamma$ is the jump across the interface, γ the surface tension coefficient, and κ the local interface curvature. The tangential shear stress balance, again employing negligible surface charge, yields

$$\left[\boldsymbol{n}^\mathsf{T} \cdot (\boldsymbol{\sigma}^f) \cdot \boldsymbol{t}_i\right]_\Gamma = 0. \tag{41}$$

Additional details are provided in [172, 203].

Numerical Approach

The simulations in this paper were conducted using the NGA computational platform, which employs a staggered grid with pressure and any scalars stored at the cell center and velocity components stored on cell faces [33]. Away from the interface, we use second-order finite difference operators that conserve mass, momentum, electric charge density, and kinetic energy. Near the interface, where discontinuities exist, special numerical methods are required. For convection, we use an unsplit, geometric, semi-Lagrangian framework. The framework is based on Owkes and Desjardins [138] and is used to transport the volume-of-fluid (VOF) representation of the phase interface, momentum, and electric charge density. Due to the staggered grid, consistent transport of both VOF and momentum is complex, requiring that the fluxes of mass used to transport VOF and momentum are discretely equal. This is achieved by discretizing VOF on a mesh twice as fine as the flow solver and computing fluxes on all faces of pressure and momentum cells, implemented as an unsplit, geometric, semi-Lagrangian with details provided in Owkes and Desjardins [139]. The resulting scheme is second-order and conserves mass, momentum, and scalars even near the interface where discontinuities exist. The interface is reconstructed using a piece-wise linear interface reconstruction (PLIC) [127] with interface normal vectors computed using the efficient least-squares VOF interface reconstruction algorithm (ELVIRA) [147]. Interface curvature is computed with the adjustable curvature evaluation scale (ACES) method [136]. NGA is fully parallelized with a message passing interface (MPI) protocol and scales well to tens of thousands of cores [36].

The electric field is computed from the electric potential using Eq. 24, which depends on the boundary conditions as well as the distribution of electric charge within the domain. The electric potential Poisson equation, Eq. 26, provides a relationship between the electric

charge, q_i, and the electric potential, ϕ_i, and is solved using the *hypre* library of routines for scalable parallel solutions of linear systems [44, 45]. The pressure and electric potential Poisson equations are solved utilizing the ghost fluid method [46] and a black box solver [30].

The challenge of obtaining accurate electric potential boundary conditions on the computational domain for electrohydrodynamic flows is well documented, for example in [172, 203] among many others. For the model used in this work, grounded ($\phi = 0$) conditions exist at the boundaries of the computational domain. For smaller computational domains with higher electric charge concentrations, some mesh-alignment of the spray can result. To eliminate this numerical artifact of the computational model, a second domain called e-Mesh is introduced that spans from the nozzle to surfaces that represent physical boundaries that are grounded. The e-Mesh grid is implemented with well-defined boundary conditions and the charge density field, which is transferred from the flow domain. The electric Poisson equation, Eq. 26, is solved in e-Mesh to obtain a potential field and values of ϕ are transferred from e-Mesh to the boundaries of the flow domain to compute ϕ in the flow domain. Solving for the potential on the e-Mesh and then interpolating to the boundaries of the smaller flow domain allows for accurate boundary conditions of ϕ on the flow domain where they are needed but not well defined. The electric potential field, ϕ, is computed on e-Mesh and values are interpolated to the Navier-Stokes mesh using a tri-cubic interpolation scheme. Figure 37b shows the enlarged and coarse e-Mesh domain surrounding the much smaller and finer Navier-Stokes mesh. The e-Mesh is fully described in [93] and demonstrated in a numerical study in [94].

Data Extraction

Simulations employed the atomization simulation statistics extraction tool (ASSET), a recently developed physics extraction tool that tracks every breakup and coalescence event occurring during an atomizing spray. Data characterizing the atomization processes were

stored in a Neo4j graphical database providing an easily accessible format. An in-depth description of the process is available in previous works [24, 159], but a brief synopsis is included here for clarity.

ASSET operates at the same time as the simulation and runs every timestep, identifying droplets, which meet the criteria for either breakup or coalescence within that step. The algorithm works by assigning two integer identification numbers to every liquid structure within the domain. The first of these values ($\mathcal{S}$) is reassigned every timestep using a connected-component algorithm to identify every independent liquid structure. A second value ($\mathcal{L}$) is transported as a scalar value with the liquid and is only reassigned when a new droplet is formed either from breakup or coalescence. The sequence of events for one timestep is as follows:

1. Transport: Volume fraction, physical properties, and $\mathcal{L}$ values are transported via flux volumes.

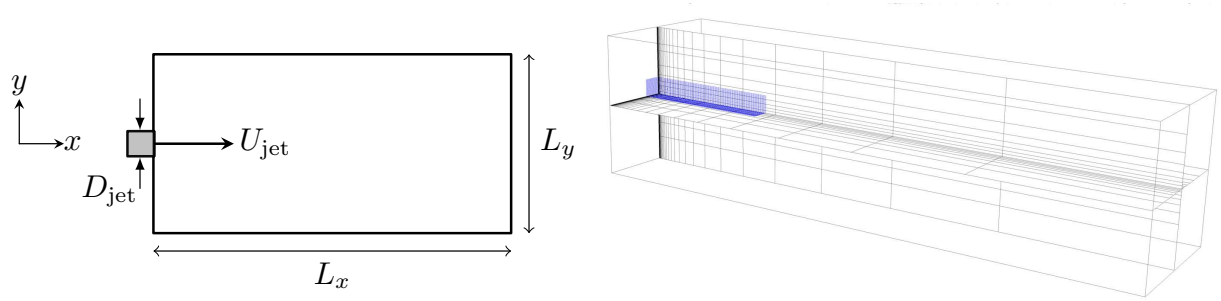
2. $\mathcal{S}$ assignment: A band-growth algorithm [65] assigns unique $\mathcal{S}$ values to structures separated by at least one full gas cell.

3. Structure data calculation: Each structure is added to a linked list with properties like volume, velocity, principal axis lengths, and average charge density.

4. Coalescence identification: Cells containing more than one $\mathcal{L}$ value processed as a single droplet, assigned a new $\mathcal{L}$

5. Breakup identification: Structures sharing an $\mathcal{L}$ but with distinct $\mathcal{S}$ values are marked as having broken up.

6. Data extraction: Properties of all identified coalescence and breakup structures are saved, along with old and new $\mathcal{L}$ values to construct a droplet "genealogy."



(a) Flow solver domain in two dimensions. (b) Navier-Stokes and e-Mesh domain.

Figure 37: Simulation domain (left); Navier-Stokes mesh in blue and e-Mesh in black (right) showing relative scale of each domain.

Table 6: Non-Dimensional simulation parameters.

Sim ID	Reynolds, Re_l $\rho_l U_{\text{jet}} D_{\text{jet}}/\mu_l$	Weber, We_g $\rho_g U_{\text{jet}}^2 D_{\text{jet}}/\gamma_l$	Electro-inertial, N_{ei} $q_l^2 D_{\text{jet}}^2/(\varepsilon_l \rho_l U_{\text{jet}}^2)$	Density ratio ρ_l/ρ_g	Visc. ratio μ_l/μ_g	Resolution* $D_{\text{jet}}/\Delta x$
Lo_00	5000	2	Uncharged	1000	50	30 (60)
Lo_05	5000	2	0.05	1000	50	30 (60)
Lo_10	5000	2	0.1	1000	50	30 (60)
Hi_00	10000	10	Uncharged	1000	50	30 (60)
Hi_10	10000	10	0.1	1000	50	30 (60)

Domain Size $L_x = 24D_{\text{jet}}$, $L_y = L_z = 8D_{\text{jet}}$ *Parenthetical value represents VOF resolution described in §4

Computational Domain and Simulations

Five liquid Diesel-type jets were simulated with liquid injection through a circular port of diameter D_{jet}, covering a range of relevant electrohydrodynamic and flow parameters. Simulation parameters are summarized in Table 6, with bulk velocity, U_{jet}. The domain size employed was $L_x = 24D_{\text{jet}}$, and $L_y = L_z = 8D_{\text{jet}}$ discretized on a uniform mesh of different sizes commensurate with the number of cells across the nozzle diameter, CAD, reported in Table 6. The Navier-Stokes computational domain is illustrated in Figure 37a. The electric potential boundary conditions implemented through e-Mesh for the charged simulation are Neumann on the upstream x-face, and Dirichlet ($\phi = 0$) on the downstream x-face and y- and z-faces. Figure 37b displays the e-Mesh surrounding the Naver-Stokes

(a) Lo_00 uncharged simulation. (b) Lo_05 charged simulation.

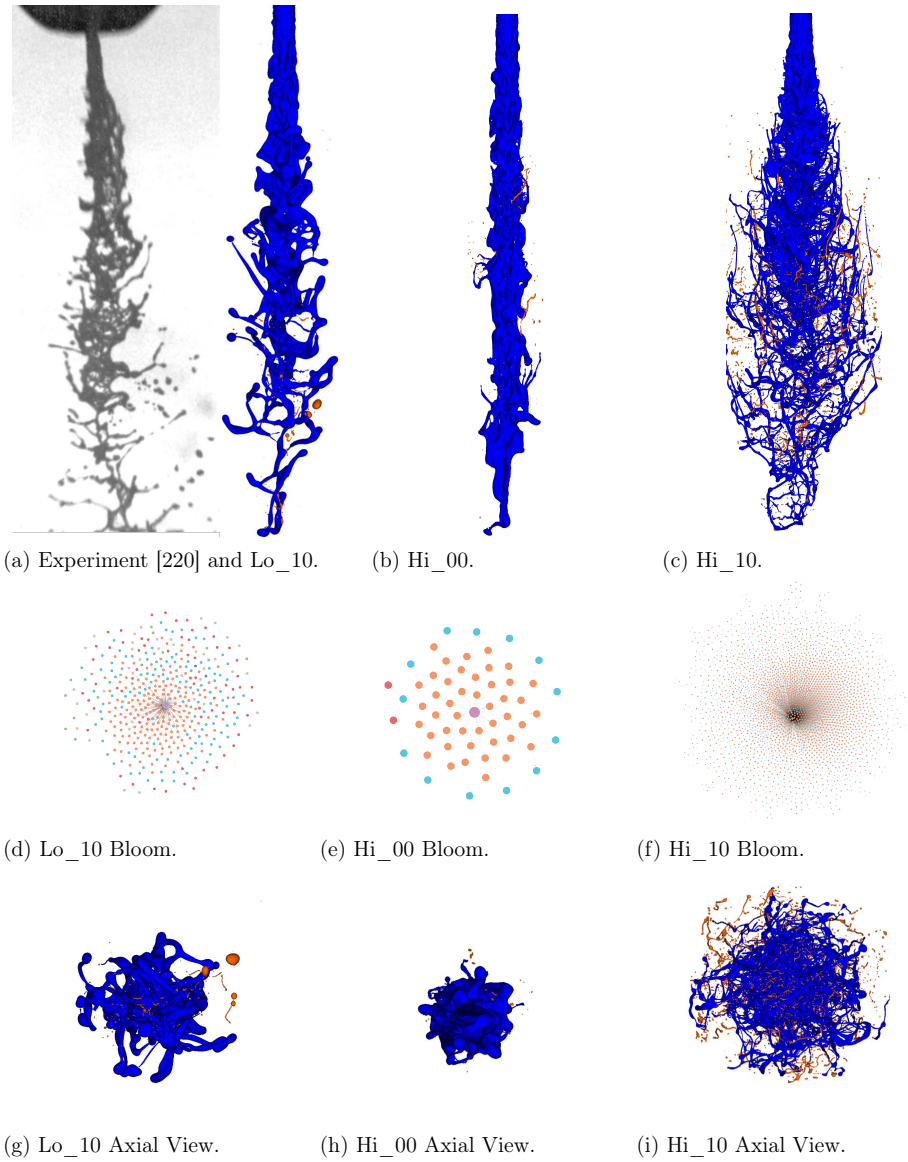
Figure 38: The Lo_00 (uncharged) and Lo_05 (low electric charge) simulations, showing minimal liquid breakup.

domain to provide a degree of scale and illustrate the e-Mesh stretching scheme employed to improve efficiency with refinement near the nozzle and coarse spacing further from the injection location. As an initial investigation, EHD parameters were chosen similar to those reported in [172] and bulk hydrodynamic parameters similar to those reported in [24]. The electric Reynolds number, Re_e—a ratio of the residence time of liquid in the domain to the charge relaxation timescale [183]—is maintained at a moderate level of approximately 780 for all charged simulations, thereby permitting some charge mobility representative of poorly conducting liquids but limiting charge relaxation within the domain.

Results

Five cases were run for analysis and comparison, varying global Weber and Reynolds numbers and electro-inertial (N_{ei}) values. The cases are outlined in Table 6. The Lo_00 and Lo_05 simulations are shown separately in Figure 38 because these cases show minimal liquid breakup for analysis, although case Lo_05, Figure 38b, exhibits minimal droplet shedding. Figure 39 displays renderings of the Lo_10, Hi_00, and Hi_10 simulations in Figures 39a, 39b, and 39c, colored by liquid volume attached to the continuous liquid core (blue) and structures detached from the core (orange), illustrating the utility of the ASSET extraction tool for analyzing primary atomization mechanisms resulting from breakup events.

Figure 39a is shown next to experimental work reported by Shrimpton and co-



(a) Experiment [220] and Lo_10. (b) Hi_00. (c) Hi_10.

(d) Lo_10 Bloom. (e) Hi_00 Bloom. (f) Hi_10 Bloom.

(g) Lo_10 Axial View. (h) Hi_00 Axial View. (i) Hi_10 Axial View.

Figure 39: Renderings of experiment reported by Yule and Shrimpton [220] (reproduced with permission) and Lo_10 simulation (a, images not the same scale), Hi_00 (b), and Hi_10 (c) simulations showing contiguous liquid core (blue) and dislodged structures (orange)); Bloom plots visualizing breakup events in simulations using ASSET (middle row, (d)-(f)); and axial view of simulations (bottom row, (g)-(i)). Full simulation parameters detailed in Table 6.

workers [179, 220] for comparison of cases with closely matched hydrodynamic and electric parameters revealing strong qualitative similarities, including the point of primary atomization onset, droplet size and other structures within the disintegrating jets. The

presence of numerous ligaments oriented radially outward from the axial centerline bears strong resemblance to the experimental photograph, and the radial orientation of liquid structures is attributable to mutual repulsion arising from the space charge field.

Unlike simpler jet-to-dripping regimes, case Hi_00 did not exhibit a clear pinch-off point. Some droplet shedding was observed in Hi_00; however, complete disintegration of the liquid core was not evident. The presence of electric charge in the liquid plays a role in primary atomization, leading to initial de-stabilization. Electric effects on secondary atomization processes are expected to be small for these simulations due to the low values of electric charge compared to inertial and surface tension forces. The mutual repulsion of unipolar electric charge leads to greater spray dispersion through the action of the Coulomb force in the liquid bulk, thereby altering droplet trajectories and velocities after separating from the liquid core. The influence of electric charge is evident in Figures 39a (Lo_01) and 39c (Hi_01), particularly compared to their uncharged companion simulations in Figures 38a (Lo_00) and 39b (Hi_00), revealing the influence of electric effects in de-stabilizing the jet. The wider plume in Figure 39b, completely filling the domain, suggests a strong interplay among inertial, viscous, surface tension, and electric effects. Generally, increasing electric charge moves the initial jet de-stabilization and point of atomization onset upstream, increases the number of structures dislodged from the liquid core and "radial fullness" of liquid in the domain, and widens the spray cone angle. The radial orientation of liquid structures in simulations with greater electric charge is attributable to Coulombic repulsion that arises from the space charge field, and is similar to that discussed elsewhere, for example in Tang and Gomez [185] among others. Coulombic repulsion leads to an electric field oriented radially from the liquid core, thereby confirming that the Coulomb force is a disruptive force that gives rise to the shape and orientation of the structures dislodged from the liquid core.

Uncharged and Charged Simulation Comparison

The uncharged and charged simulations demonstrate the robustness of the computational framework and their recently integrated tools, highlighting the notable differences in atomization processes due to electrostatic effects. Hydrodynamic and electric parameter values are listed in Table 6 for cases Hi_00 and Hi_10, identical except that there is no background electric potential and no liquid charge density in the uncharged simulation (Hi_00). The two cases are compared to elucidate the effects of electric charge and the corresponding electric field on the breakup mechanisms driving primary atomization while also demonstrating the accuracy and utility of the ASSET and e-Mesh tools. Figures 38a and 39c provide a visual comparison of the atomizing jets for the uncharged case (39a, Hi_00) and charged case (39c, Hi_10) at the simulation final time step. The presence of electric charge within a realistic Diesel type jet contributes to an earlier onset of primary atomization, and it modifies the interface shape leading to a fuller and wider plume and increased spray cone angle. The charged simulation, Figure 39c, shows the presence of droplets dislodged from the liquid core as well as interface corrugations and ligaments. At this point in the simulation, the charged case, Figure 39c, produced more than 2000 droplets; however, at the same point in the uncharged simulation, Figure 38a, only 50 droplets were created.

Figures 39e and 39f display "bloom" plots from the data stored in a Neo4j database demonstrating how the event data is organized in a logical and accessible format. The nodes are the liquid core and droplets produced from breakup events, which are represented by the lines connecting the nodes. Node colors represent the number of breakup events a structure has undergone. The liquid core is in the center of the image and the orange droplets represent primary breakup events. These images illustrate all the droplets produced during the simulation, and all the intermediate droplets between the coherent liquid core and the final droplets. Paths between nodes represent the connections among parent and child

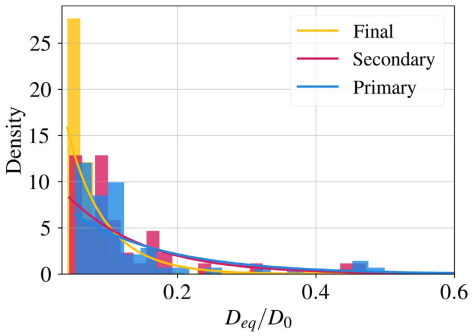
droplets. Each node stores the extracted data from the simulation at the location and time at which that droplet separated from the liquid core and became independent. Figure 39f shows many more droplets and breakup events than the uncharged bloom plot in Figure 39e.

Charged Simulations Comparison

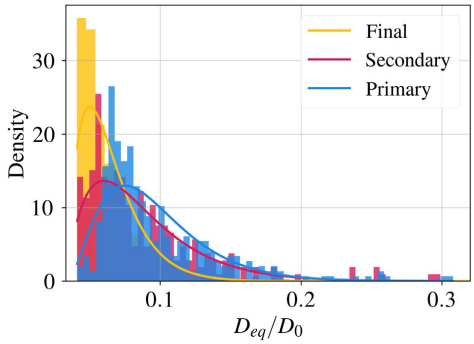
The two cases with highest electro-inertial number, $N_{\text{ei}} = 0.1$, exhibit interesting atomization features and a clear pinch-off point. Because of these characteristics, cases Lo_10 and Hi_10 are analyzed and contrasted henceforth, including estimates for the initial onset of primary atomization, breakup and coalescence, charge density distributions in the domain, and estimates for the spray cone angle. Renderings of both cases are shown in Figures 39a and 39c, illustrating a wider spray plume in case Hi_10 attributable to the higher Reynolds number and aerodynamic effects.

Figures 39d and 39f present "bloom" plots generated from the Neo4j database. In these visualizations, nodes represent droplets formed during breakup events, while the lines connecting them signify relationships between parent and child droplets (breakup or coalescence). The colors of the nodes indicate the number of breakup events each droplet has undergone. At the center of each plot lies the liquid core, with all nodes and paths eventually leading back to this core. These visualizations represent the entire droplet breakup cascade. The paths between nodes trace the lineage of droplets as they separate and evolve, while each node stores data about the droplet at the time and location it became independent from its parent droplet. The Lo_10 simulation ultimately produced 269 droplets, whereas the Hi_10 simulation generated 2,014 droplets, reflecting the significant differences in breakup dynamics between the two cases. Table 7 summarizes total droplets and mean diameter for cases Lo_00, Hi_00, and Hi_10, revealing the role of electric effects in jet de-stabilization, droplet production and size.

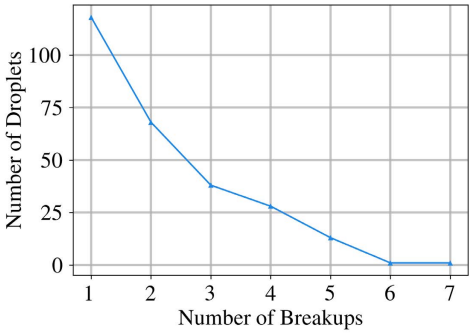
Primary Atomization Analysis



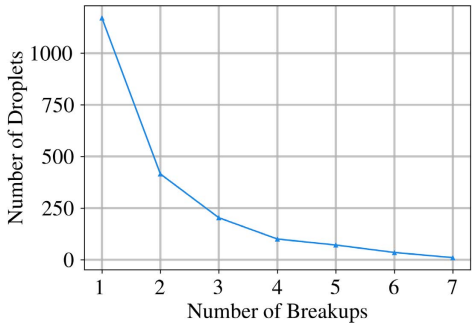
(a) Lo_10 diameter probability density.



