Air Entraining Bubbly Flows Driven by Strong Free-Surface Turbulence

by

Declan B. Gaylo

B.S., Webb Institute, 2019
S.M., Massachusetts Institute of Technology, 2021

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2026

Authored by: Declan B. Gaylo
Department of Mechanical Engineering
January 5, 2026

Certified by: Dick K.P. Yue
Philip J. Solondz Professor of Engineering, Thesis supervisor

Accepted by: Nicolas Hadjiconstantinou
Quentin Berg (1937) Professor of Mechanical Engineering
Graduate Officer, Department of Mechanical Engineering

Air Entraining Bubbly Flows Driven by Strong Free-Surface Turbulence

by
Declan B. Gaylo

Submitted to the Department of Mechanical Engineering
on January 5, 2026 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Bubbles beneath turbulent free surfaces are ubiquitous in natural and engineering processes, where predicting their size distribution is of fundamental importance. Examples include air-sea gas exchange and bubble acoustics. The bubble size distribution $N(a)$, where a is radius, is governed by the population balance equation (PBE), which has a term for each mechanism that evolves the bubble population. We consider fragmentation, entrainment, and degassing. These mechanisms are driven by turbulence near the surface, but predicting free-surface turbulence (FST) is a challenge for models, e.g., Reynolds-averaged Navier-Stokes (RANS). We use direct numerical simulation (DNS) to resolve FST. We show that turbulent Froude numbers $Fr_T^2 = \varepsilon/u_{\mathrm{rms}}\mathrm{g} > 0.1$ delineate strong FST, where near-surface turbulence is isotropic. We provide a robust definition of surface layer thickness δ_s, which collapses relevant metrics within the surface layer. For strong FST, free-surface effects are restricted to the surface layer. Towards a surface layer model for RANS, we elucidate the scaling of δ_s and energy flux into the surface layer.

While DNS resolves turbulence, measuring bubble evolution mechanisms is a challenge. We develop Eulerian label advection (ELA) to provide accurate volume-conserving bubble tracking regardless of evolution complexity. ELA allows the first direct measurement of evolution mechanisms in DNS of FST. For fragmentation, we verify that it can be treated as memoryless (assumed by the PBE) and quantify the timescale to reach $N(a) \propto a^{-10/3}$, the equilibrium for fragmentation-dominated bubble populations. From DNS of multiple FST flows, we show the large-bubble entrainment size distribution $I(a)$ scales with Fr_T^6 and $a^{-14/3}$, consistent with a mechanism we describe. We obtain the degassing rate $\Lambda(a)$, which has turbulence-driven and buoyancy-driven regimes with different scalings. We find that FST is degassing, not fragmentation, dominated, and derive the corresponding equilibrium bubble population, $N(a) = I(a)/\Lambda(a)$, which agrees with DNS measurements. Compared to $N(a) \propto a^{-10/3}$, this distinct new equilibrium has two power-law regimes, fewer large bubbles, and is very sensitive to Froude number.

The findings of this thesis contribute to fundamental understanding of strong FST and the size distribution of bubbles within it, and help pave the way for modeling and application of these flows.

Thesis Supervisor: Dick K.P. Yue
Title: Philip J. Solondz Professor of Engineering

Acknowledgments

I would first like to thank my advisor, Professor Dick Yue. His support has been invaluable to my graduate studies at MIT. He has given me the freedom and flexibility to make this work my own, while enabling and encouraging me to present the results at every opportunity. Along the way, he has always provided valuable discussion and guidance, without which I would not be the researcher I am today. Next, I would like to thank Dr. Kelli Hendrickson. She taught me scientific coding and high performance computing, and has always been available for discussions on coding, research, or more generally on how to navigate the PhD. I would also like to thank the rest of my thesis committee, Professor Themistoklis Sapsis and Professor Lydia Bourouiba, who provided valuable advice and encouragement during my work on this thesis.

I would like to thank the rest of the VFRL lab. Mary Thompson for so often going out of her way to support us, Dr. Yuming Liu for his valuable perspective on research, Dr. Grgur Tokić for his advice and friendship, and all the current and past graduate students, especially João Seixas de Medeiros for his friendship and advice and Max Pierce for being a great friend inside and outside of the lab. I would like to thank Will Harris for his friendship and inspiring my interest in acoustics. Outside of research, I would like to thank the MIT Curling Team and all my curling friends, who helped make my time at MIT fun.

I would like to thank my entire family for inspiring me from early on to pursue engineering, especially the amazing stories from my grandfather Benedict. I would like to thank my sister Emma and my parents Mark and Kym for all they have done throughout my life. Finally, Shannon for always being there for me, I can't thank you enough for your love and support.

This work was supported by the Office of Naval Research [grant numbers N00014-20-1-2059, N00014-24-1-2076]. Computational resources were provided by the Department of Defense High Performance Computing Modernization Program, the Lincoln Laboratory Supercomputing Center, and the MIT Office of Research Computing and Data.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Air Entraining Bubbly Flows

Turbulence beneath an air water free surface is present in a large variety of natural and engineered flows. In this configuration, the free surface is affected by both the disturbing force of turbulence and the restoring force of gravity. Surface tension also acts as a restoring force, but for sufficiently large scales its effect is negligible compared to gravity. The ratio of the strength of turbulence to the strength of gravity defines a turbulent Froude number (squared),

$$Fr_T^2 = \frac{u_{\text{rms}}^2}{\text{g}L_T}\,, \tag{1.1}$$

where u_{rms} is the characteristic velocity of the turbulence, $L_T = u_{\text{rms}}^3/\varepsilon$ is the characteristic length scale of the turbulence (where ε is the turbulent dissipation rate), and g is gravitational acceleration. For small Fr_T^2 turbulence is suppressed by the restoring force of gravity, and the free surface remains intact. For large Fr_T^2 turbulence overcomes gravity, and the free surface is broken up (Brocchini & Peregrine, 2001*a*). A highly visible feature of large-Fr_T^2 free-surface flows is the creation of bubbles as the broken free surface entraps air, and these flows are often referred to as self-aerating or (as we do here) air entraining. Figure 1-1 shows examples of air entraining free-surface flows, where the characteristic "white water" indicates the presence of entrained air, i.e., bubbles.

1.1.1 Examples in nature and engineering

One air entraining flow of interest is breaking waves in the ocean (e.g., figure 1-1e). The resulting bubbles significantly increase the total surface area of the interface between air and water, promoting the exchange of gases, including CO_2, between the ocean and the atmosphere (Thorpe, 1982; Wallace & Wirick, 1992; Farmer *et al.*, 1993; Melville, 1996). Additionally, when the bubbles rise to the surface and burst, they create tiny droplets known as sea spray aerosols, which have significant implications on weather prediction (Veron, 2015). Beneath the ocean surface, the bubbles have a significant effect on acoustics, through both sound generation and propagation (Medwin & Beaky, 1989; Lamarre & Melville, 1991; Deane *et al.*, 2013).