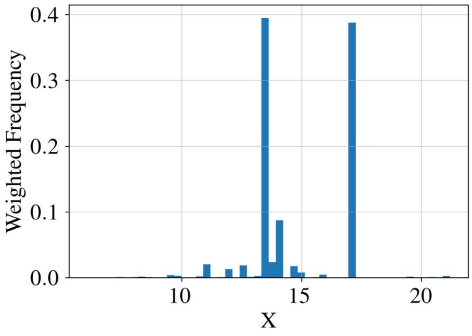
(b) Hi_10 diameter probability density.



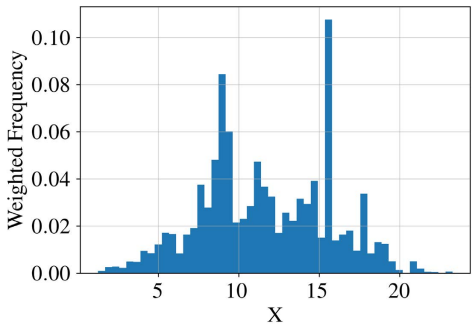
(c) Lo_10 breakup stage of final droplets.



(d) Hi_10 breakup stage of final droplets.



(e) Lo_10 volume-weighted primary breakup locations.



(f) Hi_10 volume-weighted primary breakup locations.

Figure 40: Charged simulations statistics for $N_{\text{ei}} = 0.1$. Lo_10 (left column) and Hi_10 (right column).

Spray cone angle is a useful qualitative description, and in these cases it is an effective parameter to analyze the effects of EHD and Weber and Reynolds numbers. It is clear from these numbers and the jet renderings in Figs. 39a-39c that EHD has a strong effect on the

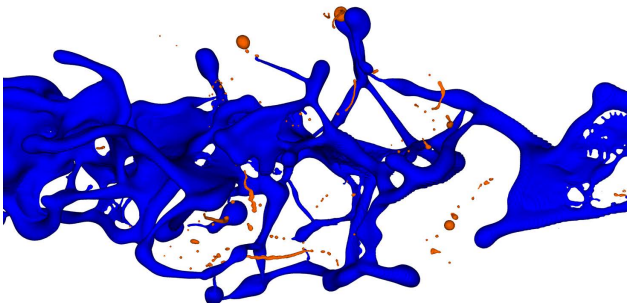


Figure 41: A magnified segment of the Lo_10 simulation, displaying ligaments stretching off of the liquid core (blue) and bulging due to Plateau-Rayleigh instability, ultimately forming large droplets and small satellite droplets (both orange). The VOF resolution for simulations is 60 $D_{\text{jet}}/\Delta x$.

Table 7: Droplet statistics and spray cone angle approximations for select simulations.

Sim ID	Total Droplets	Mean D_{eq}	Spray Cone, $\alpha_{1/2}$
Lo_10	269	$0.2137 D_{\text{jet}}$	8.6°
Hi_00	50	$0.0749 D_{\text{jet}}$	$\sim$
Hi_10	2014	$0.0797 D_{\text{jet}}$	14.0°

spray cone angle. The spray cone is defined here as a half-angle from the jet radius. For each axial and radial location, we calculate the volume of fluid in the corresponding radial ring, and define the spray boundary as the cone where 99% of the liquid volume is contained within. The half-angle, $\alpha_{1/2}$, is calculated from a right triangle, as $\alpha_{1/2} = \tan^{-1}(r_\alpha/x_\alpha)$, where r_α is the greatest radial extent meeting the threshold and x_α is the corresponding axial distance from the nozzle. Table 7 reports spray angles measured at the final timestep in each of the five simulations. Both of the uncharged simulations, Lo_00 and Hi_00, experienced negligible radial spread. The plume width increases notably with added charge. Simulations with higher Reynolds and Weber numbers showed greater radial spread, suggesting a strong interaction among inertial, surface tension, aerodynamic, and electric effects, and likely attributable to the relative thinning of liquid structures resulting from higher Reynolds and Weber numbers. These thinner structures are also more susceptible to the effects of the Coulomb force, thereby increasing their radial dispersion.

Figures 40a and 40b illustrate the droplet size distribution across different stages of the breakup process for both simulations. The data reveals that the Lo_10 simulation generated a greater proportion of larger-diameter droplets compared to the Hi_10 simulation. A clear progression is evident in the plots: larger primary structures give way to smaller secondary structures, culminating in a significant prevalence of small droplets classified as "final" droplets—those that do not undergo further breakup within the simulation. The mean diameter of final droplets was approximately $0.2137D_{\text{jet}}$ in the Lo_10 simulation and $0.0797D_{\text{jet}}$ in the Hi_10 simulation, highlighting a notable difference in the resulting droplet distributions. The sequence of breakup events leading to these final droplets is depicted in Figures 40c and 40d showing the number of breakup events required to reach the final droplet stage, with some structures undergoing as many as seven breakups before stabilizing into their final form.

The ASSET methodology permits the extraction of the locations of primary breakup events. However, sufficient data for meaningful analysis was available only for simulations Lo_10 and Hi_10. Figures 40e and 40f show the total volume of droplets shed at each location along the x-axis. This weighted distribution highlights the regions where the jet becomes most unstable and undergoes significant disintegration. In the Hi_10 simulation, two pronounced peaks appear around $X = 8$ and $X = 16$. The first corresponds to a similar spike in the number of breakup events, indicating both high droplet count and size. The second peak likely represents a large liquid structure detaching from the core, as suggested by a relative decrease in total droplet production seen in Figure 40e. The Lo_10 simulation also shows two peaks, but given its comparatively low droplet output, both peaks are more plausibly linked to large liquid structures separating from the liquid core.

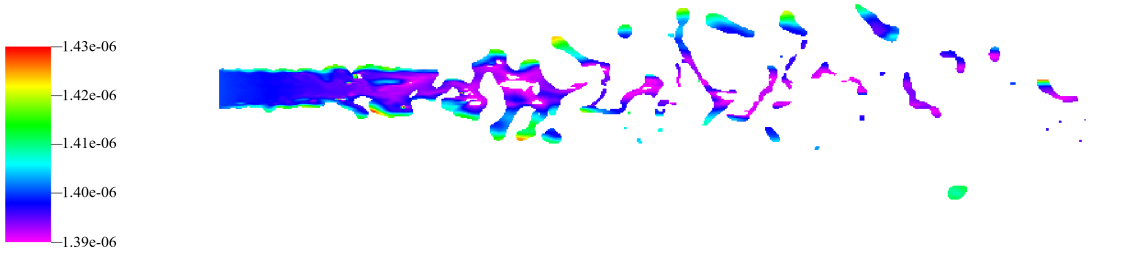
Analyzing the drop size distributions in Figures 40a and 40b confirm the prevalence of small droplets in both simulations. The Lo_10 simulation maintains larger droplets and many fewer droplets total as seen in Table 7. In both cases, there are few large droplets

and many small droplets. This distribution is largely due to capillary effects driven by the Plateau-Rayleigh instability—smooth ligaments are stretched from the liquid core, after which surface tension forces create bulges linked by thinner ligaments, ultimately leading to pinch-off. The resulting thin ligaments formed undergo the same process, producing smaller "satellite" droplets, described in experiments reported by Qian and Law [151] and later in Villermaux [207]. Figure 41, a magnified segment of the Lo_10 simulation, illustrates this phenomenon, showing several ligaments with bulges and smaller satellite ligaments and droplets. Furthermore, electric effects have been shown to reduce the influence of surface tension, leading to enhanced secondary atomization, more droplets, and a greater number of breakup and coalescence events.

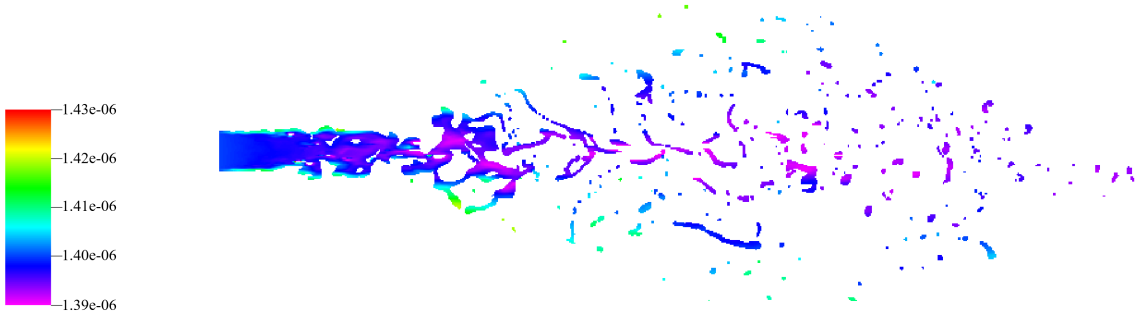
Electric Charge Density

For completeness, a preliminary assessment of electric charge density is provided here. Figure 42 illustrates x-y cut planes of the Lo_10 and Hi_10 simulations colored by electric charge density showing charge migration through the liquid. The effects of charge mobility even near the issuing nozzle can be observed in both Figure 42a and 42b with higher charge density near the surface resulting from the unipolar charge mutual repulsion. Ligaments and structures dislodged further from the centerline carry higher charge density as both jets disintegrate downstream from primary atomization mechanisms and the interplay among electric stress, inertial forces, and surface tension. Further downstream, structures originating closer to the axial centerline carry lower charge density with levels similar to those nearer jet center prior to initial disintegration.

Using the ASSET tool, values of charge density (L_Q) were extracted for every final droplet in the domains of the Lo_10 and Hi_10 simulations. Electric charge levels are shown as distributions in Figure 43a and 43b, with the initial injected charge value represented as a dashed red line. The distributions are final droplets, meaning droplets that do not experience

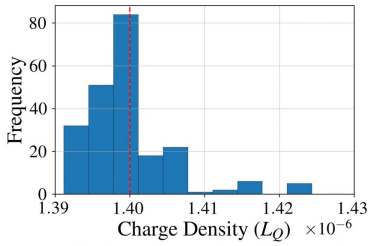


(a) Lo_10 interface cut at x-y plane.

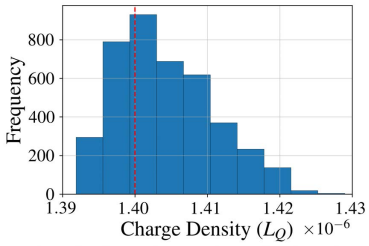


(b) Hi_10 interface cut at x-y plane.

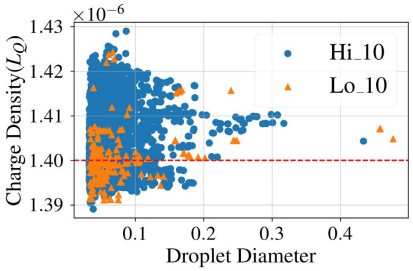
Figure 42: Axial planes x-y colored by charge density.



(a) Lo_10 charge density distribution.



(b) Hi_10 charge density distribution.



(c) Droplet charge distribution.

Figure 43: Charge density distributions for droplets in the Lo_10 (a) and Hi_10 simulations (b); scatter plot of droplet charge distribution in Lo_10 (orange) and Hi_10 (blue) simulations (c). Red dashed line shows initial value for L_Q of the injected liquid volume.

further breakup in the simulation. Charge density may vary within droplets, but the values reported in Figure 43 are mean charge values within each droplet. Most droplets contain a charge density that is very close to the initial bulk charge. Because charge migration is

included in the model, there is a distribution of charge densities in the droplets. Charge migration toward the outer surface of a liquid jet results in higher charge density away from the jet centerline and a higher value of L_Q in structures dislodged near the liquid-gas interface. This trend is more pronounced in the Hi_10 case, illustrated in Figure 43b and the scatter plot in Figure 43c in blue, compared to the charge density distribution for Lo_10 shown in Figure 43a and Figure 43c. The effect of the low electric Reynolds number in these simulations is evident in the prevailing electric charge in the liquid bulk throughout the domain with some migration to the outer surface of the jet.

Conclusions

A robust computational framework for simulating electrically charged liquid hydrocarbon jets is developed, tested and demonstrated in this work. First-principles-based numerical methods developed specifically for high-fidelity direct numerical simulations of electrohydrodynamic (EHD) flows are implemented within a conservative flow solver using an unsplit, geometric volume-of-fluid transport scheme that includes EHD effects. The numerical framework conserves mass, momentum, and the electric charge density even at the gas-liquid interface where discontinuities exist. A novel approach using a separate mesh is used to obtain accurate boundary conditions of electric potential on the computational domain. Simulations employ a recently developed physics extraction tool that tracks every breakup and coalescence event occurring during an atomizing spray. Data characterizing the atomization processes are stored in a Neo4j graphical database providing an easily accessible format. The framework demonstrated in this work confirms the robustness of the methodology and its tools. The efficient methods for simulating complex atomization of electrically charged jets enable much improved understanding of the effect of the electric field, the interplay between electric and hydrodynamic effects and their influence on atomization processes, and spray characteristics that have proven challenging to obtain experimentally.

Statistics for droplet size, breakup and coalescence events, onset of primary atomization, and spray cone angle are reported for all simulations revealing earlier and enhanced liquid breakup with increasing electric charge. The presence of electric charge de-stabilizes the liquid core, thereby advancing the onset of primary atomization, and increases spray angle and droplet spacing due to the effect of the electric force acting radially outward from the liquid core. Subsequent to initial breakup, a reduction in the effective surface tension in liquid structures due to the presence of electric charge enhances secondary atomization breakup processes in the electrified jet. The effects of charge mobility are evident, resulting from mutual repulsion of the unipolar charge with in the liquid and influencing the charge density distribution in droplets and structures dislodged during primary atomization.

The work reported here is an initial step toward a much broader application of DNS for atomizing EHD flows. The simulations reported herein demonstrate extensions to recently developed tools—a physics extraction scheme and the e-Mesh domain solver for electric boundary conditions—to a complex, multi-physics, multi-phase atomization scenario, thereby validating their robustness and the utility of the framework. The integration of these new tools within NGA improves the efficiency and accuracy of EHD simulations and significantly expands the analysis opportunities of primary atomization processes for these types of flows. Analysis of secondary atomization processes and mechanisms is a natural extension of the work reported here. Moreover, continuation of this work will permit the development of DNS databases to support future efforts, both experimental and computational. The methodologies developed in previous work [24, 93, 172, 203, 204] and demonstrated here open new, simulations-based avenues of exploration within a broader category of electrohydrodynamics.

Acknowledgments

This work was supported by Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS), National Science Foundation, under a Discovery project allocation. Simulations were performed on Purdue University's Anvil cluster through ACCESS and on the Tempest High Performance Computing System, operated and supported by University Information Technology Research Cyberinfrastructure at Montana State University. The second author gratefully acknowledges the support provided by the Department of Aerospace Engineering and College of Engineering, Texas A&M University.

A GEOMETRIC AND TOPOLOGICAL DATA ANALYSIS OF ATOMIZATION FOR MODEL DEVELOPMENT

Contribution of Authors and Co-Authors

Manuscript in following chapter

Author: Brendan Christensen

Contributions: Methodology, Data Curation, Formal Analysis, Investigation, Writing - Original Draft, Visualization

Author: Jack Ruder

Contributions: Formal Analysis, Methodology, Data Curation

Author: Alex McCleary

Contributions: Methodology, Data Curation

Author: Brittany Terese Fasy

Contributions: Supervision

Author: Mark Owkes

Contributions: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding Acquisition

Manuscript Information

Brendan Christensen, Jack Ruder, Mark Owkes, Alex McCleary, Brittany Terese Fasy

Journal of Computational Physics

Status of Manuscript:
_X_ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Elsevier Ltd.

Abstract

High-fidelity simulations of atomization present a promising approach to studying its complex physics. However, these simulations face significant challenges, often requiring hundreds of processors and generating tens to hundreds of terabytes of data, complicating their practical application. This work presents a method to extract localized geometries of droplets and their surrounding flow fields at the moment of breakup within a simulation, resulting in a comprehensive database of atomization statistics. This improves the usefulness of simulation results by storing key information on the atomization process. Parsing and comparing these geometries remains challenging, so we apply a topological data analysis tool—the directionally integrated weighted Euler characteristic transform (DIWECT)—to produce vectorized representations of the complex data. These vectors are well-suited for advanced data processing, machine-learning applications, and the development of reduced-order models. We demonstrate the effectiveness of this approach through K-means clustering of the resultant vectors and analyze the clusters to confirm that the DIWECT captures meaningful physical information in an abstract, reduced-order format. This work represents a novel, efficient, and physically relevant method with which to extract, parse, and analyze large-scale atomization simulation data with an eye toward reduced-order model development.

Introduction

Atomization—the process where a coherent liquid body breaks into smaller droplets—is a wide-ranging phenomenon with applications in numerous industries and environmental applications. From fuel injection and fire suppression to agricultural spraying and air-sea interactions, a detailed understanding of the underlying physics is essential for optimizing engineering designs and improving predictive models of natural phenomena. Despite its

widespread importance, many of the underlying physical mechanisms that drive atomization remain unresolved. There remains significant research in the field attempting to answer the fundamental questions: when and how do liquids breakup?

Experimental studies on atomization have long faced limitations due to the inherent complexity of capturing high-speed, multiphase flows. Simplified jet flows can be captured well using traditional high-speed imaging techniques [119]. However, as the Weber and Reynolds numbers increase to values relevant for many practical applications, experimental analyses of the physics become intractable because of the formation of an opaque cloud of droplets that forms around the spray. Even with recent advances in imaging techniques like ballistic [113, 168, 181] and x-ray imaging [63, 108, 210], capturing temporally continuous, quantitative data remains problematic. As a result, experimental methods have struggled to provide sufficient insight into the small-scale instabilities and breakup mechanics that occur in atomizing systems.

In response to these challenges, high-fidelity simulations—commonly known as direct numerical simulations (DNS)—have emerged as a powerful alternative [56] capable of resolving the detailed physics of atomization. High-fidelity simulations continue to produce increasingly accurate representations of the complex multiphase process, providing researchers with a tool to predict quantitative, three-dimensional, and temporally granular descriptions of liquid breakup. However, these simulations are not without their own limitations. As computational power has increased, the size of datasets generated from atomization simulations has grown dramatically. High-resolution simulations, capable of producing hundreds of terabytes of data, present immense challenges in both storage and analysis. Extracting meaningful physics from such massive datasets remains an ongoing challenge, and current methodologies are ill-equipped to fully utilize this wealth of information.

To address these limitations, the Atomization Simulation Statistics Extraction Tool

(ASSET) [24, 159] was developed to extract detailed event-based data from large-scale simulations of atomizing liquids. Unlike traditional post-processing methods that are constrained by the sheer volume of data, ASSET works coincident with a simulation to identify and record breakup and coalescence events as they occur. This approach allows for the creation of comprehensive databases that contain statistics, geometries of liquid structures, and the corresponding flow fields at the moment of each breakup event.

Furthermore, ASSET tracks the evolution of each liquid structure, allowing for the reconstruction of breakup *pathways*—the sequential processes by which liquid structures fragment into progressively smaller droplets. These pathways reveal how individual droplets emerge from larger structures, shedding light on the mechanisms that govern spray formation. By capturing these data directly from simulations, researchers gain an unprecedented level of access to temporally continuous, localized information essential for studying the full physics of atomization.

High-fidelity simulations are incredibly costly, often requiring hundreds of thousands of CPU hours on high-performance computers. Thus, predictive atomization models play a vital role, especially for practical applications where fully resolving the physics is not feasible. Reduced-order models such as Reitz-Diwakar (RD) [154], Taylor Analogy Breakup (TAB) [134], and Pilch-Erdman (PE) [145] models offer relatively fast and inexpensive ways to predict atomization behavior in specific systems. However, these models are heavily based on simplified assumptions and empirical data from idealized conditions, such as single droplets in controlled environments. As a result, they often fail to accurately represent the complex, multiscale behavior of atomization in real-world, turbulent systems. There is an increasing need for models that are informed by statistically relevant physics-based data extracted directly from simulations, enabling more accurate predictions of atomization dynamics across a broader range of conditions.

To meet this need, an essential step forward involves not only extracting more

comprehensive data from simulations but also transforming these data into a format suitable for modern analysis techniques. Concretely, it is necessary to efficiently query and parse databases full of data from physical simulations. The present work proposes using a geometric and topological data analysis tool known as the directionally integrated weighted Euler characteristic transform (DIWECT) to convert complex geometric and flow field data from atomization simulations into vectorized representations. This approach structures the data in a format well suited for machine learning and other data-driven techniques, enabling the development of reduced-order models that agree with the physics of atomization predicted by high-fidelity simulations. The resulting structured database enables efficient querying and analysis, supporting the development of reduced-order models that retain the underlying physics of atomization rather than relying on significant assumptions.

This paper focuses on demonstrating the process of generating these databases and applying DIWECT to enable future data-driven predictive models. Furthermore, the DIWECT provides a powerful framework for novel analyses, enabling the identification of previously unrecognized patterns within these large-scale, complex systems. In doing so, we aim to bridge the gap between the vast amounts of data produced by high-fidelity atomization simulations and the limited models currently used, advancing toward a more accurate and physically informed understanding of atomization.

Background - Topological Data Analysis

Modern experimental and numerical methods in all range of scientific fields are producing increasingly complex and large-scale data, thus necessitating efficient methods to analyze and visualize them. Topological data analysis (TDA) methods have emerged as a family of tools for summarizing and analyzing data in a mathematically robust and rigorous manner. For a comprehensive background on the various methodologies see *Topological Methods in Data Analysis and Visualization VI* [73]. In the context of fluid mechanics,

where massive datasets are a common challenge, TDA offers the ability to capture essential geometric components of complex fields.

There exist a few studies incorporating TDA into fluids [1, 53, 92, 103, 130, 133]. Scalar fields in fluid mechanics, such as pressure, temperature, and image pixel intensity can be analyzed using the persistent homology TDA method. These analyses have been used to identify large-scale symmetries, transition regions, and patterns in the flow [91, 158]. Similarly, studies involving vector fields, such as velocity, have leveraged topological descriptors to classify coherent flow features and reveal interactions within turbulent flows [60, 61]. While these applications illustrate the versatility of TDA in fluid mechanics, they are focused on scalar field data and large-scale flow dynamics rather than capturing the small scale, underlying physics of liquid breakup studied in this work.

In contrast, the geometric analysis of three-dimensional liquid structures in multiphase flows through TDA remains a relatively underexplored area. Chen et al. [21] introduced a skeletonization technique to capture the fundamental topology of periodic liquid jets during atomization. This method allowed for the identification of branching points in the jet structure and offered new insights into the mechanisms driving liquid breakup. To the authors' best knowledge, this is the only topological analysis of liquid structure geometries in multiphase flows. Beyond fluid mechanics, TDA has been successfully employed in fields such as medical imaging [28, 62, 148], 3D object recognition [67], and structural biology [201] to quantify, match, and analyze complex shapes. These examples demonstrate the versatility of geometric TDA in capturing intricate structural information and motivate its application to fluid systems.

Building on these developments, this work focuses on analyzing three-dimensional geometries associated with liquid structures during atomization. Concretely, we use the DIWECT—a tool within TDA—for this purpose. Unlike previous methods that emphasize whole flow fields, the present work applies TDA to isolated liquid breakup events, providing

concise, representative descriptions of complex geometries.

Background - Atomization Modeling

One of the primary objectives of this work is to provide physical data that can inform reduced-order atomization models. Achieving this requires an understanding of the information needed and the limitations of current modeling approaches. This section focuses on the role of Lagrangian droplet tracking models, as the data extracted from this study is particularly suited to their improvement.

The Lagrangian-Eulerian (LE) technique is a method used to model atomization with reduced computational cost. It represents the liquid phase as discrete “blobs" or droplets in a Lagrangian framework, while the gas phase is treated in an Eulerian framework. Prominent implementations, such as the KIVA series [3–6, 199], have been widely adopted in industry [202]. Following injection, the breakup of blobs is typically modeled using idealized instability mechanisms, such as Kelvin-Helmholtz and Rayleigh-Taylor instabilities [11, 142, 154]. Further breakup is often predicted using models like the Taylor Analogy Breakup (TAB) [134] or Pilch-Erdman (PE) [145]. These methods offer computational efficiency but rely heavily on simplifying assumptions, such as spherical droplet shapes and empirical correlations from isolated experiments.

These simplifications fail to capture the multiscale and turbulent nature of real-world atomization processes. Experiments and DNS have revealed that atomization involves complex deformation and breakup mechanisms, including bag and sheet formation, ligament destabilization, and droplet fragmentation [81, 119, 222]. These mechanisms lead to highly variable, complex droplet and flow field geometries that are far beyond the capabilities of traditional blob injection models.

Recently, there has been considerable development on the hybrid Volume-of-Fluid Lagrangian-Eulerian (VOFLE) method [42, 65, 98, 197, 225]. These models combines the

Volume-of-Fluid (VOF) technique to resolve near field jet deformation and large-scale liquid structures before transitioning to an Lagrangian description for smaller droplets (Note: Level-set (LS) is also used for interface capturing in some cases [225]). Varying local criteria are used to determine when a droplet meets the requirements for Lagrangian treatment, with most opting for a size threshold for the drop diameter, and others additionally considering the sphericity of the droplet. Only two papers have considered further breakup to the author's best knowledge. These are Yu et al. [219], who consider the RT instability for breakup and Kuo & Trujillo [98], who use the maximum entropy formalism (MEF) model. VOFLE balances computational efficiency and accuracy with an improved resolution of the early stages of jet formation and breakup.

Despite these advancements, VOFLE models face challenges associated with empirical assumptions and computational trade-offs. For example, most methods require that droplets be spherical and/or below a size threshhold before transitioning to a Lagrangian representation. This increases the necessary size and cost of the Eulerian domain. Additionally, secondary atomization is largely neglected, and when it is considered, secondary atomization models are significantly limited.

This work lays the foundation for a predictive VOFLE methodology, informed by high-resolution simulations. By extracting explicit breakup cascades, droplets can be transformed to Lagrangian structures at larger and less spherical sizes and their subsequent breakup can be more accurately predicted. The envisioned approach would identify liquid structures as they detach, compare them to a database of high-fidelity DNS results, and use stored breakup pathways to inform how a given structure will fragment. While this study does not implement the full VOFLE model, it establishes essential techniques for extracting droplet geometries, surrounding flow fields, and breakup genealogies—key steps toward improving atomization modeling.

Project Outline and Scope

This work seeks to establish a framework that bridges the gap between high-resolution simulations and reduced-order modeling of atomization processes. The ultimate goal is to use the data and processing techniques outlined in this paper to inform reduced-order atomization models. However, the development of a model is beyond the scope of the present work. The current project is structured into two key objectives:

1. **Use ASSET to extract atomization geometries:** The ASSET methodology extracts the gas-liquid interface of liquid structures, along with their surrounding flow fields, from high-fidelity simulations during breakup and coalescence events. These structures and their interactions are stored in a graph database, allowing researchers to analyze the sequence of breakup events—known as breakup pathways—that lead to the formation of smaller droplets.

2. **Vectorize the resultant databases using the DIWECT:** The geometric and flow field data extracted using ASSET are transformed into vectorized representations with the DIWECT. This process prepares the data for machine learning applications, allowing for statistical analysis and efficient shape-matching. By capturing key geometric and topological features, the DIWECT helps identify patterns and relationships within the data.

This work provides a novel approach to analyze large atomization simulation databases, yielding information necessary to inform a physics-based reduced-order model. Explicitly, it is important to analyze the DIWECT as a method for vectorizing data in a *useful* way—where useful refers to the ability of these representations to serve as effective inputs to machine learning algorithms. An ideal vectorization must preserve the essential geometric and topological characteristics of complex flow structures while reducing them to a comparable

format suitable for statistical analysis and training. This ensures that meaningful similarities and differences between structures are captured in a way that algorithms can learn from and generalize across. This quality will provide both novel, in-depth data-driven analyses into atomization as well as the potential to use these insights for predictive modeling.

Methods

Computational Platform

The advancements outlined below are built into the NGA computational platform [33, 35, 138, 139]. This platform is then used to perform the atomization simulation outlined below. NGA solves the two-phase, incompressible Navier-Stokes equations for mass and momentum conservation, expressed as

$$\frac{\partial \rho_\phi}{\partial t} + \nabla \cdot (\rho_\phi \boldsymbol{u}_\phi) = 0 \tag{42}$$

and

$$\frac{\partial \rho_\phi \boldsymbol{u}_\phi}{\partial t} + \nabla \cdot (\rho_\phi \boldsymbol{u}_\phi \otimes \boldsymbol{u}_\phi) = -\nabla p_\phi + \nabla \cdot (\mu_\phi [\nabla \boldsymbol{u}_\phi + \nabla \boldsymbol{u}_\phi^T]) + \rho_\phi \boldsymbol{g}, \tag{43}$$

where ρ_ϕ is the density, $\boldsymbol{u}_\phi$ the velocity vector, p_ϕ the pressure, μ_ϕ the dynamic viscosity, t the time, $\boldsymbol{g}$ the gravitational acceleration vector, and ϕ the phase indicator (g for gas and l for liquid). These equations are discretized on a staggered Cartesian grid, pressure stored at cell centers and velocity components at cell faces.

Numerical methods include:

- Time discretization: An iterative Crank-Nicolson scheme, with a semi-implicit correction [23]
- Gas-liquid interface tracking: A geometric volume-of-fluid (VOF) method with unsplit semi-Lagrangian fluxes [138]. The interface is reconstructed using piecewise

linear interface calculation (PLIC) [218] with efficient least-squares VoF interface reconstruction algorithm (ELVIRA) for normal vector calculations [146] on subcells to ensure mass and momentum conservation [139].

- Advection: Away from the interface, high-order finite differences for mass, momentum, and scalar transport [33] are used. Near the interface the semi-Lagrangian VOF method is used [138].

- Pressure Solver: The Poisson equation is solved with the ghost fluid method to incorporate the surface tension force [46] and a black-box solver [30].

- Surface Tension: Curvature is computed using the adjustable curvature evaluation scale (ACES) method [136].

NGA is fully parallelized via MPI, enabling efficient scaling to tens of thousands of cores [32]. Its computational efficiency and robust numerical methods have made it an effective tool for studying atomization and related multiphase flow phenomena [34, 95, 138, 171].

ASSET Methodology

The first goal in this project is to extract liquid structure and flow field geometries from high-fidelity simulations. To meet this goal, it is necessary to identify key moments and extract only those structures that provide model-relevant data. The breakup process is important because it describes the formation of droplets. However, breakup represents only part of the dynamics in large-scale atomization. Coalescence is also common in dense spray clouds. Thus, tracking both breakup and coalescence is necessary to fully describe the system.

In addition, liquid structures often deform significantly between breakup or coalescence events. These deformations provide information on the mechanisms leading to breakup, such

as an elongating ligament undergoing Plateau-Rayleigh instability or a droplet flattening before undergoing bag-type breakup.

The Atomization Simulation Statistics Extraction Tool (ASSET) is designed to capture this information by identifying and isolating breakup, coalescence, and intermediate (update) events. ASSET is an event-triggered, parallelized algorithm that operates concurrently with simulations. It runs each timestep and efficiently identifies breakup and coalescence events, gathering physical data associated with each occurrence. A comprehensive description of the algorithm is given in the original works [24, 159], however, a concise summary is included here. As we detail below, the key addition in this work is the ability to extract topologically sound geometries and to incorporate update events for more comprehensive tracking of liquid structure evolution.

The workflow for data extraction involves three key steps: event identification, geometry extraction, and velocity calculation on the extracted meshes. 1) Event identification works by monitoring changes in identification numbers, providing liquid structure history and breakup pathways. 2) Once an event is detected, the geometries of the affected liquid structures are extracted by isolating the computational cells involved in the event. To generate a watertight surface representation, a marching cubes algorithm is applied, ensuring topological accuracy and avoiding artifacts that could interfere with subsequent analyses. 3) Finally, rather than storing the full velocity field, only the velocity component normal to the liquid surface is recorded, as this component is most relevant to droplet deformation. Details on each step are provided below.

Event Identification Event identification in ASSET—described in detail by Christensen and Owkes [24]—begins by tagging liquid structures with two unique integers: $\mathcal{S}$ and $\mathcal{L}$. $\mathcal{S}$ is assigned at every timestep using a band-growth algorithm [65], distinguishing independent liquid structures within the domain. $\mathcal{L}$ is persistent across timesteps and is transported with

the liquid, only changing when breakup or coalescence occurs. $\mathcal{L}$ provides a history of the liquid, allowing us to track the breakup pathway of any given liquid structure.

The event identification algorithm distinguishes the three types of events as follows:

- **Breakup Events:** Occur when a single liquid structure (same $\mathcal{L}$) splits into multiple, structures (different $\mathcal{S}$) in non-adjacent grid cells.
- **Coalescence Events:** Identified when two or more structures with distinct $\mathcal{L}$ values flux into a single grid cell.
- **Update Events:** Triggered periodically for each structure based on a user-defined time interval relative to the last event the structure was involved in.

The original method described the extraction of statistics when an event was identified. This utility is still applied. These data are sampled both immediately preceding and following an event, providing a record of the conditions leading to and resulting from each event. Data are stored in a Neo4j graph database (see Figure 44), enabling relationships between parent and child droplets to be established and liquid structures to be tracked back to the initial coherent liquid at the start of the simulation.

Geometry Extraction When an event is identified, the region around the liquid structure(s) in the event is isolated. In fact, this is already done during the assignment of $\mathcal{S}$ through a band-growth algorithm [65], providing access to the computational cells of interest. With the cells isolated, the liquid-gas interface geometry is captured. Ensuring no erroneous topological features are captured in this step is necessary because the DIWECT methodology is highly sensitive to them. A naive implementation would be to output the PLIC representation of the interface, which is already calculated. However, as seen in Figure 45a, the edges and vertices of the PLIC interface are not aligned to produce a watertight surface, introducing non-physical holes. To ensure topological soundness, we

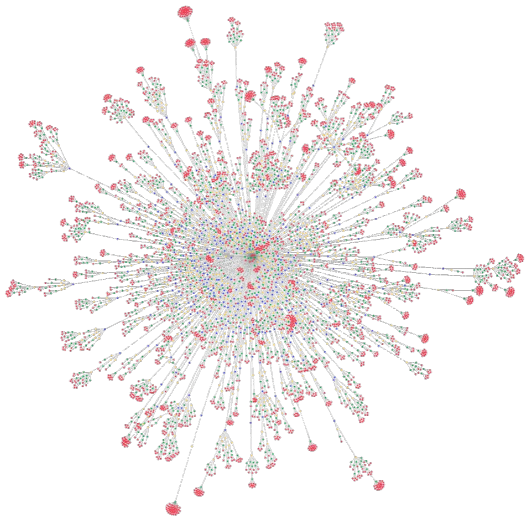
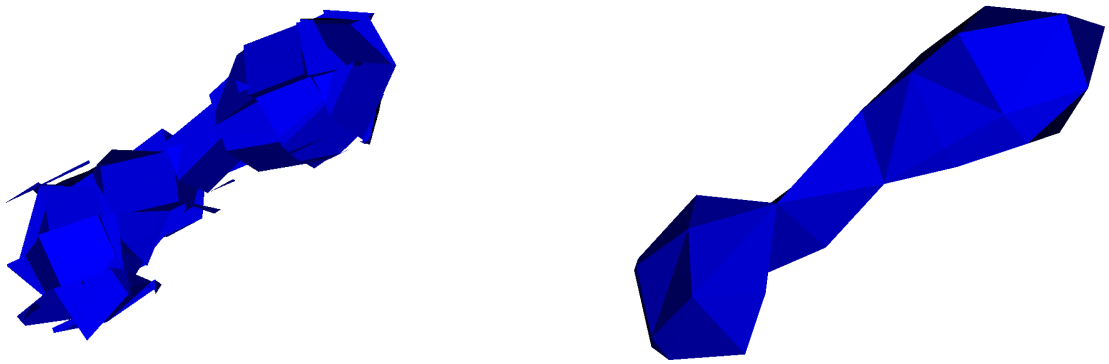


Figure 44: Neo4j graph database representing data extracted from a numerical simulation of a diesel-type jet from [24]. Dots represent droplets with lines connecting related droplets (i.e parent and child droplets). Colors represent the number of breakups taken to get to a given droplet.

implemented a formulation of the commonly used marching cubes algorithm [116] that is topologically robust [104]. This method generates a set of triangles representing the interface. The vertices and connectivity information of these triangles are stored as an unstructured mesh in the SILO format [22]. In this format, each triangle is represented as a zone, and scalar values can be associated with these zones. The marching cubes algorithm provides a topologically sound representation of the droplet, however, it does not ensure a fully accurate volume, whereas the PLIC interface is volume conserving, which is necessary for the geometric volume-of-fluid numerical methods. Producing a volume-conserving, water-tight representation and preventing the calculation of two separate interface reconstructions is an area of future work.

Velocity Calculation Saving the full velocity field around each droplet would present a significant storage requirement. To reduce the total data being saved, the flow field is simplified to include only the velocity component normal to the interface. This normal



(a) Interface geometry of a droplet computed with traditional PLIC method.

(b) Interface geometry of a droplet computed with enhanced marching cubes algorithm [104].

Figure 45: Comparison of the volume conserving PLIC interface and the topologically sound, watertight marching cubes interface.

component, represented mathematically as $\mathbf{u} \cdot \mathbf{n}$, where $\mathbf{u}$ is the velocity vector and $\mathbf{n}$ is the unit normal vector of the interface plane, is most prominently responsible for droplet deformation.

Since $\mathbf{u} \cdot \mathbf{n}$ is a scalar quantity, it can be efficiently stored in the SILO file for each triangle zone. With the orientation of the interface plane and the velocity magnitude, the interaction between the velocity and the shape of the fluid is efficiently stored together. The simplified data can then be visualized and analyzed using tools such as VisIt [22], enabling detailed studies of droplet dynamics.

The Directionally Integrated Weighted Euler Characteristic Transform (DIWECT)

With droplet geometries and their associated normal velocities now extracted and stored efficiently in a compact, structured format, the next objective is to develop a means of comparing these complex shapes and dynamics in a consistent, scalable manner. Specifically, we require a vectorized representation that captures the essential geometric and flow-related features of each droplet, enabling fast and meaningful comparison across a large database of breakup events. The Directionally Integrated Weighted Euler Characteristic Transform

(DIWECT) is introduced here as a method to provide exactly this vectorization, translating high-dimensional surface and velocity data into forms suitable to statistical analysis and machine learning.

This section details the methodology of the DIWECT. This method is an extension of the weighted Euler characteristic transform (WECT). For a more in-depth background on the WECT, see Jiang et al. [84] and Cisewski-Kehe et al. [26].

The foundation of this method are simplicial complexes. Explicitly, for a set of vertices $\mathcal{V} \subseteq \mathbb{R}^n$, a *simplex* is the convex hull of an affinely independent subset of $\mathcal{V}$. Given any simplex σ, we define its dimension $\dim \sigma$ as the number of vertices in σ minus one (e.g. a zero-simplex is a vertex, a one-simplex is an edge, a two-simplex is a triangle, etc.). A face τ of a simplex σ is any subset of its vertices ($\tau \leq \sigma$), while σ is a coface of τ ($\sigma \geq \tau$). A *geometric simplicial complex* is a finite collection of simplices K satisfying:

1. If $\sigma \in K$, then all faces $\tau \leq \sigma$ are also in K
2. If any two simplices $\sigma, \sigma' \in K$, have a nonempty intersection, then $\sigma \cap \sigma'$ is a face of both σ and σ'

A *weighted* simplicial complex (K, ω) is defined by a simplicial complex K and a weight function $\omega : K \to \mathbb{R}$, that assigns a weight to each simplex in K. The unstructured meshes extracted in §5 are simplicial complexes with the normal velocities calculated in §5 as weights on each simplex. Weights for the edges and vertices are calculated by averaging the weights of their respective co-faces.

The Euler characteristic of a simplicial complex is given as

$$\chi(K) = \sum_{\sigma \in K} (-1)^{\dim \sigma}. \tag{44}$$

For example, for three-dimensional polyhedra ($\dim 3$), the Euler Characteristic is defined

$\chi = V - E + F$ where V, E, and F are defined vertices, edges, and faces, respectively. For example, the Euler characteristic of any Platonic polyhedra (e.g. tetrahedron, cube, octahedron, etc.) is two. The Euler characteristic is a topological invariant, meaning it remains unchanged under continuous deformations such as bending or stretching, as long as there is no tearing or gluing.

The Euler characteristic is easily extended to the *weighted Euler characteristic* (χ_ω) for weighted simplicial complexes, defined as

$$\chi_\omega(K, \omega) = \sum_{\sigma \in K} \omega(\sigma) \cdot (-1)^{\dim \sigma}. \tag{45}$$

The weighted Euler characteristic alone is a limited representation of any complex, because it reduces all information to a single scalar value. However, by computing it across a sequence of subcomplexes, we can retain more information from the data. This is done using a *filtration*, which is a nested sequence of weighted simplicial complexes parameterized by $\mathbb{R}$. Given a filtration $\{K_t\}_{t \in \mathbb{R}}$, it satisfies

$$K_s \subseteq K_t \text{ for all } s \leq t. \tag{46}$$

This method specifically uses a *height filtration*, which orders simplices based on their vertex positions in a given direction $d \in \mathbb{D}^{n-1}$. The height filtration of a complex $K_{d,t}$ is defined as

$$K_{d,t} := \{\sigma \in K \mid h_d(\sigma) \leq t\}, \tag{47}$$

where $h_d(\sigma)$ is the maximum height of any vertex in σ along direction d, which is computed using the standard inner product of $\mathbb{R}^n$. In the WECT, we take the height filtration in multiple directions, thus providing a structured and mathematically rigorous way to represent complex geometries as vector-valued quantities. See Figure 46 for a demonstration of this

process.