(a) Beckton STP, Activated Sludge Tank. John Rostron, CC BY-SA 2.0

(b) Fast-moving waters of Orin Falls. NPS/Alyssa Mattei

(c) USS *Higgins* (DDG 76) operates off the coast of Haiti. U.S. Navy/Adrian White

(d) Bubbly wakes behind boats near Boston. The author

(e) North Pacific storm waves as seen from the MV *Noble Star*. NOAA

Figure 1-1: Examples of air entraining flow in (a) waste water treatment, (b) rivers, (c) ship wakes, (d) boat wakes, and (e) ocean waves. The presence of "white water" indicates entrained air.

Another air entraining flow is in shallow rivers/streams (e.g., figure 1-1b) and dam spillways (Falvey & Ervine, 1988; Chanson, 1996). Here the turbulent boundary layer grows until it reaches the free surface, which, if Fr_T^2 is large, causes air entrainment (Keller *et al.*, 1974; Wilhelms & Gulliver, 2005). Similar to breaking waves, the resulting bubbles promote gas exchange, and here the focus is often on modeling dissolved oxygen to understand the health of the river/stream (Gulliver & Rindels, 1993).

Air entraining flows are also common in engineering. When the goal is to increase air-water gas exchange, flows can be designed to promote air entrainment, such as aeration cascades in water treatment plants (Chanson, 1996). In other flows the goal is to prevent air entrainment, for example liquid-metal nuclear reactors are designed to prevent entrainment of the nodal gas barriers above the liquid metal (Patwardhan *et al.*, 2012). Of particular interest in this work is the air entraining flow around surface (or near-surface) vessels (e.g., figure 1-1c, d). The near-vessel air entrainment creates a prominent cloud of bubbles which extends far into the wake (NDRC, 1946). The larger bubbles in the cloud rise to the surface and create an observable surface slick, and the smaller bubbles make the bubbly wake acoustically detectable far behind the vessel (Trevorrow *et al.*, 1994). These features make predicting the bubbly wake important to the design and operation of naval vessels.

We highlight that in all these examples the size distribution of bubbles is of critical interest. For gas exchange one is interested in the surface area of the bubbles, and the ratio of volume to surface area depends on bubble size. For acoustics, the scattering and absorption properties of a bubble change significantly near its resonant frequency, which depends on bubble size (Medwin & Clay, 1998).

1.1.2 Challenges predicting bubble size distributions

To enable prediction and analysis of systems which depend on bubbly free-surface flows, we need numerical models which can predict the total volume and size distribution of bubbles (Zabaleta *et al.*, 2024). We see two main challenges that prevent current computational fluid dynamics (CFD) tools from accurately predicting the bubble population. The first is modeling turbulence near an air entraining free surface. Due to the large density differences between air and water, closure models for this free-surface turbulence (FST) are a challenge (Brocchini & Peregrine, 2001*b*; Hendrickson & Yue, 2019). Modeling FST is addressed in Chapter 4.

The second challenge is modeling how the population of bubbles evolves in FST. There is a huge difference between the large scales of the flow (e.g., the length of a ship $O(100)$ meters and the smallest bubbles in the flow, $O(100)$ micrometers. This $O(10^6)$ separation in scales makes resolving individual bubbles impossible. For illustration, Castro *et al.* (2016) estimate that a CFD simulation of a ship with a resolution sufficient to resolve individual bubbles would require 10 quadrillion CPU cores, orders of magnitude more cores than humanity has ever manufactured. Thus, for CFD of physical-scale bubbly flow, it is clear we need *statistical* models of how bubble populations evolve in FST, as a function of the local turbulence levels.

In the next section we discuss how the population balance equation (PBE) provides the framework for statistical modeling of bubble evolution. The PBE has been implemented in CFD (e.g. Castro & Carrica, 2013), but accurate bubble population predictions require

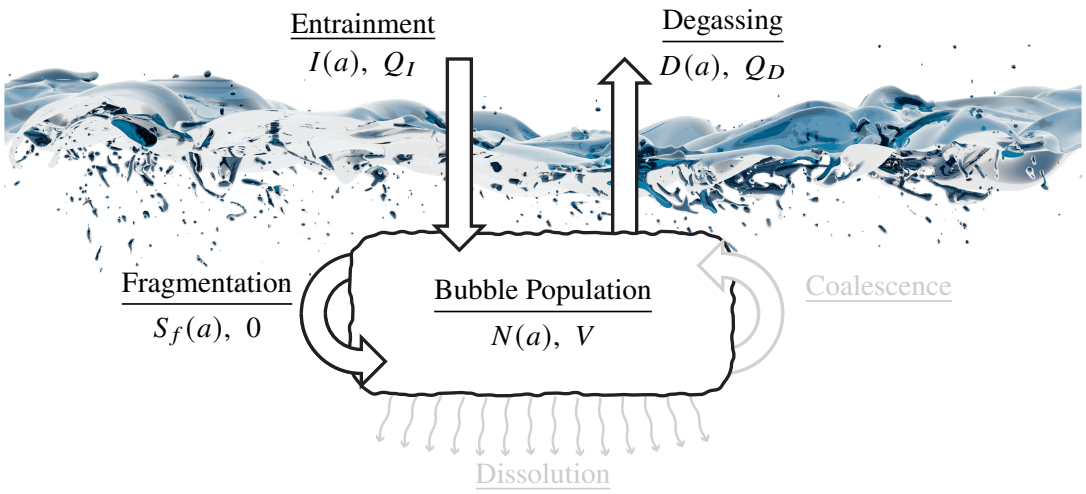


Figure 1-2: Illustration of the effect of the five physical mechanisms on the bubble population, in terms of the bulk bubble size distribution $N(a)$ (see (1.2) for definitions), and the total entrained volume V (see (1.8) for definitions). The rendering comes from DNS of an air entraining FST flow (see §6.4 for details).

accurate models of each of the underlying physical mechanisms that evolve the bubble populations. Some mechanisms, like bubble fragmentation in turbulence, are relatively well understood; however, others, like entrainment by FST, are not. A persistent barrier to understanding these mechanisms is that measuring the individual evolution mechanisms is difficult. This is difficult in experiments where visual access is a challenge, and is even difficult in direct numerical simulation (DNS), despite it providing direct access to the turbulent flow field. In Chapter 3 we develop a new numerical tool to provide robust measurement of individual evolution mechanisms in DNS, and in Chapters 5–7 we use it to gain new insight into the mechanisms in FST and how they affect the overall bubble population.