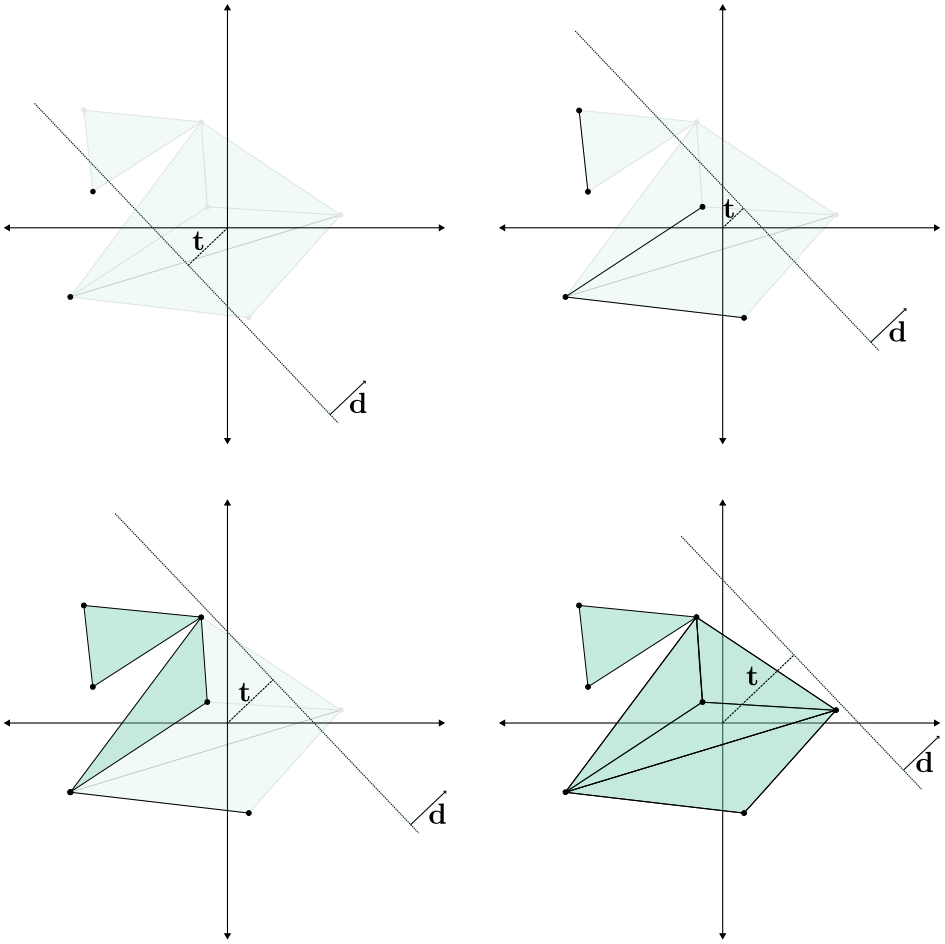


Figure 46: Height filtration in direction d. Simplices become bold when they are below the plane and contribute to the Euler characteristic for that height t.

By computing the Euler characteristic over a height filtration $\{K_d\}_{t\in\mathbb{R}}$ of a weighted simplicial complex (K,ω) for a direction $d \in \mathbb{D}^{n-1}$, we obtain the *weighted Euler characteristic function* (WECF : $\mathbb{R} \to \mathbb{R}$). This function maps a direction d and a height t to the weighted Euler characteristic of the corresponding subcomplex $K_{d,t}$. Formally

$$\mathrm{WECF}_{(K,\omega),d}(t) := \chi_\omega(K_{d,t}, \omega). \tag{48}$$

Finally, computing the weighted Euler characteristic function for all directions $d \in \mathbb{D}^{n-1}$,

we define the *weighted Euler characteristic transform* ($\mathrm{WECT} : \mathbb{D}^{n-1} \times \mathbb{R} \to \mathbb{R}$) as the function

$$\mathrm{WECT}_{(K,\omega)}(d,t) := \mathrm{WECF}_{(K,\omega),d}(t). \tag{49}$$

Theoretically, the WECT fully encodes the shape of the simplicial complex K, when computed over all directions d and heights t. In practice, the WECT is computed for a finite sampling of directions and heights. Nevertheless, with sufficiently dense sampling, the WECT retains enough information to uniquely characterize a given simplicial complex [26, 84].

The WECT is sensitive to rotations and reflections of the simplicial complex. To address this limitation, we introduce the *directionally integrated* WECT, or DIWECT which integrates the WECT over all rotations and reflections of the simplicial complex.

For a uniformly sampled set of directions $\mathcal{D} \in \mathbb{D}^{n-1}$, the DIWECT can be estimated as a directional average of the WECT and computed as

$$\mathrm{DIWECT}_{(K,\omega)}(t) = \frac{1}{|\mathcal{D}|} \sum_{d \in \mathcal{D}} \mathrm{WECT}_{(K,\omega)}(d,t). \tag{50}$$

In practice, we are not able to compute the WECT over all directions and heights, so we use an estimated form of the DIWECT, and the theoretical guarantees are necessarily limited by this discrete sampling. For this study, we selected 150 directions and 75 heights to balance two competing objectives: preserving fidelity of the droplet representation and minimizing storage requirements. As shown in Equation 50, the directional sampling is averaged over each height, yielding a 75-dimensional vector per droplet. This dimensionality was chosen to ensure that the resulting representation remains compact and do not increase the storage requirement for small droplets. For example, an octahedron, which is among the most simplified and lowest-resolution shapes extracted in our dataset, requires approximately 50 integers to represent its vertices, faces, and associated velocity components. Most

droplets in the dataset are significantly more complex than an octahedron, so a 75-element vector does not substantially increase the storage burden for the majority of cases, while still providing a decently high-resolution description of shape variation across directions. Moreover, since the DIWECT of a single orientation is sufficient to represent any rotated or reflected transformation of a droplet, we avoid the need to store multiple transformed variants, effectively reducing the total storage cost by a factor of $\mathcal{D}$, the number of sampled directions.

Algorithm 1: Methodology Summary

Run Simulation and Extract Data with ASSET
while simulation is running **do**
    Assign initial $\mathcal{S}$ and $\mathcal{L}$ identifiers in the domain
    for each timestep **do**
        Reassign $\mathcal{S}$ for each liquid structure
        Transport $\mathcal{L}$ with the liquid

        Event Identification:
        if Multiple $\mathcal{S}$ associated with one $\mathcal{L}$ **then**
            Breakup event identified
        end if
        if Multiple $\mathcal{L}$ in one cell **then**
            Coalescence event identified
        end if
        for each structure **do**
            if (Current time % update frequency = 0) **then**
                Update event identified
            end if
        end for

        Geometry Extraction:
        for each identified event **do**
            Extract the gas–liquid interface via marching cubes
            for each interface cell **do**
                Velocity Calculation:
                Compute $\mathbf{u} \cdot \mathbf{n}$, on each triangle of the mesh
            end for
            Save as unstructured mesh with normal velocities
        end for
    end for
end while

Post-Processing: Estimate the DIWECT
Aggregate all extracted data
for each unstructured mesh χ **do**
    Estimate the DIWECT of χ
end for
Output: A vectorized representation of liquid structure geometries from atomization

Results

The results presented here are derived from $\approx 120,000$ droplets extracted from a high-fidelity simulation run. The entire simulation database, made up of a full rendering of each timestep requires $\approx$ 10TB of storage space. The total storage requirement of the unstructured SILO meshes was 133GB. The full WECT was calculated with 150 directions and 75 heights and only required 1.5 GB of storage. The averaging step was conducted over each direction and the DIWECT representation of the same dataset only required 10 MB—a reduction of over 13,000-fold from the SILO meshes and 1,000,000-fold from the original simulation dataset.

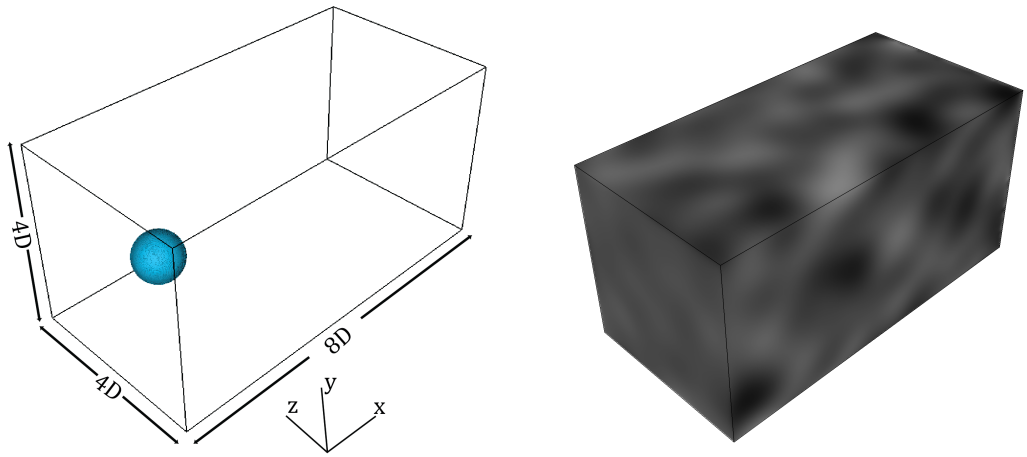
The following subsections provide proof of concept of the DIWECT as a method to uniquely and efficiently characterize liquid shapes and surrounding flow fields. First, we match droplets to their most similar counterparts using only their DIWECT values. Then, we show that the droplets can be clustered into groups which display physical relevance based on their DIWECTs.

Simulation Setup

To test the methodology, we ran a large-scale numerical simulation of an atomizing water droplet (Fig. 47a). The methods outlined in this work are designed to be scalable and capable of processing large numbers of droplets, as would be expected in most practical atomization applications. The setup in this work consists of a high-speed gas inflow at the x^- wall, which is added to an initialized homogeneous isotropic turbulence (HIT) domain and inflow, as shown in Figure 47b. This simulation ran for one month on 128 processors on the Tempest HPC at Montana State University. To reduce the size of the resultant dataset, only droplets from the second half of the simulation were analyzed.

This test case represents a common atomization scenario with broad applications, in-

cluding aerial firefighting, agricultural spraying, industrial spray cooling, and environmental aerosols. The chosen resolution balances computational efficiency with sufficient detail to capture smaller-scale droplets. The key non-dimensional parameters for this simulation are listed in Table 8.



(a) Simulation domain with initial droplet in place (D is droplet diameter).

(b) HIT-derived turbulence initialized in the domain.

Figure 47: Simulation domain setup.

Table 8: Non-dimensional parameters for the simulation. D is initial droplet diameter, dx is grid spacing, ρ_g is gas density, U_g is the bulk gas velocity, μ_l is liquid viscosity, σ is the surface tension coefficient

Parameters	Description	Value
Cells across diameter	D/dx	150
Bulk Reynolds	$\rho_g U_g D/\mu_l$	4.74×10^4
Bulk Weber	$\rho_g U_g^2 D/\sigma$	825

Droplet Matching

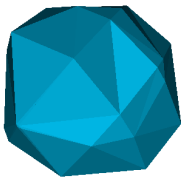
A key objective of this work is to establish a basis for efficiently comparing droplets and their local flow fields across large simulation databases. In a future scenario, a user might extract a particular droplet of interest from an ongoing simulation and automatically search a previously computed database for the most similar droplet. Then, predictions can be made from the previously simulated droplet to inform how the droplet of interest might behave.

For demonstration, we selected a set of four droplets from a separate simulation run. These droplets were pulled from another water drop atomization case, but with a larger initial diameter. Four common shapes from the database were chosen. Each droplet's DIWECT signature was then compared against the database developed from the simulation in §5 to find the most similar entries according to Euclidean distance in DIWECT-space. These results can be seen in Figures 48-51, which show the reference droplet on the left, and the closest matched droplets on the right. AR is the aspect ratio, defined as the shortest principal axis divided by the longest, and V, which is the volume of the droplet.

The matching results show very nice qualitative agreement for the shapes. There exists some variation between the shape parameters, but it is important to note that there are many interacting factors that lead to the final DIWECT calculation, notably, the DIWECT captures information about the normal velocity components. Thus, comparing a simple shape parameter and the size of droplets is a limited assessment of the DIWECT's capability of to compare similar structures.

Clustering Results

To further demonstrate the utility of the DIWECT, we apply clustering to the resulting vectors using Euclidean distance as a measure of similarity. By comparing the resulting clusters with known physical parameters, we assess how effectively the DIWECT preserves



(a) Droplet 1: AR = 0.868, $V = 3.3 \times 10^{-12}$

(b) AR = 0.703, $V = 5.22 \times 10^{-12}$

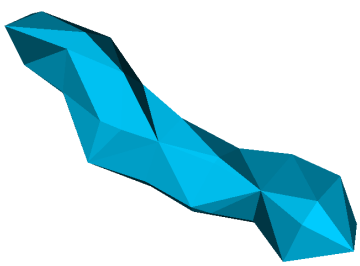
(c) AR = 0.782, $V = 2.99 \times 10^{-12}$

(d) AR = 0.755, $V = 5.74 \times 10^{-12}$

(e) AR = 0.780, $V = 1.44 \times 10^{-11}$

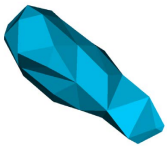
Figure 48: Droplet 1 and the four closest matches based on DIWECT

relevant physical information. Given the large size of the dataset ($\approx 120,000$ samples), computational efficiency is a key consideration. To ensure efficiency of our analysis at scale, we consider methods better than $O(n^2)$ where n is the number of samples. Another consideration is the dimensionality of the data. The discretization produces a database of 75-dimensional vectors, so the dimensionality must be considered to avoid sparse sampling in high-dimensional space, otherwise known as the curse of dimensionality. Projecting the data into a lower dimension improves the capability of non-parametric methods and enhances computational efficiency. Principal Component Analysis (PCA) is therefore applied to the dataset before clustering. We found that 20 components was sufficient to explain the variability of our data (See Figure 52). To determine the appropriate number of PCA components, we used the elbow method, which identifies the point beyond which additional components contribute diminishing returns in explained variance. Following the application

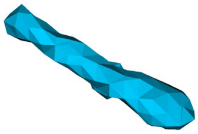


(a) Droplet 2: AR = 0.304, $V = 4 \times 10^{-12}$

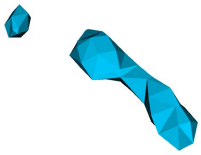
(b) AR = 0.405, $V = 7.6 \times 10^{-12}$



(c) AR = 0.478, $V = 7.8 \times 10^{-12}$



(d) AR = 0.213, $V = 2.7 \times 10^{-11}$



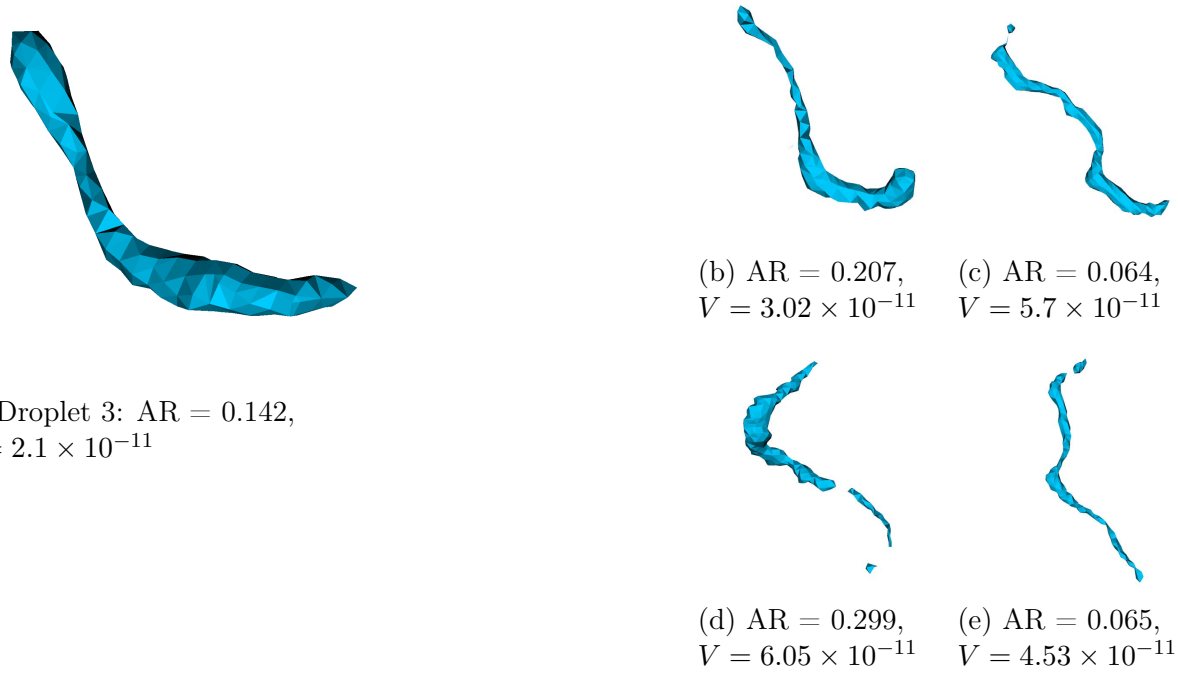
(e) AR = 0.174, $V = 1.3 \times 10^{-11}$

Figure 49: Droplet 2 and the four closest matches based on DIWECT

of PCA, we cluster the data using the K-means algorithm [114]. We choose K-means because of its scalability, with a complexity of $O(nki)$, where n is the number of droplets, k is the number of clusters, and i is the dimensionality after PCA. This methodology provides both computational efficiency and a systematic way to assess, in a quantitative manner, whether the DIWECT representation retains physical information from atomization systems.

We begin with relatively few clusters ($K = 4$) to illustrate how the droplets naturally separate into groups with similar morphologies. Figure 53 provides a random sampling of droplets from each cluster. Some droplets appear similar because they are either observed at consecutive update events (when only minor shape changes occur) or they represent large structures that experience a small breakup event (e.g. droplet shedding off of a ligament), which is not visible at the scale of the image.

From inspection of Figure 53, cluster 0 appears to contain small, elongate droplets,



(a) Droplet 3: AR = 0.142, $V = 2.1 \times 10^{-11}$

(b) AR = 0.207, $V = 3.02 \times 10^{-11}$

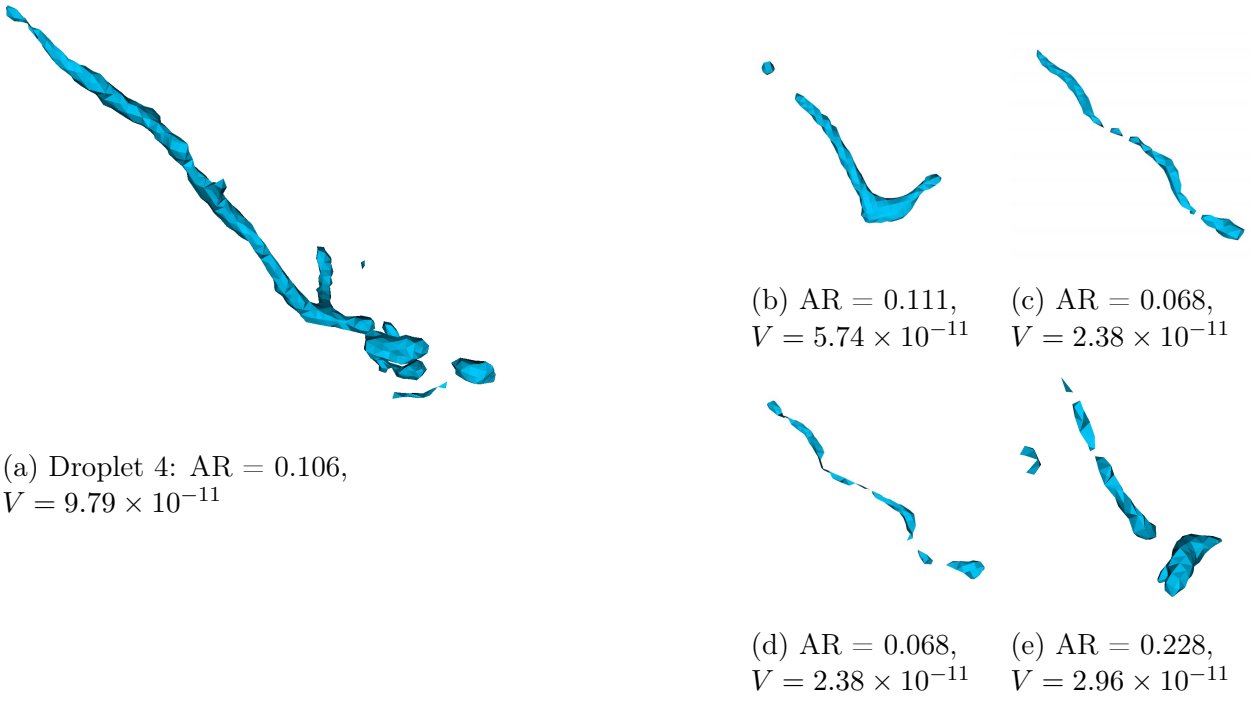
(c) AR = 0.064, $V = 5.7 \times 10^{-11}$

(d) AR = 0.299, $V = 6.05 \times 10^{-11}$

(e) AR = 0.065, $V = 4.53 \times 10^{-11}$

Figure 50: Droplet 3 and the four closest matches based on DIWECT

whereas cluster 1 has small but more spherical droplets. Clusters 2 and 3 generally contain larger, more complex shapes. A selection of statistics describing these clusters is shown in Figure 54, with mean values and standard deviations indicated by error bars. The quantities plotted—volume, aspect ratio, slip velocities, and cluster population—capture important physical characteristics of the breakup process. Aspect ratio is defined as the ratio of the smallest principal axis length to the largest principal axis length (L_{min}/L_{max}); lower values indicate more elongate shapes, while higher values indicate more spherical droplets. Slip velocities refer to the difference between the droplet velocity and the surrounding gas velocity in the x-, y-, and z-directions, which influences the aerodynamic forces acting on the droplet. The slip velocity is a slightly different (though related) quantity than the normal components of the velocity used as weights in the DIWECT. The slip velocity is calculated as a weighted sum of the gas velocity upstream of the droplet, which should represent the portion of the flow that impinges on the droplet.