1.2 Population balance equation for statistical modeling of bubble populations

The bubble population in a region of interest can be described by the bubble size distribution $N(a)$ (dimensions $[1/L]$), where $N(a)\delta a$ is defined to be the number of bubbles of effective radius[1] $[a, a+\delta a]$. The evolution of $N(a)$ is described by a Boltzmann-type PBE,

$$\partial N/\partial t(a) = S_d(a) + S_f(a) + S_c(a) + I(a) - D(a)\,, \tag{1.2}$$

with source terms (dimensions $[1/LT]$) describing each of the five physical mechanisms that evolve the bubble population (Sporleder *et al.*, 2012):

- $S_d(a)$ Dissolution where air dissolves into the surrounding water.

[1]For non-spherical bubbles, the effective radius a is based on the bubble's volume v through $a = (3v/4\pi)^{1/3}$.

- $S_f(a)$ Fragmentation where turbulence breaks a bubble into multiple bubbles.
- $S_c(a)$ Coalescence where multiple bubbles join into one.
- $I(a)$ Entrainment where a bubble is created by entrapment of air at the free surface.
- $D(a)$ Degassing where a bubble bursts at the free surface.

Figure 1-2 illustrates how these five mechanisms evolve the bubble population in air entraining free-surface flow. The PBE describes the evolution of the statistical distribution of the bubble population, so each of the physical processes are also described by a statistical distribution. At a high level, one goal of this work is to elucidate these processes' distributions and, through (1.2), predict the resulting $N(a)$.

1.2.1 Physical mechanisms not considered in this work

We start with the physical processes that will not be investigated in this work: dissolution and coalescence. Avoiding the details from chemistry, we assume a velocity $U_{\text{dissolution}}$ which characterizes the rate at which air dissolves into the surrounding water per unit interfacial area. A bubble's volume scales like $\propto a^3$ and interfacial area like $\propto a^2$, which leads to a characteristic timescale $T_{\text{dissolution}} \propto a/U_{\text{dissolution}}$ for dissolution to affect the size of a bubble. We will consider bubble evolution of sufficiently large bubbles over sufficiently short timescales T such that $T \ll T_{\text{dissolution}}$ and the effect of dissolution is negligible. Using bubbly ship wakes as an example, we are interested in the bubble evolution surrounding and immediately behind ship ($T \sim 10$ seconds), rather than far behind the where dissolution becomes relevant ($T \sim 10$ minutes) (Trevorrow *et al.*, 1994). Our DNS models air and water as immiscible (see Chapter 2), in effect setting $U_{\text{dissolution}} = 0$ exactly.

Coalescence happens when multiple bubbles collide and then merge. Therefore, the frequency of coalescence is primarily a function of the density of bubbles. The density of bubbles can be described by the void fraction, which gives the average proportion of volume that is occupied by air (as opposed to water). We will consider bubble evolution in flows with moderate void fractions such that coalescence does not significantly affect the bubble population. To address very large void fractions, incorporating coalescence is an area for future work (see §8.2).

1.2.2 Physical mechanisms considered in this work

For flows with negligible dissolution and coalescence, the bubble evolution is described by the PBE

$$\partial N/\partial t(a) = S_f(a) + I(a) - D(a)\,. \tag{1.3}$$

Fragmentation, entrainment, and degassing will be the focus of Chapters 5, 6, and 7 respectively, and how they interact through (1.3) informs our understanding of $N(a)$. To describe their interactions, it is first useful to introduce how each of these terms are modeled. Fragmentation can be split into two terms,

$$S_f(a) \equiv S_f^+(a) - S_f^-(a)\,, \tag{1.4}$$

where $S_f^-(a)$ describes the destruction of bubbles of radius a by fragmentation and $S_f^+(a)$ the creation of bubbles of radius a as daughters of the fragmentation of larger bubbles. For more

details on the creation term see Chapter 5. For the destruction term, we model individual fragmentation events as independent and memoryless (as verified in Chapter 5), in which fragmentation is a Poisson process. This gives

$$S_f^-(a) = \Omega(a)N(a)\,, \tag{1.5}$$

where $\Omega(a)$ is the fragmentation rate (dimensions $[1/T]$). Similarly, if the degassing statistics of individual bubbles are independent and memoryless, we can define a degassing rate $\Lambda(a)$ (dimensions $[1/T]$) and

$$D(a) = \Lambda(a)N(a)\,. \tag{1.6}$$

We can now split (1.3) into two groups

$$\partial N/\partial t(a) = \left[I(a) + S_f^+(a)\right] - \left[\Lambda(a) + \Omega(a)\right]N(a)\,, \tag{1.7}$$

where the first group is terms describing the creation of bubbles of radius a, and the second group described destruction of bubbles of radius a. The destruction terms are linear with $N(a)$.

1.2.3 Equilibrium and non-equilibrium bubble populations

In addition to $\partial N/\partial t(a)$ described by the PBE, it can also be useful to consider the change in the total volume of bubbles, $\mathrm{d}V/\mathrm{d}t$ where $V = (4\pi/3)\int N(a)a^3\,\mathrm{d}a$. Integrating (1.3) using the same bubble volume weighted integral,

$$\mathrm{d}V/\mathrm{d}t = Q_I - Q_D\,, \tag{1.8}$$

where $Q_I \equiv (4\pi/3)\int I(a)a^3\,\mathrm{d}a$ is the flux of air volume from above to beneath the free surface (entrainment flux) and $Q_D \equiv (4\pi/3)\int D(a)a^3\,\mathrm{d}a$ is the flux from beneath to above the free surface (degassing flux). Fragmentation only moves air between bubble sizes, so it does not contribute to $\mathrm{d}V/\mathrm{d}t$. As opposed to decaying bubble populations with negligible entrainment ($\mathrm{d}V/\mathrm{d}t < 0$) such as during the quiescent period of breaking waves (Deane & Stokes, 2002) or in the far wake of a ship, our interest is the behavior of air entraining flows where Q_I is relevant ($\mathrm{d}V/\mathrm{d}t \geq 0$).

For air entraining flows, we can classify the bubble population as being in either a non-equilibrium regime where $\mathrm{d}V/\mathrm{d}t > 0$ and $\partial N/\partial t(a) \neq 0$, or an equilibrium regime where $\mathrm{d}V/\mathrm{d}t = 0$ and $\partial N/\partial t(a) = 0$. In this work we obtain a new equilibrium solution to the PBE, and, by elucidating the individual terms of (1.7), we also inform modeling of non-equilibrium bubble populations.

1.3 Equilibrium bubble populations in plunging breaking waves

As discussed in §1.1, one flow where the bubble population is of interest is breaking waves in the ocean. For breaking waves, Garrett *et al.* (2000) predicted that the population of large bubbles follows $N(a) \propto a^{-10/3}$. "Large" refers to bubbles larger than the Hinze scale a_H (Hinze, 1955). Deane & Stokes (2002) observed this $-10/3$ power law in laboratory experiments of plunging breaking waves. Since then, it has been observed by many in both laboratory experiments and numerical simulations of breaking waves (see review by Deike, 2022). In this section we review the modeling assumptions used by Garrett *et al.* (2000). While the resulting $N(a > a_H) \propto a^{-10/3}$ describes bubble populations beneath plunging breaking waves, this work will show that this distribution does not apply universally to air entraining flows (Chapter 7).

1.3.1 Derivation of -10/3 equilibrium solution

Garrett *et al.* (2000) consider a population of bubbles where very large bubbles are entrained and then successively fragment into smaller and smaller bubbles. Garrett *et al.* (2000) call this a fragmentation cascade and note analogies to the energy cascade description of turbulence in the Kolmogorov inertial subrange. We can also describe fragmentation cascades in terms of the PBE. Bubbles in the cascade (bubbles sized between the large entrained bubbles and the Hinze scale) are only affected by fragmentation:

$$\partial N/\partial t(a) = S_f(a)\,. \tag{1.9}$$

At equilibrium $\partial N/\partial t(a) = 0$ and, recalling (1.4), we have $0 = S_f^+(a) - S_f^-(a)$. In words, the rate at which bubbles of radius a are created as the daughters when larger bubbles fragment is equal to the rate at which bubbles of radius a fragment. With a model of fragmentation, this can be solved to obtain the associated $N(a)$.

Fragmentation is governed by the balance between the disturbing force of turbulence on a bubble and the restoring force of surface tension. The ratio between the two is given by the bubble Weber number

$$We_B = \frac{2\varepsilon^{2/3}(2a)^{5/3}}{(\sigma/\rho_w)}\,, \tag{1.10}$$

where ε is the turbulent dissipation rate, a is the radius of the parent bubble, σ is the surface-tension coefficient, and ρ_w the density of water. The Hinze scale is defined as the bubble Weber number We_H (and corresponding radius a_H) below which surface tension largely prevents fragmentation (Hinze, 1955). For $We_B \gg We_H$ (i.e., $a \gg a_H$), fragmentation is unaffected by surface tension, and (as a result of the scaling of turbulence within the Kolmogorov inertial subrange) the fragmentation rate follows (Martínez-Bazán *et al.*, 1999*a*)

$$\Omega(a \gg a_H) \propto \varepsilon^{1/3} a^{-2/3}\,. \tag{1.11}$$

This gives the destruction term, $S_f^-(a) = \Omega(a)N(a)$. For the creation term, Garrett *et al.* (2000) assume a simple model where all bubbles fragment into exactly m identically sized

daughters. This assumption gives $S_f^+(a) = m^{2/3}\Omega(m^{1/3}a)N(m^{1/3}a)$ (see derivation by Gaylo *et al.*, 2021). If we assume a power-law solution to the large bubble population,

$$N(a) \propto a^\beta \quad \text{for } a \gg a_H\,, \tag{1.12}$$

then $0 = S_f^+(a) - S_f^-(a)$ simplifies to

$$0 = \Omega(a)N(a)\left[m^{\beta/3+10/9} - 1\right]\,, \tag{1.13}$$

which is solved by $\beta = -10/3$.

1.3.2 Evaluation of modeling assumptions

The wide agreement on $\beta = -10/3$ for plunging breaking waves (Deike, 2022) suggests that Garrett *et al.* (2000) provide a good model of the bubble population in that flow. However, we will show in this work that it does not apply universally to bubble populations in air entraining flows. To start, we consider the assumptions that went into the derivation in §1.3.1.

Fragmentation dominance

The key assumption in §1.3.1 is that fragmentation is dominant over degassing, $S_f(a) \gg D(a)$, leading to the simplified PBE (1.9). In Chapter 7 we will study air entraining FST and show that it leads to a degassing-dominated bubble population where $S_f(a) \ll D(a)$, creating a distinct power law $\beta \neq -10/3$.

Locality in fragmentation cascades

There can only be a fragmentation cascade if the daughter bubbles of fragmentation are only slightly smaller on average than the parent bubble, a property called locality (Chan *et al.*, 2021*b*). If bubble fragmentation were non-local, a fragmentation event of a parent bubble $a \gg a_H$ would often produce daughter bubbles of radii $a < a_H$, which would not further fragment. In §1.3.1, by applying (1.11) to the creation and destruction term, we assumed that it was possible for both daughter and parent to be larger than the Hinze scale ($a \gg a_H$ and $m^{1/3}a \gg a_H$). This is where the assumption of locality enters.

Chan *et al.* (2021*c*) study the fragmentation cascade in simulations of plunging breaking waves and measure locality, confirming that fragmentation is strongly local. In Chapter 5 we study fragmentation in simulations of homogeneous isotropic turbulence and measure the "speed" that air moves through the fragmentation cascade. We find this speed is finite, which also confirms locality.

Entrainment only at large scales

In §1.3.1 we followed Garrett *et al.* (2000) and assumed that bubbles are entrained only at some large scale, below which entrainment is negligible, $S_f(a) \gg I(a)$. Gaylo *et al.* (2021)

address the case where entrainment is not restrained to large bubbles. They consider the PBE

$$\partial N/\partial t(a) = S_f(a) + I(a)\,, \tag{1.14}$$

and assume a cut-off power law for the entrainment size distribution,

$$I(a) \propto \begin{cases} a^\gamma & a < a_{\max} \\ 0 & \text{otherwise} \end{cases}. \tag{1.15}$$

Fragmentation is modeled the same as in §1.3.1. By solving (1.14) for $\partial N/\partial t(a) = 0$, they obtain the power law solution

$$\beta = \frac{\gamma + 4}{1 - (a_{\max}/a)^{\gamma+4}} - 10/3\,, \tag{1.16}$$

For $a \ll a_{\max}$, there are two regimes of fragmentation-dominated bubble populations based on the entrainment power law γ:

$$\beta = \begin{cases} -10/3 & \gamma \geq -4 \\ \gamma + 2/3 & \gamma < -4 \end{cases}. \tag{1.17}$$

For weak entrainment ($\gamma \geq -4$), most volume is entrained in the largest bubbles, and the population still obtains $\beta = -10/3$. For strong entrainment ($\gamma < -4$), most volume is entrained in the smallest bubbles and entrainment rather than the fragmentation cascade set β. In Chapter 6 we show that entrainment of large bubbles in FST is strong ($\gamma = -14/3$ implying $\beta = -4$); however, we also find that degassing is dominant rather than fragmentation so (1.17) does not apply.

1.4 Thesis Outline

While the equilibrium bubble size distribution is relatively well understood for super Hinze-scale bubbles in plunging breaking waves ($\beta = -10/3$), section 1.1.1 illustrates that there are many other air entraining free-surface flows where predicting the bubble size distribution is of interest. Recent observations of different large-bubble power laws in simulations of the wake behind a dry transom stern ($\beta \in [-5, -4]$) (Hendrickson *et al.*, 2019) highlight that $\beta = -10/3$ is not necessarily a universal description of bubble populations in air entraining free-surface flows. The purpose of this work is the development of more general models of bubble populations which, through the PBE (1.3), consider all the relevant physical mechanisms in air entraining flow.