(a) Droplet 4: AR = 0.106, $V = 9.79 \times 10^{-11}$

(b) AR = 0.111, $V = 5.74 \times 10^{-11}$

(c) AR = 0.068, $V = 2.38 \times 10^{-11}$

(d) AR = 0.068, $V = 2.38 \times 10^{-11}$

(e) AR = 0.228, $V = 2.96 \times 10^{-11}$

Figure 51: Droplet 4 and the four closest matches based on DIWECT

Figure 54 shows that clusters 0 and 1 are composed of large numbers of small droplets; cluster 0 tends to have more elongate droplets (lower aspect ratio), while cluster 1 contains droplets that are more spherical. In contrast, clusters 2 and 3 contain fewer but larger structures, with cluster 3 showing the greatest mean volume. Slip velocities do not vary drastically among clusters, suggesting that size and shape features contribute more toward the similarity or difference of droplets' respective DIWECT values.

Next, we increase the number of clusters to analyze. This will provide further analysis of how the DIWECT differentiates physical features. In the $K = 4$ clusters, it was clear that the size and shape of droplets were ultimately the key features in differentiating them (i.e. see Figs. 54a and 54b). While these are key factors in atomization, velocity contributes significantly to breakup mechanisms. The effects of velocity should become more apparent, as there are enough clusters to capture droplets with similar shapes, but differing local velocity fields. This will provide further confidence that the WECT can successfully differentiate

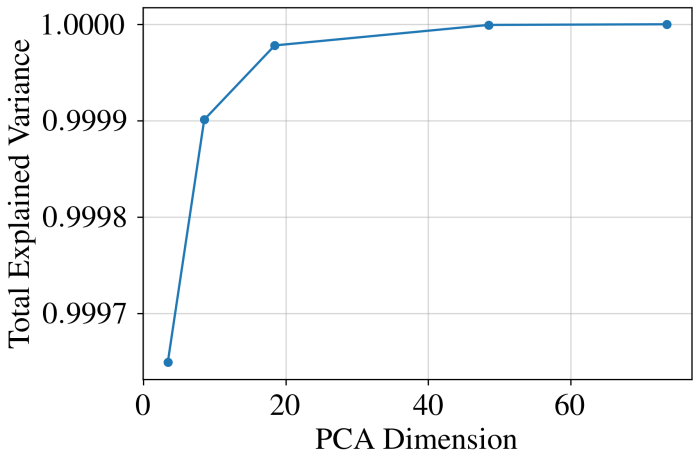
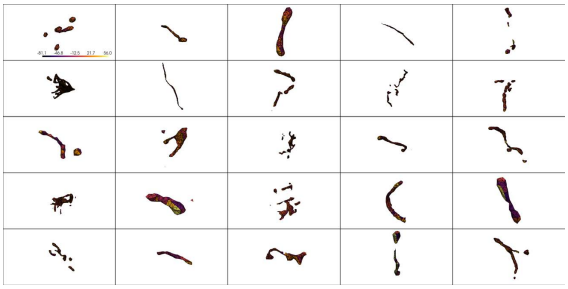


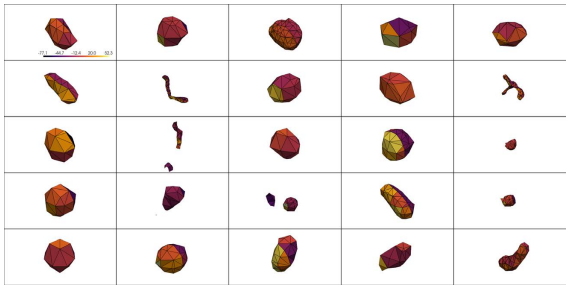
Figure 52: Total displayed variance of data with varying PCA dimensions.

structures by not only geometry, but also by their surrounding flow conditions.

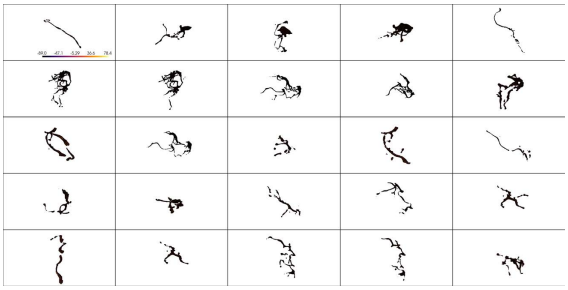
The results for $K = 16$ are shown in Fig. 55. The large cluster 1 in the original $K = 4$ collection is now spread out among clusters 1, 9, and 14, which display small-volume droplets with relatively high aspect ratios (see Fig. 56 for visual confirmation). Among these, cluster 14 exhibits a lower mean aspect ratio compared to clusters 1 and 9 (≈ 0.5 vs. ≈ 0.6). While clusters 1 and 9 appear geometrically similar, they were clustered separately, likely suggesting an influence from velocity. Notably, cluster 9 displays a stronger x-direction slip velocity than cluster 1 (≈ 35m/s vs. ≈ 25m/s), suggesting that droplets in this grouping were subjected to differing flow conditions, potentially experiencing higher shear. This differentiation implies that variations in velocity fields contribute to different breakup behaviors, which are now captured by the greater number of clusters.



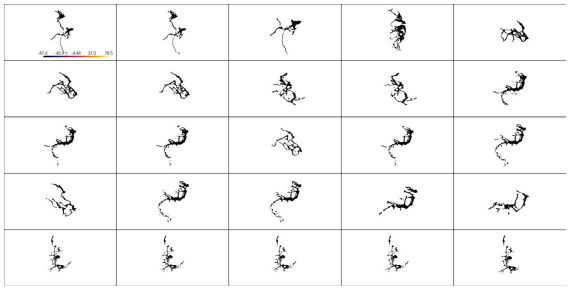
(a) K-means cluster 0.



(b) K-means cluster 1.

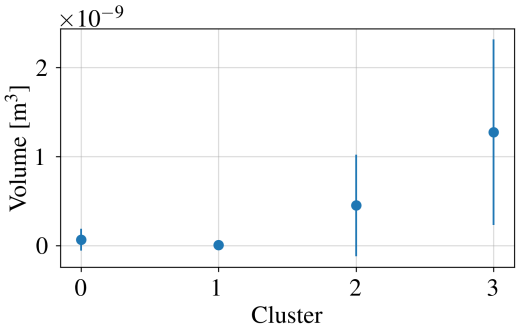


(c) K-means cluster 2.

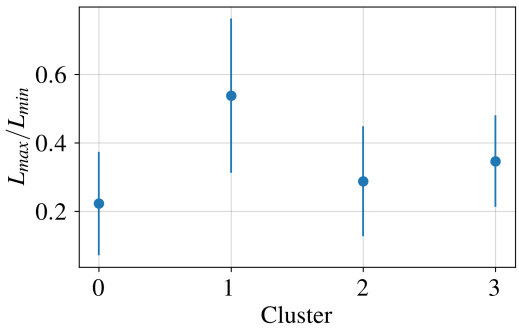


(d) K-means cluster 3.

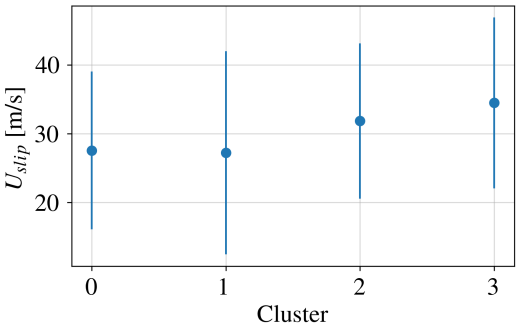
Figure 53: Visualization of the droplets in K-means clusters for $K = 4$.



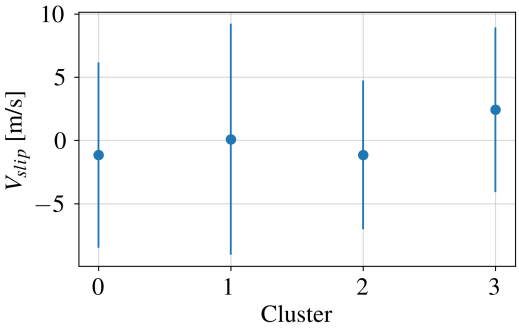
(a) Mean volume



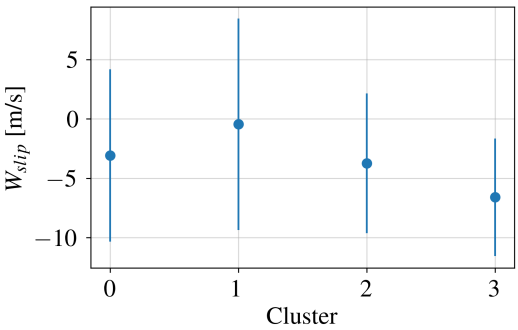
(b) Mean aspect ratio



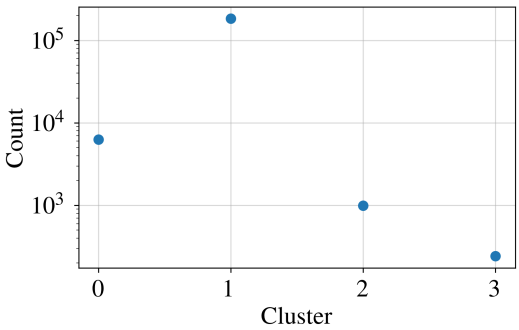
(c) Mean x-direction velocity



(d) Mean y-direction velocity

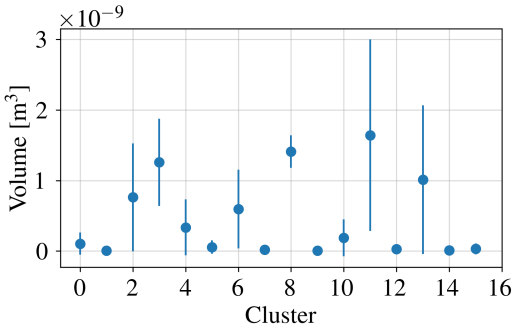


(e) Mean z-direction velocity

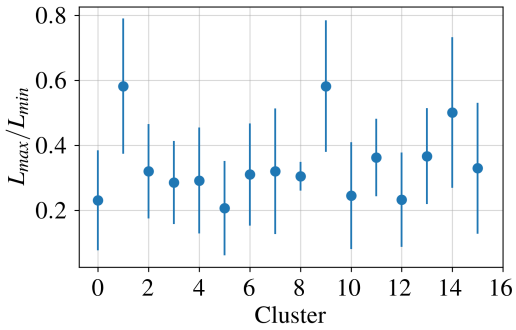


(f) Number of droplets

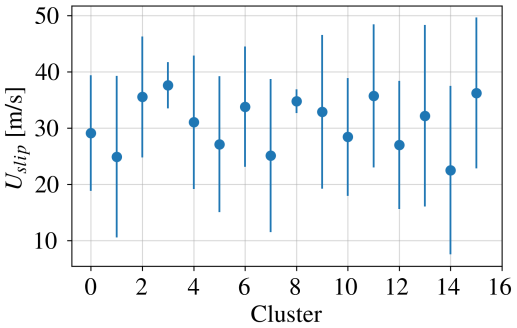
Figure 54: Statistics for clusters created by k-means clustering.



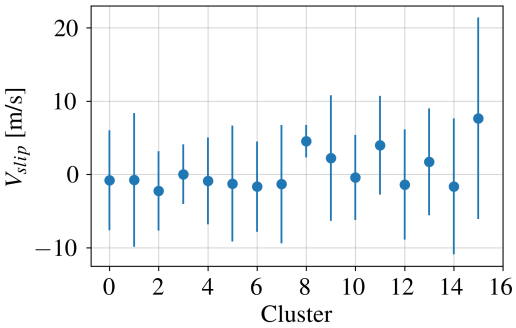
(a) Mean volume



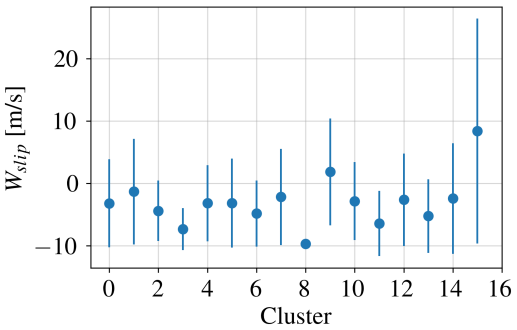
(b) Mean aspect ratio



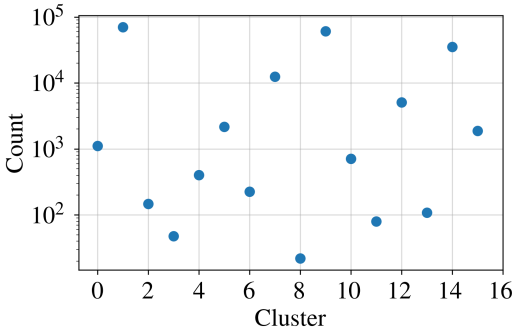
(c) Mean x-direction velocity



(d) Mean y-direction velocity

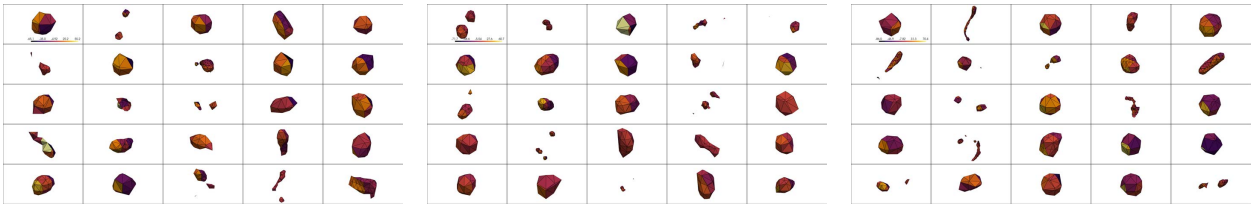


(e) Mean z-direction velocity



(f) Number of droplets

Figure 55: Statistics for clusters created by k-means clustering.



(a) Sample from cluster 1 (b) Sample from cluster 9 (c) Sample from cluster 14

Figure 56: Renderings of samples of droplets from clusters 1,9, and 14 in $K = 16$ clustering output.

Conclusions

This work has taken two significant steps toward a more comprehensive analysis of atomizing flows. First, we have shown that it is possible to extract liquid shapes and associated flow fields from high-fidelity simulations in a way that directly targets breakup and coalescence events. Previous approaches generally rely on full-field outputs [32, 60, 174, 221], but the method presented here isolates each relevant structure and captures localized information at the moment of breakup. This capability represents a shift from broad snapshots of the entire simulation domain to data that provides insights into model-relevant dynamics.

Second, we have demonstrated a means of transforming these complex, high-dimensional datasets into vectorized representations using the DIWECT. By approximating a mapping of geometry and local velocity fields to concise vectorized descriptors, we can readily apply machine learning or other data-driven tools to an otherwise intractable volume of simulation data. Our clustering examples illustrate that droplets sharing similar shape, size, and flow conditions appear in common groups, reaffirming that the DIWECT captures physically meaningful features.

While these contributions move atomization research closer to fully data-driven reduced-order modeling, there are limitations that must be addressed in future work. ASSET's

requirement that distinct liquid structures be separated by a full grid cell can sometimes cause the output geometries to misrepresent connected components, either by grouping individual structures together or artificially separating continuous ones. This issue is evident in Figs. 53 and 56 and presents a challenge for the DIWECT, which is sensitive to discontinuities in interface representation. This artifact can be reduced with increasing simulation resolution, which is possible with continually improving with numerical methods and computational efficiency. However, more robust treatment of liquid structures at the moment of breakup would improve the methodology. Implementations of numerical methods such as reconstruction with 2 planes (R2P) [58] in output simulations would greatly improve these artifacts.

Despite the limitations, these methods show promise that physical data can be recovered from the DIWECT. As noted in §5, given an infinite number of sampling directions and heights, the DIWECT is a one-to-one mapping of any shape. Practically, the method must be discretized, so we chose 150 directions and 75 heights for the present application, then an estimate of the DIWECT was computed. This resulted in a data reduction from 133 GB to 10 MB. Applying the methodology presented in this work to a predictive tool will necessitate tuning these sampling parameters, however this remains outside the scope of the present work.

The key results in this work show that the DIWECT can be used for finding the closest match which can determine the closest match to a given droplet in a database of droplets. This will be useful as a breakup model where the database can be used to predict the dynamics of a droplet. Additionally, clustering was performed by first computing PCA on the DIWECT database to further reduce the data size and to combat the curse of dimensionality. The clustering results first demonstrate the utility of the DIWECT as a way to represent droplets by classifying them into logical bins, and second to highlight the ability of the DIWECT to aid in analyzing simulation results where droplets can be clustered

to understand the number of droplets in various types.

Ultimately, this work bridges the gap between massive simulation datasets and meaningful data for model development. The methods outlined here represent significant improvements in the field of atomization simulations, by allowing greater utilization of massive, costly simulation runs. Future work will focus on model development to create a novel category of physics-based models without significant empirical assumptions.

LOCAL WEBER NUMBER CHARACTERIZATION

The *local Weber number* is defined analogously to the classical Weber number as

$$We_{\text{local}} = \frac{\rho_g \ \boldsymbol{U}_{\text{slip}}{}^2 D_{\text{eq}}}{\sigma}, \tag{51}$$

where D_{eq} is the volume-equivalent diameter of the liquid fragment (or droplet) in question, ρ_g is the gas density, σ is the surface tension, and $\boldsymbol{U}_{\text{slip}}$ is the *local* slip velocity between the droplet and the surrounding gas.

This quantity is vital for predictions of breakup. Commonly used reduced-order secondary atomization models such as the Taylor analogy breakup (TAB) [134], Pilch-Erdman (PE) [145], and Reitz-Diwakar (RD) [154] all use some value for a critical Weber number to predict the onset of breakup from aerodynamic forces. The local slip velocity ($\boldsymbol{U}_{\text{slip}}$) proves the most difficult quantity to accurately obtain. The slip velocity is the difference between the *unperturbed* gas velocity in the region around a droplet and the droplet's own velocity. Determining a gas velocity that is not affected by the droplet's boundary layer effects proves challenging, especially in dense droplet fields.

Eulerian-Lagrangian models often must calculate an unperturbed velocity field when a droplet is converted into a Lagrangian tracker particle [98, 112]. Generally these droplets are assumed to be small and spherical, so their effect on the flow is assumed to be relatively small. The data extracted from ASSET in the preceding Chapters 2, 3, and 4 show that there are a significant quantity of large droplets, which are likely to interact with the flow. Thus, a robust calculation of the local slip velocity is necessary to fully understand the effects of aerodynamic forces throughout a large-scale atomizing system. This represents a novel method to calculate the local Weber that takes into account highly complex, dense spray fields and the target droplet's own boundary layer.

Overview of the local slip-velocity calculation.

1. **Identify the droplet region.** The ASSET methodology uses a *band-growth* algorithm to identify individual droplets (or liquid fragments). This process gives us:

   - The cells belonging to each droplet,
   - The droplet's volume V,
   - Its centroid (center of gravity) $\boldsymbol{x}_{\mathrm{cg}}$.

2. **Initialize a level-set function.** We define

$$G = \mathrm{VOF} - 0.5,$$

   where VOF is the volume-of-fluid field. Consequently, $G > 0$ indicates a predominantly liquid region, whereas $G < 0$ indicates predominantly gas. This makes G a signed-distance function from the liquid-gas interface.

3. **Obtain the dominant gas velocity direction.** In order to decide which direction is "upstream," we compute an approximate gas velocity vector $\boldsymbol{U}_{g,\mathrm{dom}}$ by volume-weighting the gas velocity in cells near the droplet. The unit direction of this velocity is

$$\hat{\boldsymbol{u}}_g \;=\; \frac{\boldsymbol{U}_{g,\mathrm{dom}}}{\|\boldsymbol{U}_{g,\mathrm{dom}}\|}.$$

   For sampling, we proceed *against* this direction, i.e. $-\hat{\boldsymbol{u}}_g$.

4. **Compute the droplet's equivalent radius.** From the droplet volume V, we define the equivalent spherical diameter,

$$D_{\mathrm{eq}} = \left(\frac{6V}{\pi}\right)^{\frac{1}{3}},$$

   and the corresponding radius $r_{\mathrm{eq}} = \frac{1}{2}D_{\mathrm{eq}}$.