Different air entraining flows may have certain specific mechanisms driven by the large-scale flow structure (e.g., entrapment of a cavity by a plunging breaking wave (Deike *et al.*, 2016; Chan *et al.*, 2021*a*; Gao *et al.*, 2021)); however, strong turbulence beneath the free surface is a common feature of many air entraining flows Brocchini & Peregrine (2001*a*). Thus, for generality, our focus will be on the mechanisms driven by this free-surface turbulence (FST). Using direct numerical simulation (DNS) of air entraining FST, along

with a new numerical measurement technique, we elucidate these mechanisms and how they scale, particularly with turbulent Froude number Fr_T^2. Using these insights, we discover a new regime for the equilibrium bubble size distribution of super capillary-scale bubbles in FST, which is quantitatively distinct from $\beta = -10/3$ for plunging breaking waves.

The thesis is organized as follows,

- **Chapter 2: Numerical Methods for Incompressible Turbulent Bubbly Flow**

  The numerical methods used to perform DNS of bubbly flow are reviewed. The governing Navier-Stokes equations are introduced, and a second-order finite-volume solver, MPF-Solver, is described. The conservative volume of fluid (cVOF) method (Weymouth & Yue, 2010) is described in detail, as the bubble tracking algorithm developed in Chapter 3 is closely tied to cVOF. Also relevant to bubble tracking, the accuracy of different methods for identifying and labeling bubbles is evaluated.

- **Chapter 3: ELA Method for Volume-Conservative Bubble Tracking**

  A previous barrier to quantifying bubble evolution mechanisms is that they are difficult to measure, even in DNS. To solve this, we develop Eulerian label advection (ELA) to track the evolution of bubbles. As opposed to Lagrangian methods, the Eulerian nature makes it robust, independent of the complexity of the bubble evolution. The method inherits the volume-conservation of cVOF, meaning all air is tracked. The output of ELA is a matrix-based description of bubble evolution, from which individual bubble evolution mechanisms are easily accessible. ELA allows, for the first time, direct measurement of bubble evolution mechanisms in air entraining FST (used in Chapter 5–7).

- **Chapter 4: Characterizing the Surface Layer of Strong FST**

  The near-surface turbulence in air entraining FST is characterized. $Fr_T^2 = 0.1$ is found to be the critical value above which gravity effects are weak enough for near-surface turbulence to be nearly isotropic. For $Fr_T^2 > 0.1$ (strong FST), we show the effects of the free surface are constrained to a surface layer of thickness $\delta_s/L_T \propto Fr_T^2$. We show that, even at these large Fr_T^2, the free surface is largely intact and (compared to bubbles or droplets) dominates near-surface dynamics. Scaling by δ_s and turbulence properties measured at the bottom of the surface layer collapses relevant turbulence metrics across a wide range of Fr_T^2. We discuss how these results inform reduced-order modeling of FST.

- **Chapter 5: Bubble Fragmentation in HIT**

  The fragmentation term $S_f(a)$ is quantified using ELA in DNS of bubble fragmentation in homogeneous isotropic turbulence (HIT). We identify three fundamental timescales which characterize the statistics relevant to the PBE. In addition to the often-studied bubble lifetime τ_ℓ, these are the relaxation time τ_r and convergence time τ_c. We find $\tau_r \ll \tau_\ell$, which validates that fragmentation events can be modeled as independent and memoryless, a core assumption of the PBE. τ_c provides the characteristic time for fragmentation-dominated bubble populations to obtain the $\beta = -10/3$ equilibrium solution and provides a new constraint on fragmentation models.

- **Chapter 6: Bubble Entrainment in FST**

  Through a mechanistic argument, we obtain the entrainment size distribution $I(a)$ for large (super capillary scale) bubbles entrained by FST. We find that entrainment scales with bubble radius $a^{-14/3}$ and Fr_T^6, independent from weak surface tension effects. The sensitivity to Froude number is significantly stronger than previously thought (Yu *et al.*, 2020). We quantify $I(a)$ using ELA in DNS of a flow which isolates entrainment by FST, and a canonical free-surface shear flow, characteristic of the flow behind the transom of a ship. In both flows, the agreement with our model is very strong ($R^2 = 0.990$ and $R^2 = 0.891$ respectively). We also find evidence of $I(a) \propto a^{-14/3}$ in previous open channel flow experiments (Wei *et al.*, 2019), highlighting the generality of the FST entrainment mechanism we describe.

- **Chapter 7: Bubble Degassing in FST**

  Through a mechanistic argument, we obtain the degassing rate $\Lambda(a)$, which has two regimes split by a_Λ. One where bubble rise is driven by turbulence ($a < a_\Lambda$) and one where it is driven by buoyancy ($a > a_\Lambda$). We measure $\Lambda(a)$ using ELA in the same DNS of free-surface shear flow as Chapter 6 and obtain a good agreement with our model ($R^2 = 0.761$), including the regime change at a_Λ.

  By quantifying degassing, we show that bubble populations in free-surface shear flow are dominated by degassing rather than fragmentation ($\Lambda(a) \gg \Omega(a)$). Through the PBE (1.3), we show degassing dominance leads to an equilibrium solution $N(a) = I(a)/\Lambda(a)$, distinct from the solution discussed in section 1.3. Using the $I(a)$ from Chapter 6 and $\Lambda(a)$ from this chapter, without introducing any additional fitting parameters, the degassing-dominated bubble population we predict agrees well with the $N(a)$ we measure in the free-surface shear flow ($R^2 = 0.849$). The proprieties of the degassing-dominated bubble population make it easily distinguishable from fragmentation-dominated bubble populations.

Chapter 2

Numerical Methods for Incompressible Turbulent Bubbly Flow

The first two sections of this chapter describe `MPFSolver`, the numerical solver that is used throughout this work to perform DNS of two-phase incompressible flows. Section 2.1 provides a general overview of the solver, and section 2.2 focuses on volume of fluid (VOF) method used to capture the two phases (air and water). Finally, section 2.3 describes methods for identifying connected regions of air, i.e., bubbles, based on the VOF description. In addition to providing a description of how the simulations throughout this thesis are performed, an aim of this chapter is to introduce the numerical concepts on which the new bubble-tracking algorithm presented in Chapter 3 is built.