5. **Determine the sampling location.** We define a point $\boldsymbol{x}'$ located one droplet diameter *upstream* of the droplet's centroid:

$$\boldsymbol{x}' \;=\; \boldsymbol{x}_{\mathrm{cg}} \;-\; D_{\mathrm{eq}}\,\hat{\boldsymbol{u}}_g.$$

Conceptually, this point is chosen so that we gather the gas velocity in the region that will "impinge" on the droplet.

6. **Loop over the domain and compute weighted gas velocities.** For each grid cell with center $\boldsymbol{x}_j$, we skip any cells predominantly filled with liquid (i.e., where VOF > 0.5). For the remaining (gas) cells, we define a Gaussian weight

$$w_j \;=\; \exp\!\Big(-\frac{\|\boldsymbol{x}' - \boldsymbol{x}_j\|^2}{2\,r_{\mathrm{eq}}^2}\Big).$$

This weight decreases as the cell center is farther from $\boldsymbol{x}'$.

7. **Adjust weights near other droplets.** If a cell is within some threshold distance of another droplet interface (i.e. $|G| < th$), we scale the Gaussian weight by $\frac{G}{th}$, reflecting the effect of the boundary region on the gas velocity in those cells.

8. **Compute the weighted average gas velocity.** After summing over all suitable gas cells, the average gas velocity at $\boldsymbol{x}'$ is

$$\boldsymbol{U}_g^{\mathrm{avg}} \;=\; \frac{\displaystyle\sum_{j\in\mathcal{S}} w_j\,\boldsymbol{u}_g(\boldsymbol{x}_j)}{\displaystyle\sum_{j\in\mathcal{S}} w_j}, \tag{52}$$

where $\boldsymbol{u}_g(\boldsymbol{x}_j)$ is the gas velocity in cell j, and $\mathcal{S}$ is the set of all gas cells.

9. **Calculate the slip velocity and local Weber number.** Finally, define the local

slip velocity as

$$\boldsymbol{U}_{\text{slip}} \;=\; \boldsymbol{U}_g^{\text{avg}} \;-\; \boldsymbol{u}_l,$$

where $\boldsymbol{u}_l$ is the droplet's velocity (e.g., the volume-weighted average of liquid velocities in the droplet cells). This completes the computation of $\boldsymbol{U}_{\text{slip}}$, which can then be substituted into Eq. (51) to obtain We_{local}.

Significance. This locally resolved Weber number is valuable in characterizing breakup phenomena for droplets in complex flows, where the relative velocity between gas and droplet can differ significantly from a uniform bulk flow assumption. By focusing on a local upstream neighborhood and using weighting approach, this method provides a more accurate slip velocity, thus improving our ability to predict and interpret droplet breakup events in simulations.

CONCLUSIONS AND OUTLOOK

Summary

This dissertation presents a methodology to enhance the use of numerical simulations for gas-liquid atomizing flows and applies it to deepen our understanding of the underlying physics of these complex systems. Gaining insight into atomization improves the efficiency of numerous industrial, pharmaceutical, environmental, and energy-related processes. Ultimately, this work advances data extraction and analysis from simulations, providing access to previously unavailable quantitative data and offering new insights into practical atomizing systems. The methods developed here not only improve understanding of atomization physics, but also generate essential data for reduced-order models.

This dissertation presents advancements in the computational study of gas-liquid atomizing flows through the development and application of the Atomization Simulation Statistics Extraction Tool (ASSET). The work is structured around three main contributions: (1) the initial development of ASSET for efficient data extraction from high-fidelity atomization simulations, (2) the application of ASSET to study two practical atomization systems—single droplet breakup and electro-hydrodynamic (EHD) liquid jets, and (3) the enhancement of ASSET to extract detailed liquid and surrounding flow field geometries, ultimately preparing the extracted data for future reduced-order modeling.

Developing ASSET

The first major contribution—and the foundation for the other contributions—is the development of ASSET, an algorithmic tool built into high-fidelity simulations to extract data concurrently with the simulation run. As described in Chapter 2, ASSET identifies breakup and coalescence events as they occur, storing them in a Neo4j graph database. This provides access to localized, quantitative data that would otherwise require extensive post-

processing of tens-hundreds of terabytes of output data. Additionally, ASSET tracks liquid structures as they evolve throughout the simulation, developing breakup-pathways, which gives novel insights into how and when liquids break up. This tool represents a significant advancement in atomization research, allowing for quantitative analysis of directly applicable breakup physics.

Applying ASSET

With ASSET developed, the second phase of this research focused on applying it to study two practical atomization systems: single droplet breakup in high-speed gas flows and electro-hydrodynamic (EHD) liquid jets.

1. Droplet Aerobreakup: High-fidelity numerical simulations were conducted to investigate aerodynamic breakup mechanisms in Newtonian and non-Newtonian droplets, as described in Chapter 3. The system studied is inspired by aerial firefighting, which has significant real-world application. Applying ASSET allowed for an examination of breakup mechanisms at the complex boundary of multiple breakup regimes. The study confirmed that ligament formation drives breakup, and further qualitative analyses provided insights into instability formation and growth in droplet breakup.

2. EHD Liquid Jets: The role of electrostatic forces in atomization was explored through simulations of charged diesel-type liquid jets in Chapter 4. By incorporating ASSET into these simulations, it was possible to track the onset of primary atomization and assess the influence of electric charge on droplet size distribution, spray angle, and breakup efficiency. The results demonstrated that increased electric charge enhances atomization by destabilizing liquid structures earlier in the breakup process.

These studies not only offered simulations with which to assess ASSET's capabilities, but also provided insights into two important and relevant atomization systems.

Extracting Data for Model Development

The final contribution of this dissertation involves expanding ASSET's capabilities to extract data that are directly relevant for model development. Concretely, Chapter 5 describes a method to extract gas-liquid interface and local flow field geometries around droplets from simulations. The extracted geometries are then transformed into vectorized representations using topological data analysis techniques. The vectorized format is well suited for advanced data processing techniques, like machine-learning, which can be used to efficiently and accurately match liquid shapes and flow fields to inform simulations using databases of known droplet breakup pathways. The resultant database is drastically reduced from hundreds of terabytes to megabytes. This foundation allows for droplet breakup to be predicted with known physics rather than fully simulating the process, which is costly.

Another key addition is the ability to calculate a more accurate local Weber number from simulations, described in Chapter 6. The local Weber number can be used in reduced-order models to predict when a droplet is likely to break up due to aerodynamic forces. Most often, the Weber number is calculated using bulk characteristics of the flow, which neglects the complex, scale-variable, three-dimensional nature of gas flow in atomization systems. The method developed in this dissertation determines the local slip velocity between a droplet and the surrounding gas by identifying the dominant gas velocity direction and sampling upstream of the droplet where the impinging flow originates. A weighted average of the upstream velocities ensures that the effects of other droplet's boundary layers and wakes is accounted for in the calculation. This approach represents a novel calculation of a physically important non-dimensional quantity.

Future Work

The improvements described in this work advance atomization research, and specifically gas-liquid numerical simulations. However, significant challenges remain, and improvements or extensions to ASSET could continue driving progress in the field.

1. Reduced-order model development: A clear extension of this work is the development of a reduced-order model for atomization. Numerical simulations are incredibly expensive and impractical for most applications, and existing reduced-order models are highly limited in the systems they can accurately predict. Thus, the data extracted and vectorized in Chapter 5 should be used to develop a physics-based, universal atomization model. The data are best suited for a combined Eulerian-Lagrangian approach, where the liquid core is simulated and any droplets that break off are treated in a Lagrangian manner. When breakup occurs, droplets can be compared to databases of known droplet pathways to predict how and when the current Lagrangian droplet is likely to break up. This lookup process can be efficiently conducted using the DIWECT.

2. Droplet aerobreakup simulations: The study in Chapter 3 analyzes Newtonian and non-Newtonian droplets breaking up under aerodynamic forces. The parameter range is relatively narrow in that work and several extensions could be studied in future works:

    - A more physically accurate shear-thinning model to handle low shear-rates (e.g. Carreau-Yasuda model [12]) would provide greater applicability to real-world systems.
    - Expanding the range of Weber and Reynolds numbers to better define the transition from the Rayleigh-Taylor piercing (RTP) regime to the shear-induced entrapment (SIE) regime, where bag-type breakup is not likely to occur. This

could update aerobreakup regime maps and investigate the mechanisms in high-energy regimes that remain difficult to analyze with experimental methods alone [126, 191, 193].

- Since the system studied in Chapter 3 is inspired by aerial firefighting, incorporating viscoelasticity into the droplet viscosity model would better capture the behavior of fire-retardants, which are complex, non-Newtonian fluids exhibiting both shear-thinning and viscoelastic properties.

3. Resolution considerations: The greatest challenge with atomization simulations and in CFD in general is (and has always been) the incredible resolution requirements to capture the smallest relevant scales. As discussed in Chapter 3, great strides have been made in numerical methods and computational efficiency, but still, the small-scale mechanisms in atomization simulations remain under-resolved. Specifically, the ligament and thin sheet structures observed in these simulations cannot be accurately resolved with current methods. The power and capabilities of these simulations are likely to continue improving, however, at great cost. In the interim, implementing a two-plane reconstruction like that described by Giliberto and Desjardins [54] would allow for more accurate prediction of breakup mechanisms that occur below the resolvable grid scale.

4. Improvements to extracted geometries: The gas-liquid interface and flow field geometries extracted in Chapter 5 are computed using a marching cubes algorithm with topological guarantees [104]. This method ensures that the extracted shapes are water-tight, which is important for topological data analysis methods. However, the algorithm does not conserve volume, limiting the utility of the extracted shapes. While it effectively captures the general shape and topology of the original structure, its usefulness for further quantitative analysis is restricted. Enhancing this method to

improve volume accuracy would significantly increase its applicability.

5. Comprehensive aerial firefighting study:One of the most high-impact applications of atomization research is aerial firefighting. A better understanding of breakup mechanisms will enable more accurate predictions of spray distribution and deposition, leading to improved engineering of these systems. Aerial firefighting operates across a wide range of scales, from meters to microns, making it particularly challenging to model. However, with ASSET, structures can be extracted when the reach a size threshhold and ported into finer scale simulations as initial conditions and simulated again. The data collected at each scale can then be used to inform models, providing unprecedented physical insight for reduced-order modeling of this vital system.

REFERENCES CITED

[1] Alberto Roman Afanador, Stéphane Zaleski, Gretar Tryggvason, and Jiacai Lu. Effect of topology changes on the breakup of a periodic liquid jet. *Computers & Fluids*, 228:105059, 2021.

[2] Arpit Agarwal and Mario F. Trujillo. A closer look at linear stability theory in modeling spray atomization. *International Journal of Multiphase Flow*, 109:1–13, 12 2018.

[3] AA Amsden, JD Ramshaw, PJ O'rourke, and JK Dukowicz. Kiva: A computer program for two-and three-dimensional fluid flows with chemical reactions and fuel sprays. Technical report, Los Alamos National Lab., NM (USA), 1985.

[4] Anthony Amsden. Kiva3v. a block-structured kiva program for engines with vertical or canted valves. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 1997.

[5] Anthony A Amsden. Kiva3. a kiva program with block-structured mesh for complex geometries. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 1993.

[6] Anthony A Amsden, Peter J O'Rourke, and T Daniel Butler. Kiva-ii: A computer program for chemically reactive flows with sprays. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 1989.

[7] WH Anderson, RE Brown, PJ Blatz, NA Louie, and J Burchfield. Investigation of rheological properties of aerial- delivered fire retardant. Technical report, North. Forest Fire Lab., U.S. Forest Serv., 1974.

[8] Edgar L Andreas and Kerry A Emanuel. Effects of sea spray on tropical cyclone intensity. *Journal of the atmospheric sciences*, 58(24):3741–3751, 2001.

[9] Amar S Basu. Droplet morphometry and velocimetry (dmv): a video processing software for time-resolved, label-free tracking of droplet parameters. *Lab on a Chip*, 13(10):1892–1901, 2013.

[10] J. C. Baygents, N. J. Rivette, and H. A. Stone. Electrohydrodynamic deformation and interaction of drop pairs. *Journal of Fluid Mechanics*, 368:359–375, 1998.

[11] Jennifer C. Beale and Rolf D. Reitz. Modeling spray atomization with the kelvin-helmholtz/rayleigh-taylor hybrid model. *Atomization and Sprays*, 9, 1999.

[12] Robert Byron Bird, Robert Calvin Armstrong, and Ole Hassager. *Dynamics of polymeric liquids. Vol. 1: Fluid mechanics.* John Wiley and Sons Inc., New York, NY, 1987.

[13] JB Blaisot and Jerôme Yon. Droplet size and morphology characterization for dense sprays by image processing: application to the diesel spray. *Experiments in fluids*, 39(6):977–994, 2005.

[14] Francisco J. Blanco-Rodríguez and J. M. Gordillo. On the sea spray aerosol originated from bubble bursting jets. *Journal of Fluid Mechanics*, 886, 2020.

[15] L Bravo, D Kim, M Tess, M Kurman, F Ham, and C Kweon. High resolution numerical simulations of primary atomization in diesel sprays with single component reference fuels. *27th Annual Conference on Liquid Atomization and Spray Systems*, 2015.

[16] L. Broniarz-Press, M. Ochowiak, M. Matuszak, and S. Włodarczak. The effect of shear and extensional viscosity on atomization in medical inhaler. *International Journal of Pharmaceutics*, 468, 2014.

[17] M. Broumand and M. Birouk. Liquid jet in a subsonic gaseous crossflow: Recent progress and remaining challenges. *Progress in Energy and Combustion Science*, 57:1–29, 2016.

[18] Corentin Calbrix, Alexei Stoukov, Axelle Cadiere, Benoit Roig, and Dominique Legendre. Numerical simulation of aerial liquid drops of canadair cl-415 and dash-8 airtankers. *International Journal of Wildland Fire*, 32:1515–1528, 2023.

[19] Qin-Liu Cao, Wen-He Liao, Benveniste Natan, Feng Feng, and Wei-Tao Wu. Secondary atomization of non-Newtonian kerosene gel at low Weber numbers: A numerical study. *Aerospace Science and Technology*, 120:107280, January 2022.

[20] J. Chang, A.J. Kelly, and J.M. Crowley. *Handbook of electrostatic processes*. Marcel Dekker, Inc., Jan 1995.

[21] Xianyang Chen, Jiacai Lu, Stéphane Zaleski, and Grétar Tryggvason. Characterizing interface topology in multiphase flows using skeletons. *Physics of Fluids*, 34(9), 2022.

[22] Hank Childs, Eric Brugger, Brad Whitlock, Jeremy Meredith, Sean Ahern, David Pugmire, Kathleen Biagas, Mark Miller, Cyrus Harrison, Gunther H. Weber, Hari Krishnan, Thomas Fogal, Allen Sanderson, Christoph Garth, E. Wes Bethel, David Camp, Oliver Rübel, Marc Durant, Jean M. Favre, and Paul Navrátil. Visit: An end-user tool for visualizing and analyzing very large data. In *High Performance Visualization–Enabling Extreme-Scale Scientific Insight*, pages 357–372. Chapman & Hall, October 2012.

[23] Haecheon Choi and Parviz Moin. Effects of the computational time step on numerical solutions of turbulent flow. *Journal of Computational Physics*, 113, 7 1994.

[24] Brendan Christensen and Mark Owkes. Efficient extraction of atomization processes from high-fidelity simulations. *Computers and Fluids*, 254, 3 2023.

[25] CA Chryssakis, DN Assanis, and FX Tanner. Atomization models. *Handbook of Atomization and Sprays: Theory and Applications*, pages 215–231, 2011.

[26] Jessi Cisewski-Kehe, Brittany T. Fasy, Dhanush Giriyan, and Eli Quist. The weighted euler characteristic transform for image shape classification. *arXiv.org*, 2023.

[27] R.T. Collins, M.T. Harris, and O.A. Basaran. Breakup of Electrified Jets. *Journal of Fluid Mechanics*, 588:75–129, sep 2007.

[28] Lorin Crawford, Anthea Monod, Andrew X Chen, Sayan Mukherjee, and Raúl Rabadán. Predicting clinical outcomes in glioblastoma: an application of topological and functional data analysis. *Journal of the American Statistical Association*, 115(531):1139–1150, 2020.

[29] Joseph M Crowley. Electrostatic fundamentals, 1995.

[30] J. E. Dendy. Black box multigrid. *Journal of Computational Physics*, 48, 1982.

[31] Suraj S. Deshpande, Soumil R. Gurjar, and Mario F. Trujillo. A computational study of an atomizing liquid sheet. *Physics of Fluids*, 27:082108, 8 2015.

[32] O Desjardins and H Pitsch. Detailed numerical investigation of turbulent atomization of liquid jets. *Atomization and Sprays*, 20:311–336, 2010.

[33] Olivier Desjardins, Guillaume Blanquart, Guillaume Balarac, and Heinz Pitsch. High order conservative finite difference scheme for variable density low mach number turbulent flows. *Journal of Computational Physics*, 227:7125–7159, 7 2008.

[34] Olivier Desjardins, Jeremy O. McCaslin, Mark Owkes, and Peter Brady. Direct numerical and large-eddy simulation of primary atomization in complex geometries. *Atomization and Sprays*, 23, 2013.

[35] Olivier Desjardins, Vincent Moureau, and Heinz Pitsch. An accurate conservative level set/ghost fluid method for simulating turbulent atomization. *Journal of Computational Physics*, 227:8395–8416, 9 2008.

[36] Olivier Desjardins and H Pitsch. Detailed numerical investigation of turbulent atomization of liquid jets. *Atomization and Sprays*, 1(July), 2010.

[37] W Dettmer, PH Saksono, and D Perić. On a finite element formulation for incompressible newtonian fluid flows on moving domains in the presence of surface tension. *Communications in Numerical Methods in Engineering*, 19(9):659–668, 2003.

[38] Qingming Dong and Amalendu Sau. Unsteady electrorotation of a viscous drop in a uniform electric field. *Physics of Fluids*, 35(4), 2023.

[39] Benedikt Dorschner, Luc Biasiori-Poulanges, Kevin Schmidmayer, Hazem El-Rabii, and Tim Colonius. On the formation and recurrent shedding of ligaments in droplet aerobreakup. *Journal of Fluid Mechanics*, 904:A20, December 2020.

[40] W. Du and S. Chaudhuri. A Multiphysics Model For Charged Liquid Droplet Breakup In Electric Fields. *International Journal of Multiphase Flow*, 90:46–56, 2017.

[41] Fabien Evrard, Robert Michael Chiodi, Berend van Wachem, and Olivier Desjardins. Piecewise-parabolic interface reconstruction from first moments. Technical report, Los Alamos National Laboratory (LANL), Los Alamos, NM (United States), 2023.

[42] Fabien Evrard, Fabian Denner, and Berend van Wachem. A multi-scale approach to simulate atomisation processes. *International Journal of Multiphase Flow*, 119:194–216, 2019.

[43] G. M. Faeth, L. P. Hsiang, and P. K. Wu. Structure and breakup properties of sprays. *International Journal of Multiphase Flow*, 21, 1995.

[44] Robert D Falgout, Jim E Jones, and Ulrike Meier Yang. Pursuing scalability for hypre's conceptual interfaces. *ACM Transactions on Mathematical Software (TOMS)*, 31(3):326–350, 2005.

[45] Robert D Falgout, Jim E Jones, and Ulrike Meier Yang. The design and implementation of hypre, a library of parallel high performance preconditioners. In *Numerical solution of partial differential equations on parallel computers*, pages 267–294. Springer, 2006.

[46] Ronald P. Fedkiw, Tariq Aslam, Barry Merriman, and Stanley Osher. A non-oscillatory eulerian approach to interfaces in multimaterial flows (the ghost fluid method). *Journal of Computational Physics*, 152, 1999.

[47] J.Q. Feng. Electrohydrodynamic Behaviour Of A Drop Subjected To A Steady Uniform Electric Field At Finite Electric Reynolds Number. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, pages 2245–2269, 1999.

[48] J.Q. Feng and T.C. Scott. A Computational Analysis Of Electrohydrodynamics Of A Leaky Dielectric Drop In An Electric Field. *Journal of Fluid Mechanics*, 1996.

[49] Daniel Fredrich, Erik Weiand, and Andrea Giusti. Electrostatic fields for the control of evaporating charged fuel sprays. *International Journal of Multiphase Flow*, 160:104312, 2023.

[50] E Fuentes, H Coe, D Green, G De Leeuw, and G Mcfiggans. Laboratory-generated primary marine aerosol via bubble-bursting and atomization. *Atmos. Meas. Tech*, 3:141–162, 2010.

[51] D. Fuster, J. P. Matas, S. Marty, S. Popinet, J. Hoepffner, A. Cartellier, and S. Zaleski. Instability regimes in the primary breakup region of planar coflowing sheets. *Journal of Fluid Mechanics*, 736:150–176, 12 2013.

[52] J. Galle, S. Defruyt, C. Van de Maele, R. Piloto Rodriguez, Q. Denon, A. Verliefde, and S. Verhelst. Experimental investigation concerning the influence of fuel type and properties on the injection and atomization of liquid biofuels inanoptical combustion chamber. *Biomass and Bioenergy*, 57, 2013.