2.1 Direct numerical simulation

2.1.1 Governing equations

For the flows of interest in this work, it is assumed that both fluids are incompressible Newtonian fluids. Noting that density ρ and viscosity μ at a point in space depends on which phase is present, the governing equation is the incompressible Navier-Stokes equation (N-S), written in single-fluid form but with variable density and viscosity:

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0\,, \tag{2.1a}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}\mathbf{u} = -\frac{1}{\rho}\boldsymbol{\nabla} p + \frac{1}{\rho}\boldsymbol{\nabla} \cdot \boldsymbol{\tau} - \mathrm{g}\hat{\mathbf{e}}_z + \sigma\kappa\delta_s\hat{\mathbf{n}} + \mathbf{f}\,, \tag{2.1b}$$

where $\mathbf{u}$ is the fluid velocity, p is the pressure, $\boldsymbol{\tau} = \mu\left(\boldsymbol{\nabla}\mathbf{u} + \boldsymbol{\nabla}\mathbf{u}^{\mathsf{T}}\right)$ is viscous stress tensor, g is gravitational acceleration (the positive-z direction is up), and $\mathbf{f}$ denotes any additional applied forces (see section 2.1.5). For surface tension σ is the surface tension coefficient, δ_s is the interfacial Dirac delta function, $\hat{\mathbf{n}}$ is the interface normal vector, and κ is the interface curvature.

For the flows of interest in this work, we will treat the two phases as immiscible. This

means the mixture can be described by a binary field called the fluid color function,

$$c(\mathbf{x}, t) \equiv \begin{cases} 1 & \text{if} \quad \mathbf{x} \in \text{`dark' fluid} \\ 0 & \text{if} \quad \mathbf{x} \in \text{`light' fluid} \end{cases} . \tag{2.2}$$

In incompressible flow the volume of each fluid is conserved, so the color function is also conserved. This leads to a third governing equation,

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla} c = 0 \, . \tag{2.3}$$

`MPFSolver` uses the convention that the 'dark' fluid is water and the 'light' fluid is air. Based on the color function, we can express the density and viscosity as

$$\rho = \rho_a (1 - c) + \rho_w c \, , \tag{2.4a}$$

$$\mu = \mu_a (1 - c) + \mu_w c \, . \tag{2.4b}$$

Here ρ_w and μ_w are the values for water and ρ_a and μ_a are for air. We note that the choice of which fluid is 'dark' versus 'light' is arbitrary and all the numerical methods discussed here are agnostic to this choice. In section 2.3 when describing bubble identification and Chapter 3 when describing bubble tracking, it is more convenient to use the convention that $c = 1$ is air and $c = 0$ is water.

Without change of notation, we nondimensionalize the governing equations by a characteristic velocity scale U, length scale L, and the properties of water ρ_w and μ_w. (2.1a) and (2.3) remain unchanged, and (2.1b) becomes

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla} \mathbf{u} = -\frac{1}{\rho} \boldsymbol{\nabla} p + \frac{1}{\rho Re} \boldsymbol{\nabla} \cdot \boldsymbol{\tau} - \frac{\hat{\mathbf{e}}_z}{Fr^2} + \frac{\kappa \delta_s}{We} \hat{\mathbf{n}} + \mathbf{f} \, , \tag{2.5}$$

where nondimensionalized density and viscosity,

$$\rho = \lambda (1 - c) + c \, , \tag{2.6a}$$

$$\mu = \eta (1 - c) + c \, , \tag{2.6b}$$

replace (2.4), with the ratios defined $\lambda \equiv \rho_a / \rho_w$ and $\eta \equiv \mu_a / \mu_w$. Unless otherwise noted, we use $\lambda = 0.00123$ and $\eta = 0.0159$, characteristic of air and water. In (2.5), the parameters that describe a flow are Reynolds number $Re = UL/\nu_w$ (where $\nu_w = \mu_w / \rho_w$), Froude number (squared) $Fr^2 = U^2 / \mathrm{g}L$, and Weber number $We = U^2 L / (\sigma / \rho_w)$.

For calculations, we separate the pressure into a pseudo hydrostatic component and a pseudo dynamic component, $p = p_h + p_d$. The pseudo hydrostatic component is defined in reference to the top of the domain in z,

$$p_h(x, y, z, t) \equiv \frac{1}{Fr^2} \int_z^{Z_{max}} \rho(x, y, z', t) \, \mathrm{d}z' \, . \tag{2.7}$$

We say "pseudo" here because ρ is a function of time so p_h is not actually static. Still, this

separation is useful numerically because $\boldsymbol{\nabla} p_h$ cancels out the gravity term in the z-component of (2.5) such that the remaining $\boldsymbol{\nabla} p_d$ term does not have a strong anisotropy.

In summary, two-phase incompressible flow is described by velocity $\mathbf{u}$, color function c, and pseudo dynamic pressure p_d and is governed by the three equations,

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0\,, \tag{2.8a}$$

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla} c = 0\,, \tag{2.8b}$$

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{1}{\rho[c]} \boldsymbol{\nabla} p_d + \mathbf{f}_{\mathrm{RHS}}[\mathbf{u}, c]\,. \tag{2.8c}$$

For conciseness, we group terms of (2.5) that will be treated explicitly into an acceleration term $\mathbf{f}_{\mathrm{RHS}}$ which we note is a function of $\mathbf{u}$ and c. Density is a function only of c.

2.1.2 Temporal discretization

We use a two-stage Runge-Kutta method described by Dommermuth *et al.* (2004) to provide a second-order in time discretization. For each stage, $\mathbf{f}_{\mathrm{RHS}}$ is found explicitly and a pressure projection method is used to calculate p_d. For a given velocity field $\mathbf{u}^k$ and color function field c^k at time t^k we seek the velocity field $\mathbf{u}^{k+1}$ and color function field c^{k+1} at time $t^{k+1} = t^k + \Delta t$. For the two-stage Runge-Kutta method, this is done over two iterations. For the first iteration: a predictor step

$$\mathbf{u}^{*k+1/2} = \mathbf{u}^k + \Delta t\ \mathbf{f}_{\mathrm{RHS}}\left[\mathbf{u}^k, c^k\right]\ ; \tag{2.9a}$$

accounts for the explicit terms; a Poisson equation

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho[c^k]} \boldsymbol{\nabla} {p_d}^{k+1/2} \right) = \frac{1}{\Delta t} \boldsymbol{\nabla} \cdot \mathbf{u}^{*k+1/2}\,, \tag{2.9b}$$

is solved to find a dynamic pressure ${p_d}^{k+1/2}$ which enforces (2.8a); The velocity is corrected by the new pressure

$$\mathbf{u}^{k+1/2} = \mathbf{u}^{*k+1/2} - \Delta t\ \frac{1}{\rho[c^k]} \boldsymbol{\nabla} {p_d}^{k+1/2}\,; \tag{2.9c}$$

and the color function is updated

$$c^{k+1/2} = cVOF\left[\mathbf{u}^k, c^k\right]\ . \tag{2.9d}$$

For the second iteration, the same steps are repeated:

$$\mathbf{u}^{*k+1} = \frac{\mathbf{u}^k + \mathbf{u}^{k+1/2}}{2} + \frac{\Delta t}{2}\ \mathbf{f}_{\mathrm{RHS}}\left[\mathbf{u}^{k+1/2}, c^{k+1/2}\right]\ , \tag{2.10a}$$