[53] Marcio Gameiro, Konstantin Mischaikow, and Thomas Wanner. Evolution of pattern complexity in the cahn–hilliard theory of phase separation. *Acta Materialia*, 53(3):693–704, 2005.

[54] Joseph Giliberto and Olivier Desjardins. Modeling the break-up of thin films in pressure-swirl atomization. In *ILASS-Americas 34th Annual Conference on Liquid Atomization and Spray Systems*, Ithaca, NY, USA, May 19-22 2024. Sibley School of Mechanical and Aerospace Engineering, Cornell University.

[55] Global Wildfire Information System. GWIS Wildfire Statistics Estimates, 2025. Accessed: 2025-02-03.

[56] Mikhael Gorokhovski and Marcus Herrmann. Modeling primary atomization. *Annual Review of Fluid Mechanics*, 40:343–366, 2008.

[57] D. R. Guildenbecher, C. López-Rivera, and P. E. Sojka. Secondary atomization. *Experiments in Fluids*, 46(3):371–402, March 2009.

[58] Austin Han, Robert Chiodi, and Olivier Desjardins. Capturing thin structures in vof simulations with two-plane reconstruction. arXiv, 2024.

[59] Y Hardalupas, AMKP Taylor, and JH Whitelaw. Characteristics of the spray from a diesel injector. *International Journal of Multiphase Flow*, 18(2):159–179, 1992.

[60] Josef Hasslberger, Sebastian Ketterl, Markus Klein, and Nilanjan Chakraborty. Flow topologies in primary atomization of liquid jets: A direct numerical simulation analysis. *Journal of Fluid Mechanics*, 859:819–838, 1 2019.

[61] Josef Hasslberger, Svenja Marten, and Markus Klein. A theoretical investigation of flow topologies in bubble- and droplet-affected flows. *Fluids*, 4, 2019.

[62] Taosong He, Lichan Hong, Dongqing Chen, and Zhengrong Liang. Reliable path for virtual endoscopy: Ensuring complete examination of human organs. *IEEE transactions on visualization and computer graphics*, 7(4):333–342, 2001.

[63] Theodore J. Heindel. X-ray imaging techniques to quantify spray characteristics in the near field. *Atomization and Sprays*, 28, 2018.

[64] Kelli Hendrickson, Gabriel D. Weymouth, and Dick K.P. Yue. Informed component label algorithm for robust identification of connected components with volume-of-fluid method. *Computers and Fluids*, 197, 1 2020.

[65] M. Herrmann. A parallel eulerian interface tracking/lagrangian point particle multi-scale coupling procedure. *Journal of Computational Physics*, 229:745–759, 2 2010.

[66] M. Herrmann. On simulating primary atomization using the refined level set grid method. *Atomization and Sprays*, 21, 2011.

[67] Masaki Hilaga, Yoshihisa Shinagawa, Taku Kohmura, and Tosiyasu L Kunii. Topology matching for fully automatic similarity estimation of 3d shapes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 203–212, 2001.

[68] J. O. Hinze. Fundamentals of the hydrodynamic mechanism of splitting in dispersion processes. *AIChE Journal*, 1, 1955.

[69] Cyril W Hirt and Billy D Nichols. Volume of fluid (vof) method for the dynamics of free boundaries. *Journal of computational physics*, 39(1):201–225, 1981.

[70] B Vajdi Hokmabad, S Faraji, T Ghaznavi Dizajyekan, B Sadri, and E Esmaeilzadeh. Electric field-assisted manipulation of liquid jet and emanated droplets. *International journal of multiphase flow*, 65:127–137, 2014.

[71] J. T. Holgate and M. Coppins. Shapes, Stability, And Hysteresis Of Rotating And Charged Axisymmetric Drops In A Vacuum. *Physics of Fluids*, 30(6):064107, June 2018.

[72] J. Hoshen and R. Kopelman. Percolation and cluster distribution. i. cluster multiple labeling technique and critical concentration algorithm. *Physical Review B*, 14, 1976.

[73] Ingrid Hotz, Talha Bin Masood, Filip Sadlo, and Julien Tierny. *Topological methods in data analysis and visualization VI*. Springer, 2021.

[74] Wei-Fan Hu, Ming-Chih Lai, and Yuan-Nan Young. A hybrid immersed boundary and immersed interface method for electrohydrodynamic simulations. *Journal of Computational Physics*, 282:47–61, 2015.

[75] Kang Y. Huh, Eunju Lee, and Ja-Ye Koo. Diesel spray atomization model considering nozzle exit turbulence conditions. *Atomization and Sprays*, 8:453–469, 1998.

[76] E. A. Ibrahim, H. Q. Yang, and A. J. Przekwas. Modeling of spray droplets deformation and breakup. *Journal of Propulsion and Power*, 9, 1993.

[77] Takashi Ikeda, Koichi Kotani, Yuichiro Maeda, and Hideki Kohno. Preliminary study on application of x-ray ct scanner to measurement of void fractions in steady state two-phase flows. *Journal of Nuclear Science and Technology*, 20(1):1–12, 1983.

[78] Isaac M. Jackiw and Nasser Ashgriz. On aerodynamic droplet breakup. *Journal of Fluid Mechanics*, 913:A33, 2021.

[79] Mohit Jain, R. Surya Prakash, Gaurav Tomar, and R. V. Ravikrishna. Secondary breakup of a drop at moderate Weber numbers. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2177):20140930, May 2015.

[80] M. Jalaal and K. Mehravaran. Transient growth of droplet instabilities in a stream. *Physics of Fluids*, 26(1):012101, January 2014.

[81] D. Jarrahbashi and W. A. Sirignano. Invited article: Vorticity dynamics for transient high-pressure liquid injection. *Physics of Fluids*, 26, 9 2014.

[82] D. Jarrahbashi, W. A. Sirignano, P. P. Popov, and F. Hussain. Early spray development at high gas density: Hole, ligament and bridge formations. *Journal of Fluid Mechanics*, 792:186–231, 3 2016.

[83] A. Jaworek and A.T. Sobczyk. Electrospraying Route To Nanotechnology: An Overview. *Journal of Electrostatics*, 66(3-4):197–219, 2008.

[84] Qitong Jiang, Sebastian Kurtek, and Tom Needham. The weighted euler curve transform for shape and image analysis. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 3685–3694, 2020.

[85] D. D. Joseph, G. S. Beavers, and T. Funada. Rayleigh–Taylor instability of viscoelastic drops at high Weber numbers. *Journal of Fluid Mechanics*, 453:109–132, February 2002.

[86] David Katoshevski. Analysis of a sea spray: Effect of simultaneous growth and fragmentation on droplet/particle size distribution of a multicomponent aerosol. *Atomization and Sprays*, 11, 2001.

[87] AJ Kelly. the electrostatic atomization of hydrocarbons. *Journal of the Institute of Energy*, 1984.

[88] K. Kim and R.J. Turnbull. Generation Of Charged Drops Of Insulating Liquids By Electrostatic Spraying. *Journal of Applied Physics*, 1976.

[89] A. Kourmatzis, P.X. Pham, and A.R. Masri. Air assisted atomization and spray density characterization of ethanol and a range of biodiesels. *Fuel*, 108:758–770, 2013.

[90] A. Kourmatzis and J. S. Shrimpton. Electrohydrodynamics and charge injection atomizers: A review of the governing equations and turbulence. *Journal of Atomization and Sprays*, 19:1045–1063, Dec 2009.

[91] Miroslav Kramár, Rachel Levanger, Jeffrey Tithof, Balachandra Suri, Mu Xu, Mark Paul, Michael F Schatz, and Konstantin Mischaikow. Analysis of kolmogorov flow and rayleigh–bénard convection using persistent homology. *Physica D: Nonlinear Phenomena*, 334:82–98, 2016.

[92] Kapilanjan Krishan, Huseyin Kurtuldu, Michael F Schatz, Marcio Gameiro, Konstantin Mischaikow, and Santiago Madruga. Homology and symmetry breaking in rayleigh-bénard convection: Experiments and simulations. *Physics of Fluids*, 19(11), 2007.

[93] Venkata Krisshna, Wanjiao Liu, and Mark Owkes. High-fidelity simulations of a rotary bell atomizer with electrohydrodynamic effects. *International Journal of Multiphase Flow*, 168:104566, 2023.

[94] Venkata Krisshna and Mark Owkes. Investigating atomization characteristics in an electrostatic rotary bell atomizer. *International Journal of Multiphase Flow*, page 104814, 2024.

[95] Will Krolick and Mark Owkes. Primary atomization instability extraction using dynamic mode decomposition. *Atomization and Sprays*, 28:1061–1079, 2018.

[96] Stefan A. Krzeczkowski. Measurement of liquid droplet disintegration mechanisms. *International Journal of Multiphase Flow*, 6(3):227–239, June 1980.

[97] Navin Kumar Chandra, Shubham Sharma, Saptarshi Basu, and Aloke Kumar. Shock-induced aerobreakup of a polymeric droplet. *Journal of Fluid Mechanics*, 965:A1, June 2023.

[98] Chia-Wei Kuo and Mario F Trujillo. Simulation of liquid jet atomization and droplet breakup via a volume-of-fluid lagrangian–eulerian strategy. *Physics of Fluids*, 34(11), 2022.

[99] S. Edward Law. Agricultural electrostatic spray application a review of significant research and development during the 20th century. *Journal of Electrostatics*, 51:25–42, 2001.

[100] Arthur H Lefebvre and Vincent G McDonell. *Atomization and sprays*. CRC press, 2017.

[101] Dominique Legendre. Fluid dynamics of airtanker firefighting. *Annual Review of Fluid Mechanics*, 56(1):577–603, 2024.

[102] W Lehr and W Hiller. Electrostatic atomization of liquid hydrocarbons. *Journal of Electrostatics*, 30:433–440, 1993.

[103] Rachel Levanger. *A Comparison Framework For Interleaved Persistence Modules and Applications of Persistent Homology to Problems in Fluid Dynamics*. PhD thesis, Rutgers, The State University of New Jersey, 2017.

[104] Thomas Lewiner, Hélio Lopes, Antônio Wilson Vieira, and Geovan Tavares. Efficient implementation of marching cubes' cases with topological guarantees. *Journal of graphics tools*, 8(2):1–15, 2003.

[105] Hongfei Li, Steven Cryer, John Raymond, and Lipi Acharya. Interpreting atomization of agricultural spray image patterns using latent dirichlet allocation techniques. *Artificial Intelligence in Agriculture*, 4, 2020.

[106] Xiaoyi Li and Marios C. Soteriou. High fidelity simulation and analysis of liquid jet atomization in a gaseous crossflow at intermediate weber numbers. *Physics of Fluids*, 28, 8 2016.

[107] Yang Li, Zhikun Xu, Xiaoyun Peng, Tianyou Wang, and Zhizhao Che. Numerical simulation of secondary breakup of shear-thinning droplets. *Physics of Fluids*, 35(1):012103, January 2023.

[108] Zhilong Li, Wenbo Zhao, and Zhijun Wu. Understanding transient internal flow processes in high-pressure nozzles using synchrotron radiation x-ray phase contrast imaging technology. *Atomization and Sprays*, 31, 2021.

[109] V. A. Likhanov and O. P. Lopatin. Model for calculating the characteristics of fuel injection and atomization in diesel when working on alcohol-fuel emulsions. *IOP Conference Series: Materials Science and Engineering*, 862, 2020.

[110] Y. Ling, D. Fuster, G. Tryggvason, and S. Zaleski. A two-phase mixing layer between parallel gas and liquid streams: Multiphase turbulence statistics and influence of interfacial instability. *Journal of Fluid Mechanics*, 859:268–307, 1 2019.

[111] Yue Ling, Daniel Fuster, Stéphane Zaleski, and Grétar Tryggvason. Spray formation in a quasiplanar gas-liquid mixing layer at moderate density ratios: A numerical closeup. *Physical Review Fluids*, 2, 1 2017.

[112] Yue Ling, Stéphane Zaleski, and Ruben Scardovelli. Multiscale simulation of atomization with small droplets represented by a lagrangian point-particle model. *International Journal of Multiphase Flow*, 76:122–143, 2015.

[113] Mark A Linne, Megan Paciaroni, James R Gord, and Terrence R Meyer. Ballistic imaging of the liquid core for a steady jet in crossflow. *Applied Optics*, 2005.

[114] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[115] J.M. López-Herrera, M.A. Herrada, and A.M. Gañán-Calvo. Electrokinetic modelling of cone-jet electrosprays. *Journal of Fluid Mechanics*, 964:A19, 2023.

[116] W.E. Lorensen and H. Cline. Marching cubes: A high resolution 3d surface construction algorithm. *Computer graphics*, 21(1):7–12, 1987.

[117] A Lozano and F Barreras. Experimental study of the gas flow in an air-blasted liquid sheet. *Experiments in fluids*, 31(4):367–376, 2001.

[118] Yuh-Fun Maa, Phuong-Anh Nguyen, Theresa Sweeney, Steven Shire, and Chung Hsu. Protein inhalation powders: Spray drying vs spray freeze drying. *Pharmaceutical Research*, 16, 7 1999.

[119] P. H. Marmottant and E. Villermaux. On spray formation. *Journal of Fluid Mechanics*, 498:73–111, 1 2004.

[120] J Melcher and C Smith. Electrohydrodynamic charge relaxation and interfacial perpendicular-field instability. *Phys. Fluids*, Jan 1969.

[121] James R Melcher. Charge relaxation on a moving liquid interface. *The Physics of Fluids*, 10(2):325–332, 1967.

[122] T Menard, PA Beau, S Tanguy, FX Demoulin, A Berlemont, et al. Primary break-up: Dns of liquid jet to improve atomization modelling. *WIT Transactions on Engineering Sciences*, 50, 2005.

[123] Thibault Ménard, Sebastien Tanguy, and Alain Berlemont. Coupling level set/vof/ghost fluid methods: Validation and application to 3d simulation of the primary break-up of a liquid jet. *International Journal of Multiphase Flow*, 33(5):510–524, 2007.

[124] Jomela C. Meng and Tim Colonius. Numerical simulation of the aerobreakup of a water droplet. *Journal of Fluid Mechanics*, 835:1108–1135, January 2018.

[125] Yogeshwar Nath Mishra, Elias Kristensson, and Edouard Berrocal. Reliable lif/mie droplet sizing in sprays using structured laser illumination planar imaging. *Optics express*, 22(4):4480–4492, 2014.

[126] VV Mitkin and TG Theofanous. The physics of aerobreakup. iv. strain-thickening liquids. *Physics of Fluids*, 29(12), 2017.

[127] Keith W Morton and Michael John Baines. *Numerical methods for fluid dynamics V*, volume 5. Oxford University Press, 1995.

[128] James A. Mueller and Fabrice Veron. A sea state–dependent spume generation function. *Journal of Physical Oceanography*, 39(9):2363 – 2372, 2009.

[129] Robert P. Mun, Brian W. Young, and David V. Boger. Atomisation of dilute polymer solutions in agricultural spray nozzles. *Journal of Non-Newtonian Fluid Mechanics*, 83, 1999.

[130] Florent Nauleau, Thibault Bridel-Bertomeu, Fabien Vivodtzev, Héloïse Beaugendre, and Julien Tierny. Topological data analysis for numerical method comparisons of 2d turbulent flows. In *Journée Visu 2021*, 2021.

[131] C-L Ng, R Sankarakrishnan, and KA Sallam. Bag breakup of nonturbulent liquid jets in crossflow. *International Journal of Multiphase Flow*, 34(3):241–259, 2008.

[132] J.A. Nicholls. Stream and droplet breakup by shock waves. In *NASA-SP-194 Rocket Combustion Instability*, pages 126–128. 1972.

[133] José Núñez, Ahtziri González, and Eduardo Ramos. Topological data analysis of lagrangian orbits in natural convection flows confined in a cylinder. *Physical Review Fluids*, 7(12):123501, 2022.

[134] P.J. O'Rourke and A. A. Amsden. The tab method for numerical calculation of spray droplet breakup. *SAE Technical Papers*, 1987.

[135] Stanley Osher and James A Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of computational physics*, 79(1):12–49, 1988.

[136] Mark Owkes, Eric Cauble, Jacob Senecal, and Robert A. Currie. Importance of curvature evaluation scale for predictive simulations of dynamic gas–liquid interfaces. *Journal of Computational Physics*, 365:37–55, 7 2018.

[137] Mark Owkes and Olivier Desjardins. A discontinuous galerkin conservative level set scheme for interface capturing in multiphase flows. *Journal of Computational Physics*, 249:275–302, 2013.

[138] Mark Owkes and Olivier Desjardins. A computational framework for conservative, three-dimensional, unsplit, geometric transport with application to the volume-of-fluid (vof) method. *Journal of Computational Physics*, 270:587–612, 8 2014.

[139] Mark Owkes and Olivier Desjardins. A mass and momentum conserving unsplit semi-lagrangian framework for simulating multiphase flows. *Journal of Computational Physics*, 332:21–46, 3 2017.

[140] Cesar I Pairetti, Santiago Marquez Damian, Norberto M Nigro, Stéphane Popinet, and Stéphane Zaleski. Mesh resolution effects on primary atomization simulations. *Atomization and Sprays*, 30(12), 2020.

[141] H. Paknemat, a. R. Pishevar, and P. Pournaderi. Numerical simulation of drop deformations and breakup modes caused by direct current electric fields. *Physics of Fluids*, 24(10):102101, 2012.

[142] Mark A. Patterson and Rolf D. Reitz. Modeling the effects of fuel spray characteristics on diesel engine combustion and emission. *SAE Technical Papers*, 1998.

[143] P. X. Pham, A. Kourmatzis, and A. R. Masri. Local characteristics of fragments in atomizing sprays. *Experimental Thermal and Fluid Science*, 95:44–51, 7 2018.

[144] Phuong X. Pham, Agisilaos Kourmatzis, and Assaad R. Masri. Simultaneous volume-velocity measurements in the near field of atomizing sprays. *Measurement Science and Technology*, 28, 10 2017.

[145] M Pilch and C A Erdman. Use of breakup time data and velocity history data to predict the maximum size of stable fragments for acceleration-induced breakup of a liquid drop. *Int. J. Multiphase Flow*, 13:741–757, 1987.

[146] James Edward Pilliod and Elbridge Gerry Puckett. Second-order accurate volume-of-fluid algorithms for tracking material interfaces. *Journal of Computational Physics*, 199:465–502, 9 2004.

[147] James E Pilliod Jr. Second-order accurate volume-of-fluid algorithms for tracking material interfaces. *Journal of Computational Physics*, 199(2), 1997.

[148] Stephen M Pizer, Guido Gerig, Sarang Joshi, and Stephen R Aylward. Multiscale medial shape-based analysis of image objects. *Proceedings of the IEEE*, 91(10):1670–1679, 2003.

[149] Johanna Potyka and Kathrin Schulte. A volume of fluid method for three dimensional direct numerical simulations of immiscible droplet collisions. *International Journal of Multiphase Flow*, 170:104654, 2024.

[150] Surya Prakash, Suhas S Jain, Jeffery A Lovett, B N Raghunandan, R V Ravikrishna, and Gaurav Tomar. Detailed numerical simulations of atomization of a liquid jet in a swirling gas crossflow. *Atomization and Sprays*, pages 1–28, 2019.

[151] J Qian and Chung King Law. Regimes of coalescence and separation in droplet collision. *Journal of fluid mechanics*, 331:59–80, 1997.

[152] Lord Rayleigh. On the instability of jets. *Proceedings of the London Mathematical Society*, s1-10, 1878.

[153] Lord Rayleigh. On the equilibrium of liquid conducting masses. *Philisophical Magazine, Series 5*, 14:184–186, Apr 1882.

[154] R. D. Reitz and R. Diwakar. Structure of high-pressure fuel sprays. *SAE Technical Papers*, 1987.

[155] ARH Rigit and John S Shrimpton. Spray characteristics of charge injection electrostatic atomizers with small-orifice diameters. *Atomization and Sprays*, 16(4), 2006.

[156] Nicolas Rimbert, S Castrillon Escobar, Renaud Meignen, M Hadj-Achour, and Michel Gradeck. Spheroidal droplet deformation, oscillation and breakup in uniform outer flow. *Journal of Fluid Mechanics*, 904:A15, 2020.

[157] Clément Rouaix, Alexei Stoukov, Yannick Bury, David Joubert, and Dominique Legendre. Liquid jet breakup in gaseous crossflow injected through a large diameter nozzle. *International Journal of Multiphase Flow*, 163:104419, 2023.

[158] A Roy, RAI Haque, AJ Mitra, S Tarafdar, and T Dutta. Characterizing fluid dynamical systems using euler characteristic surface and euler metric. *Physics of Fluids*, 35(8), 2023.

[159] Clark Rubel and Mark Owkes. Extraction of droplet genealogies from high-fidelity atomization simulations. *Atomization and Sprays*, 29:709–739, 2019.

[160] Murray Rudman. Volume-tracking methods for interfacial flow calculations. *International journal for numerical methods in fluids*, 24(7):671–691, 1997.

[161] F. J. Salvador, S. Ruiz, Marco Crialesi-Esposito, and Ignacio Blanquer. Analysis on the effects of turbulent inflow conditions on spray primary atomization in the near-field by direct numerical simulation. *International Journal of Multiphase Flow*, 102:49–63, 5 2018.

[162] Felix Savart. Mémoire sur la constitution des veines liquides lancées par des orifices circulaires en mince paroi. *Annales de chimie et de physique*, 53:337–386, 1833.

[163] DA Saville. Electrohydrodynamic stability: effects of charge relaxation at the interface of a liquid jet. *Journal of Fluid Mechanics*, 48(4):815–827, 1971.

[164] D.A. Saville. Electrohydrodynamics: the taylor-melcher leaky dielectric model. *Annual review of fluid mechanics*, 29(1):27–64, 1997.

[165] David P Schmidt, Idriss Nouar, PK Senecal, J Rutland, JK Martin, Rolf D Reitz, and Jeffrey A Hoffman. Pressure-swirl atomization in the near field. *SAE transactions*, pages 471–484, 1999.

[166] Ory Schnitzer, Itzchak Frankel, and Ehud Yariv. Electrokinetic flows about conducting drops. *Journal of Fluid Mechanics*, 722:394–423, 2013.

[167] Ory Schnitzer and Ehud Yariv. The taylor–melcher leaky dielectric model as a macroscale electrokinetic description. *Journal of Fluid Mechanics*, 773:1–33, 2015.

[168] David Sedarsky, Edouard Berrocal, and Mark Linne. Numerical analysis of ballistic imaging for revealing liquid breakup in dense sprays. *Atomization and Sprays*, 20, 2010.

[169] JA Sethian. Level set methods and fast marching methods. *Cambridge University*, 1999.

[170] Shubham Sharma, Awanish Pratap Singh, S. Srinivas Rao, Aloke Kumar, and Saptarshi Basu. Shock induced aerobreakup of a droplet. *Journal of Fluid Mechanics*, 929:A27, December 2021.

[171] Patrick Sheehy and Mark Owkes. Numerical study of electric reynolds number on electrohydrodynamic (ehd) assisted atomization. *Atomization and Sprays*, 27, 2017.

[172] Patrick Sheehy and Mark Owkes. Numerical study of electric reynolds number on electrohydrodynamic (ehd) assisted atomization. *Atomization and Sprays*, 27(7), 2017.

[173] JD Sherwood. Breakup of fluid droplets in electric and magnetic fields. *Journal of Fluid Mechanics*, 188:133–146, 1988.

[174] J. Shinjo and A. Umemura. Simulation of liquid jet primary breakup: Dynamics of ligament and droplet formation. *International Journal of Multiphase Flow*, 36:513–532, 7 2010.

[175] J. Shinjo and A. Umemura. Surface instability and primary atomization characteristics of straight liquid jet sprays. *International Journal of Multiphase Flow*, 37:1294–1304, 12 2011.

[176] John Shrimpton and Agissilaos Kourmatzis. Direct numerical simulation of forced flow dielectric ehd within charge injection atomizers. *IEEE transactions on dielectrics and electrical insulation*, 17(6):1838–1845, 2010.

[177] John S Shrimpton and Yossapong Laoonual. Dynamics of electrically charged transient evaporating sprays. *International journal for numerical methods in engineering*, 67(8):1063–1081, 2006.

[178] JS Shrimpton. Pulsed charged sprays: application to disi engines during early injection. *International Journal for Numerical Methods in Engineering*, 58(3):513–536, 2003.

[179] J.S. Shrimpton and A.J. Yule. Atomization, combustion, and control of charged hydrocarbon sprays. *Journal of Atomization and Sprays*, 11:365–396, Feb 2001.

[180] C.M. Sipperley and W.D. Bachalo. Volumetric imaging and multi-angle illumination for dense sprays characterization. *Proceedings of ILASS Americas, 25th Annual Conference on Liquid Atomization and Spray Systems*, 2013.

[181] Pierre R. Slangen, Pierre Lauret, Frederic Heymes, Laurent Aprin, and Nicolas Lecysyn. High-speed imaging optical techniques for shockwave and droplets atomization analysis. *Optical Engineering*, 55:121706, 6 2016.

[182] Gregory J Smallwood, Ömer L Gülder, and David R Snelling. The structure of the dense core region in transient diesel sprays. In *Symposium (international) on combustion*, volume 25, pages 371–379. Elsevier, 1994.

[183] Otmar M Stuetzer. Magnetohydrodynamics and electrohydrodynamics. *The Physics of Fluids*, 5(5):534–544, 1962.

[184] C. Sula, H. Grosshans, and M. V. Papalexandris. Assessment of droplet breakup models for spray flow simulations. *Flow, Turbulence and Combustion*, 105:889–914, 9 2020.

[185] K Tang and A Gomez. On the structure of an electrostatic spray of monodisperse droplets. *Physics of Fluids*, 6(7):2317–2332, Jan 1994.

[186] F. X. Tanner. Liquid jet atomization and droplet breakup modeling of non-evaporating diesel fuel sprays. *SAE Technical Papers*, 1997.

[187] G. Taylor. Disintegration of Water Drops in an Electric Field. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 280(1382):383–397, jul 1964.

[188] G. I. Taylor and a. D. McEwan. The stability of a horizontal fluid interface in a vertical electric field. *Journal of Fluid Mechanics*, 22(01):1, mar 2006.

[189] Geoffrey Taylor and Proc R Soc Lond A. Electrically Driven Jets. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences (1969)*, pages 453–475, 1969.

[190] Geoffrey Ingram Taylor. Studies in electrohydrodynamics. I. The circulation produced in a drop by an electric field. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 291(1425):159–166, 1966.

[191] T. G. Theofanous and G. J. Li. On the physics of aerobreakup. *Physics of Fluids*, 20(5):052103, May 2008.

[192] T. G. Theofanous, V. V. Mitkin, C. L. Ng, C-H. Chang, X. Deng, and S. Sushchikh. The physics of aerobreakup. II. Viscous liquids. *Physics of Fluids*, 24(2):022104, February 2012.

[193] T.G. Theofanous. Aerobreakup of Newtonian and Viscoelastic Liquids. *Annual Review of Fluid Mechanics*, 43(1):661–690, January 2011.

[194] TG Theofanous, GJ Li, and Truc-Nam Dinh. Aerobreakup in rarefied supersonic gas flows. *J. Fluids Eng.*, 126(4):516–527, 2004.

[195] TG Theofanous, GJ Li, Truc-Nam Dinh, and C-H Chang. Aerobreakup in disturbed subsonic and supersonic flow fields. *Journal of Fluid Mechanics*, 593:131–170, 2007.

[196] G Tomar, D Gerlach, G Biswas, N Alleborn, A Sharma, F Durst, SWJ Welch, and A Delgado. Two-phase electrohydrodynamic simulations using a volume-of-fluid approach. *Journal of Computational Physics*, 227(2):1267–1285, 2007.

[197] Gaurav Tomar, Daniel Fuster, Stéphane Zaleski, and Stéphane Popinet. Multiscale simulations of primary atomization. *Computers & Fluids*, 39(10):1864–1874, 2010.

[198] Antonio J. Torregrosa, Raúl Payri, F. Javier Salvador, and Marco Crialesi-Esposito. Study of turbulence in atomizing liquid jets. *International Journal of Multiphase Flow*, 129, 8 2020.

[199] David J Torres and Mario F Trujillo. Kiva-4: An unstructured ale code for compressible gas flow with sprays. *Journal of Computational Physics*, 219(2):943–975, 2006.

[200] Elias Trautner, Josef Hasslberger, Sebastian Ketterl, and Markus Klein. Primary atomization of liquid jets: Identification and investigation of droplets at the instant of their formation using direct numerical simulation. *International Journal of Multiphase Flow*, 160:104360, 2023.

[201] Katharine Turner, Sayan Mukherjee, and Doug M Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[202] U.S. Department of Energy. Modeling of combustion processes improves engine efficiency, January 2006.

[203] BP Van Poppel, Olivier Desjardins, and JW Daily. A ghost fluid, level set methodology for simulating multiphase electrohydrodynamic flows with application to liquid fuel injection. *Journal of Computational Physics*, 229(20):7977–7996, 2010.

[204] Bret P Van Poppel, Brendan Christensen, Mark Owkes, and Venkata Krisshna. Framework for analyzing atomization in electrically charged diesel-type jets. In *ILASS Americas 34th Conference on Liquid Atomization and Spray Systems*, pages 1–10, 2024.

[205] Fabrice Veron. Ocean spray. *Annual Review of Fluid Mechanics*, 47(1):507–538, 2015.

[206] Emmanuel Villermaux. Fragmentation. *Annu. Rev. Fluid Mech.*, 39(1):419–446, 2007.

[207] Emmanuel Villermaux. Fragmentation versus cohesion. *Journal of Fluid Mechanics*, 898:P1, 2020.

[208] C Wang, S Chang, H Wu, L Ding, and J M Thompson. Theoretical modeling of spray drop deformation and breakup in the multimode breakup regime. *Atomization and Sprays*, 25:857–869, 2015.

[209] C. Wang, S. Chang, H. Wu, and J. Xu. Modeling of drop breakup in the bag breakup regime. *Applied Physics Letters*, 104, 4 2014.

[210] Y. J. Wang, Kyoung Su Im, K. Fezzaa, W. K. Lee, Jin Wang, P. Micheli, and C. Laub. Quantitative x-ray phase-contrast imaging of air-assisted water sprays with high weber numbers. *Applied Physics Letters*, 89, 2006.

[211] Yue Wang, Won Geun Lee, Rolf D. Reitz, and Ramachandra Diwakar. Numerical simulation of diesel sprays using an eulerian-lagrangian spray and atomization (elsa) model coupled with nozzle flow. *Proceedings of: SAE 2011 World Congress & Exhibition*, 4 2011.

[212] Zi-Yu Wang, Hui Zhao, Wei-Feng Li, Jian-Liang Xu, and Hai-Feng Liu. Secondary breakup of shear thickening suspension drop. *Physics of Fluids*, 33(9):093103, September 2021.

[213] Tomoaki Watanabe, Yasuhiko Sakai, Kouji Nagata, Yasumasa Ito, and Toshiyuki Hayase. Vortex stretching and compression near the turbulent/non-turbulent interface in a planar jet. *Journal of Fluid Mechanics*, 758:754–785, 11 2014.

[214] Constantin Weber. Zum zerfall eines flüssigkeitsstrahles. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 11(2):136–154, 1931.

[215] K-J Wu, RD Reitz, and FV Bracco. Measurements of drop size at the spray edge near the nozzle in atomizing liquid jets. *The Physics of fluids*, 29(4):941–951, 1986.

[216] Zonglu Xie, Qingming Dong, Amalendu Sau, and Zhentao Wang. Self-propulsion of a Quincke droplet on a superhydrophobic wall under low electric Reynolds number $Re_E \leq 1$. *Physics of Fluids*, 36(8), 2024.

[217] Xin Yang and Henry Potter. A novel method to discriminate active from residual whitecaps using particle image velocimetry. *Remote Sensing*, 13(20):4051, 2021.

[218] David L. Youngs. Time-dependent multi-material flow with large fluid distortion. In K. W. Morton and M. J. Baines, editors, *Numerical Methods in Fluid Dynamics*, pages 273–285. Academic Press, 1 1982.

[219] Hongjiang Yu, L Goldsworthy, Mohammadmahdi Ghiji, Paul A Brandner, and Vikrambhai Garaniya. A parallel volume of fluid-lagrangian parcel tracking coupling procedure for diesel spray modelling. *Computers & Fluids*, 150:46–65, 2017.

[220] AJ Yule, JS Shrimpton, AP Watkins, W Balachandran, and D Hu. Electrostatically atomized hydrocarbon sprays. *Fuel*, 74(7):1094–1103, 1995.

[221] A. Zandian, W. A. Sirignano, and F. Hussain. Planar liquid jet: Early deformation and atomization cascades. *Physics of Fluids*, 29, 6 2017.

[222] A. Zandian, W. A. Sirignano, and F. Hussain. Understanding liquid-jet atomization cascades via vortex dynamics. *Journal of Fluid Mechanics*, 843:293–354, 5 2018.

[223] F Zhao, M.-C Lai, and D L Harrington. Automotive spark-ignited direct-injection gasoline engines. *Progress in Energy and Combustion Science*, 25:437–562, 1999.

[224] Hui Zhao, Hai-Feng Liu, Jian-Liang Xu, and Wei-Feng Li. Secondary breakup of coal water slurry drops. *Physics of Fluids*, 23(11):113101, November 2011.

[225] Davide Zuzio, Jean-Luc Estivalèzes, and Bastien DiPierro. An improved multiscale eulerian–lagrangian method for simulation of atomization process. *Computers & Fluids*, 176:285–301, 2018.

APPENDICES

APPENDIX A

LIQUID IDENTIFICATION AND COALESCENCE CODE

The following is a pseudocode describing how the liquid identification number is transported and coalescence is identified.

Algorithm 2: $\mathcal{L}$-transport and Coalescence Pseudocode

for $i = 1 \rightarrow N_{\text{cell}}$ **do** ▷ Loop over cells in domain
  ▷ **Form list of $\mathcal{L}$ for this cell**
  $\mathcal{L}_{\text{count}} = 0$
  if cell has interface or VoF > 0.5 **then** ▷ Search through cells with liquid
    if $\mathcal{L} > 0$ **then**
      $\mathcal{L}_{\text{count}} \leftarrow \mathcal{L}_{\text{count}} + 1$ ▷ Keep track of the number of unique $\mathcal{L}$ values within the cell
      $\mathcal{L}_{\text{cell}}(\mathcal{L}_{\text{count}}) \leftarrow \mathcal{L}$ ▷ Keep track of $\mathcal{L}$ value in the cell
    end if
  end if
  for $f = 1 \rightarrow N_{\text{faces}}$ **do** ▷ Loop over all faces of a cell
    for $n = 1 \rightarrow \mathcal{L}_{\max}$ **do** ▷ Loop through all $\mathcal{L}$ that can flux through a face
      if $\mathcal{L}_{\text{flux},n} \neq 0$ **then** ▷ Check to see if there is an n^{th} $\mathcal{L}$ flux into the cell
        if $\mathcal{L}_{\text{flux},n} \neq \mathcal{L}_{\text{cell}}(1 : \mathcal{L}_{\text{cell}})$ **then** ▷ Check to see if $\mathcal{L}$ fluxes matches any of the current $\mathcal{L}$ in the cell
          $\mathcal{L}_{\text{count}} \leftarrow \mathcal{L}_{\text{count}} + 1$ ▷ Update $\mathcal{L}$ counter in the cell
          $\mathcal{L}_{\text{cell}}(\mathcal{L}_{\text{count}}) \leftarrow \mathcal{L}_{\text{flux}}$ ▷ Keep track of all $\mathcal{L}$ values in the cell
        end if
      end if
    end for
  end for
  ▷ **Check for coalescence or merge event**
  if $\mathcal{L}_{\text{count}} \geq 2$ **then** ▷ More than one unique $\mathcal{L}$ for this cell
    for $n = 2 \rightarrow \mathcal{L}_{\text{count}}$ **do**
      nMerge = nMerge+1 ▷ Keep track of the number of coalescence events
      $\mathcal{L}1(\text{nMerge}) = \min(\mathcal{L}_{\text{cell}}(1), \mathcal{L}_{\text{cell}}(n))$ ▷ Make a list of the $\mathcal{L}$'s which need to be merged
      $\mathcal{L}2(\text{nMerge}) = \max(\mathcal{L}_{\text{cell}}(1), \mathcal{L}_{\text{cell}}(n))$
    end for
  end if
end for
▷ **Compute new $\mathcal{L}$ after coalescence event**
for $n = 1 \rightarrow$ nMerge **do** ▷ Loop over the all identified coalescence events
  if $\mathcal{L}1(n) = \mathcal{L}1(n+1)$ **then** ▷ Combine identical events
    if $\mathcal{L}2(n) = \mathcal{L}2(n+1)$ **then**
      $\mathcal{L}_{new} \leftarrow min(\mathcal{L}1, \mathcal{L}2)$ ▷ Assign the smaller $\mathcal{L}$ to the new merged structure
      $\mathcal{L}_{old} \leftarrow max(\mathcal{L}1, \mathcal{L}2)$
    end if
  end if
end for
▷ **Update $\mathcal{L}$ in domain**
where $(\mathcal{L} = \mathcal{L}_{old})$: $\mathcal{L} \leftarrow \mathcal{L}_{new}$ ▷ Update $\mathcal{L}$'s throughout domain

APPENDIX B

NEO4J DATA INPUT

Data from our simulation are exported in the form of a CSV file. Each row within the CSV represents a breakup event within the simulation and each column contains statistics we extracted from the event. Column headers are listed in the following Cypher code, following the "csvline" variable (i.e. csvline.OldLID and csvline.NewLID are the old and new $\mathcal{L}$'s associated with each droplet). We import the CSV into Neo4j to create nodes and relationships. Each node corresponds to a row within the CSV and the relationships are created using the identification numbers.

```
// Import data from CSV and create split/merge nodes
LOAD CSV WITH HEADERS FROM "link_to_csv" AS csvline CREATE (n:
    droplet {id: TOINTEGER(csvline.NewLID), Event: csvline.
    Merge_Split, OldLID: TOINTEGER(csvline.OldLID), OldSID:
    TOINTEGER(csvline.OldSID), NewLID: TOINTEGER(csvline.NewLID),
    NewSID: TOINTEGER(csvline.NewSID), Volume: TOFLOAT(csvline.Vol)
    , Event_Time: TOFLOAT(csvline.Time), X: TOFLOAT(csvline.X), Y:
     TOFLOAT(csvline.Y), Z: TOFLOAT(csvline.Z), U: TOFLOAT(
    csvline.U), V: TOFLOAT(csvline.V), W: TOFLOAT(csvline.W),
    U_gas: TOFLOAT(csvline.U_gas), V_gas: TOFLOAT(csvline.V_gas),
     W_gas: TOFLOAT(csvline.W_gas), L1_old: TOFLOAT(csvline.
    L1_old), L2_old: TOFLOAT(csvline.L2_old), L3_old: TOFLOAT(
    csvline.L3_old), L1_new: TOFLOAT(csvline.L1_new), L2_new:
    TOFLOAT(csvline.L2_new), L3_new: TOFLOAT(csvline.L3_new),
    Vol_old: TOFLOAT(csvline.Vol_old)})

// Create merge relations between droplets
MATCH (n:droplet{Event:"Split"}),(d:droplet{Event:"Split"}),(m:
    droplet{Event:"Merge"})
WHERE n.NewLID = m.NewLID AND d.NewLID = m.OldLID
CREATE (d)-[:Merge]->(n)

// Create merge relations between droplets and core
MATCH (n:droplet{Event:"None"}),(d:droplet{Event:"Split"}),(m:
    droplet{Event:"Merge"})
WHERE n.NewLID = m.NewLID AND d.NewLID = m.OldLID
CREATE (d)-[:Merge]->(n)

// Deletes merge nodes
MATCH (d:droplet)
WHERE d.Event = 'Merge'
DELETE d;

// Creates split relations
MATCH (n:droplet),(d:droplet)
```

```
WHERE n.Event = "Split" and n.OldLID = d.NewLID
CREATE (d)-[:Split]->(n);
```

APPENDIX C

PYTHON NEO4J DRIVER SCRIPT

The below script uses the py2neo library to query a Neo4j graph database from Python. This script was used to analyze the data presented in Fig. 10. First, communication with a graph is established. The graph must be running in Neo4j when the script is executed. Then, a time series is created and result arrays are initialized. Following this, the Cypher queries are executed. To use a Python variable within the Cypher query, a $ must be placed immediately before the variable, then the variable is defined in the second argument in the ".run" function. The ".evaluate()" function ensures that only values from the query are returned, otherwise results appear as a dictionary. Results from the query are then appended to the results lists to be used for further analysis.

```
import numpy as np
from py2neo import Graph
# Using Python Neo4j API "py2neo" to query a graph database
gs = Graph("<Graph URI>",password="<Graph Password>") # Graph must
    be running on the Neo4j platform for this to work. First input
    is graph URI and second is a required password

# Set up time array
dt = 5e-5
time = []
x = np.arange(0,2.65e-3,dt)

# Convert time class from numpy.float64 to python native float
for t in x:
    time.append(t.item())

# Initialize result arrays
nd = [] # total number of droplets
n1 = [] # number of primary droplets
n2 = [] # total number of secondary droplets

# Run Cypher Queries

# Total Droplets
for t in time:
    g = gs.run("MATCH(n:droplet),(c:core) WHERE n.Event_Time <= $t
        RETURN count(n)",t=t).evaluate() # evaluate() returns only
         values from the query
    nd.append(g)
# Primary Droplets
for t in time:
    g = gs.run("MATCH(n:primary),(c:core) WHERE n.Event_Time <= $t
        RETURN count(n)",t=t).evaluate()
```

```
    n1.append(g)
# All secondary
for t in time:
    g = gs.run("MATCH(n:droplet),(c:core) WHERE n.Event_Time <= $t
        AND NOT n:primary RETURN count(n)",t=t).evaluate()
    n2.append(g)
```