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho[c^{k+1/2}]} \boldsymbol{\nabla} {p_d}^{k+1} \right) = \frac{2}{\Delta t} \boldsymbol{\nabla} \cdot \mathbf{u}^{*k+1}\,, \tag{2.10b}$$

$$\mathbf{u}^{k+1} = \mathbf{u}^{*k+1} - \frac{\Delta t}{2} \frac{1}{\rho[c^{k+1/2}]} \boldsymbol{\nabla} {p_d}^{k+1} , \tag{2.10c}$$

$$c^{k+1} = cVOF \left[\frac{\mathbf{u}^k + \mathbf{u}^{k+1/2}}{2}, c^k \right] . \tag{2.10d}$$

For (2.9d) and (2.10d), the (discretized) color function field is updated using the conservative Volume of Fluid (cVOF) method (Weymouth & Yue, 2010), which is discussed in section 2.2. Note that the pressure correction step is done before updating the color function field. This ensures that the two terms on the right side of the momentum equation (2.8c) are calculated using the same density field.

The time step Δt is chosen dynamically based on the criteria described by Campbell (2014, §5.3.4), with two exceptions. First, for the vicious criteria, we use

$$\Delta t < \frac{1}{6} Re \ \min \left[\Delta x_d\right]^2 \ (\rho/\mu) \tag{2.11}$$

where the constant $1/6$ comes from Tryggvason *et al.* (2011, §3.1).[1] The large difference in μ and ρ between air and water coupled with numerical interpolation means that special care must be taken with how ρ/μ is calculated near fluid interfaces. In Appendix A we derive a definition of $\tilde{\nu} = \mu/\rho$ which ensures (2.11) guarantees linear stability. Second, when surface tension is modeled, an additional criteria,

$$\Delta t < \sqrt{\frac{1}{8\pi} We \ (\lambda + 1) \ \left(\min \left[\Delta x_d\right]\right)^3} , \tag{2.12}$$

is also considered. Generally, we find the dynamic selection of Δt is driven by the cVOF Courant restriction,

$$\Delta t \sum_{d=1}^{\mathcal{N}} \left| \frac{u_d}{\Delta x_d} \right| \leq C , \tag{2.13}$$

where $\mathcal{N} = 3$ is the number of dimensions and $C = 1/2$ (Weymouth & Yue, 2010).

2.1.3 Spatial discretization

We use a second-order finite-volume approach using the marker-and-cell (MAC) staggered-grid method (Harlow & Welch, 1965), where scalar quantities are described at cell centers and the velocity components are described at the respective cell-face centers. This is done on three-dimensional Cartesian grids of size $N_i \times N_j \times N_k$. Figure 2-1 provides a two-dimensional illustration of the MAC staggered grid. Following the standard finite-volume approach, the integral of a quantity within a discrete control volume defines the cell-centered value. For the color function field, this defines the VOF field,

$$f_{ijk}(t) \equiv \frac{\int_{\Omega_{ijk}} c(\mathbf{x}, t) \ \mathrm{d}V}{\Delta \Omega_{ijk}} , \tag{2.14}$$

[1] Campbell (2014) uses $3/14$ intended for a two-step Adams-Bashforth scheme.

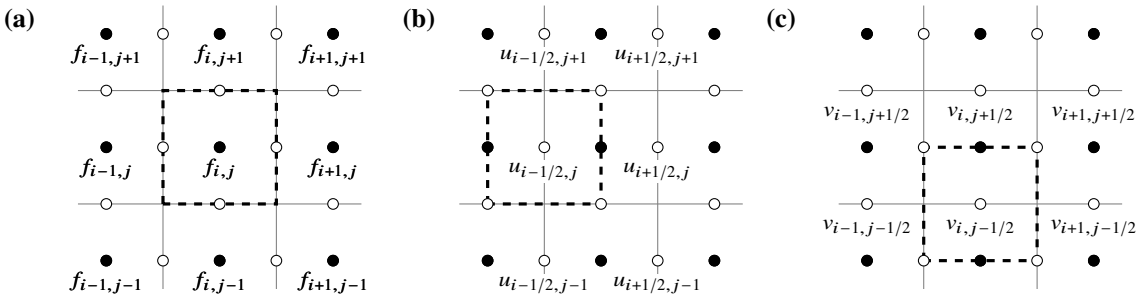


Figure 2-1: Two-dimensional illustration of MAC staggered-grid mesh, where the dashed box indicates the control volume for the: (a), pressure and VOF grid; (b), u-velocity grid; and (c), v-velocity grid.

where Ω_{ijk} is the region of each cell in the density grid with volume $\Delta\Omega_{ijk} \equiv \int_{\Omega_{ijk}} \mathrm{d}V$. Section 2.2 discusses how careful integration of (2.8b) over Ω_{ijk} obtains a volume-conservative scheme to advance f in time (Weymouth & Yue, 2010). Performing the same integration on (2.6), we obtain the (non-dimensional) fluid properties for each cell in the density grid,

$$\rho_{ijk} = \lambda(1 - f_{ijk}) + f_{ijk}\,, \tag{2.15a}$$

$$\mu_{ijk} = \eta(1 - f_{ijk}) + f_{ijk}\,. \tag{2.15b}$$

In some simulations it is useful to smooth the discontinuities in fluid properties, in which case we use the filter described by Tryggvason *et al.* (2011, §7.1.4) to generate a smoothed VOF field, $\tilde{f}_{ijk}$, which is used in place of f_{ijk} in (2.15).

For u, the velocity in the x-direction, the grid is staggered in x (see figure 2-1). The velocity at the center of this staggered control volume is

$$u^k_{i-1/2\,jk} = \frac{\int_{\Omega_{i-1/2\,jk}} u^k(\mathbf{x})\ \mathrm{d}V}{\Delta\Omega_{i-1/2\,jk}}\,. \tag{2.16}$$

The same can be defined for v, the velocity in the y direction, and w, the velocity in the z direction. By integrating over these cells and applying divergence theorem, discrete forms of (2.9a), (2.9c), (2.10a), and (2.10c) can be derived that are second-order in space (Tryggvason *et al.*, 2011, Chapter 3). For the advection term in $\mathbf{f}_{\mathrm{RHS}}$, we use the second-order centered scheme.

For the viscous term in $\mathbf{f}_{\mathrm{RHS}}$, a challenge for two-phase flow is calculating the viscosity. The diagonal terms of $\boldsymbol{\tau}$ need the cell centered value μ_{ijk}, which is easily available from MAC using (2.15b). However, for the off-diagonal terms there is not a consistent second-order method for two-phase flows using MAC (Tryggvason *et al.*, 2011). Campbell (2014) uses a harmonic mean approximation which is exact when the interface normal aligns with a specified Cartesian axis; however, this approach cannot be applied to bubbly flows where the interface normal can be in any orientation. Instead, we follow Yu (2019) and use a simple

arithmetic mean,

$$\mu_{i+1/2\,j+1/2\,k} = \frac{\mu_{i\,j\,k} + \mu_{i+1\,j\,k} + \mu_{i\,j+1\,k} + \mu_{i+1\,j+1\,k}}{4}\,. \tag{2.17}$$

The effect of this interpolation method on the viscous timestep restriction (2.11) is discussed in Appendix A. Tryggvason *et al.* (2011) note that the arithmetic mean approach is more robust than the harmonic mean, but, in effect, increases the viscosity in a region near the interface.

For the surface tension term in $\mathbf{f}_{\text{RHS}}$, we use a continuous surface force method (Brackbill *et al.*, 1992), as implemented by Yu (2019). To calculate interface curvature, we first interpolate the VOF field to the respective velocity grid. For each cell, we calculate the interface height in the 3×3 stencil of neighboring cells perpendicular to the dominate interface direction in that cell, as determined by interface reconstruction described in section 2.2.1. For each cell in the stencil, we first attempt to calculate the interface height using the forward and backward search described by Popinet (2009, Algorithm 4), but fall back to a sum over 7 cells when this fails. Based on the 9 height functions in the 3×3 stencil, the curvature is calculated using a standard second-order finite-difference scheme (Francois *et al.*, 2006, eq. 33–36). Because we calculate curvature after interpolating the VOF field to velocity grid, no interpolation of the curvature is necessary (cf. Francois *et al.*, 2006, eq. 37).

Finally, we have (2.9b) and (2.10b) to calculate the pseudo dynamic pressure p_d. We highlight that this is the only step that is solved implicitly. Tryggvason *et al.* (2011) provide a description of the discrete form of (2.9b) and (2.10b) for MAC as well as the accompanying boundary conditions. The result is a variable coefficient Poisson equation, which we use the `hypre` library[2] (Falgout *et al.*, 2006) to solve using the generalized minimal residual (GMRES) method .

2.1.4 Minimum grid size

To accurately resolve the physics described by (2.1), it is essential that the grid size Δ is small enough to capture the necessary scales.

Resolving turbulence

In this work, we directly model the flow without any turbulence closure models. This DNS approach requires us to capture the smallest scale of turbulence, the Kolmogorov microscale,

$$\eta_T \sim \varepsilon^{-1/4} Re^{-3/4} \tag{2.18}$$

where ε is the turbulent dissipation rate. For `MPFSolver`, previous convergence studies by Yu (2019) and an additional convergence study here focused specifically on bubble evolution (see Appendix G) show that a grid size

$$\eta_T/\Delta > 1 \tag{2.19}$$

[2] `https://computing.llnl.gov/projects/hypre-scalable-linear-solvers-multigrid-methods`

is sufficient to resolve the turbulence and its effects on bubble evolution. While DNS fully resolves turbulent flows described by N-S without any additional modeling, computational limitations on the number of grid cells mean that, with current high-performance computing, we are typically limited to $Re \lesssim 10^3$.

Resolving surface tension

For simulations that include surface tension (finite *We*), the cell Weber number,

$$We_\Delta = \frac{u_{\text{rms}}^2 \Delta}{4\pi(\sigma/\rho_w)}\,, \tag{2.20}$$

estimates the ratio between the grid and the minimum characteristic radius of curvature of an interface deformed by inertial turbulence. Popinet (2018) suggest

$$We_\Delta \lesssim 1 \tag{2.21}$$

ensures surface tension forces are resolved by the grid. Alternatively, some compare the ratio of the grid size to the Hinze scale $a_H \propto \varepsilon^{-2/5}(\sigma/\rho_w)^{3/5}$ (see (5.1) in Chapter 5) and suggest

$$\Delta/a_H \lesssim 1\,. \tag{2.22}$$

This ratio is related to (2.20) through

$$\Delta/a_H \propto We_\Delta^{3/5}\,(\Delta/L_T)^{2/5} \tag{2.23}$$

where $L_T = u_{\text{rms}}^3/\varepsilon$ is the characteristic length scale of the turbulence. The ratio of $L_T/\eta_T \propto Re^{3/4}$, meaning that, for a given *Re*, Δ/L_T in (2.22) is usually fixed by the requirement to resolve the Kolmogorov scale (2.19). In practice, we find (2.22) is usually redundant to (2.21).

2.1.5 Linear forcing

To simulate quasi-steady turbulence, it is necessary to inject energy into the flow to balance the energy lost to heat through dissipation. A common method to do this is linear forcing (Lundgren, 2003; Rosales & Meneveau, 2005). In (2.1b) we define a body force linearly proportional to the turbulent fluctuations

$$\mathbf{f} \equiv A\mathbf{u}''\,, \tag{2.24}$$

where, to guarantee the forcing does not impart a net change in momentum, $\mathbf{u}''$ is from Favre averaging,

$$\mathbf{u}'' \equiv \mathbf{u} - \langle\rho\mathbf{u}\rangle/\langle\rho\rangle\,. \tag{2.25}$$

With no mean flow ($\langle\rho\mathbf{u}\rangle = 0$), the turbulent kinetic energy (TKE) budget becomes

$$\frac{1}{2}\frac{\mathrm{D}\langle\rho\mathbf{u}\cdot\mathbf{u}\rangle}{\mathrm{D}t} = -\langle\rho\rangle\varepsilon + \langle A\rho\mathbf{u}\cdot\mathbf{u}\rangle + \text{production at domain boundaries}\,. \tag{2.26}$$

ε is the mean turbulent dissipation rate (units $[L^2/T^3]$),

$$\varepsilon \equiv \langle \boldsymbol{\tau} : \nabla \mathbf{u} \rangle \, / \langle \rho \rangle \, , \tag{2.27}$$

To obtain quasi-steady turbulence, we want no (statistical) change in TKE. Assuming the value of A in (2.24) is a constant throughout the domain and that there is no TKE production at the boundaries, the right side of (2.26) is made zero by

$$A = \varepsilon / 2\tilde{k} \, , \tag{2.28}$$

where $\tilde{k}$ is the TKE density

$$\tilde{k} = \tfrac{1}{2} \langle \rho \mathbf{u} \cdot \mathbf{u} \rangle / \langle \rho \rangle \, . \tag{2.29}$$

In practice, A can be either a fixed value, or can be calculated each time step based on a target dissipation rate $\varepsilon_{\text{target}}$ and $\tilde{k}$ measured from the current flow field (Rosales & Meneveau, 2005), i.e.,

$$A \equiv \varepsilon_{\text{target}} / 2\tilde{k} \, . \tag{2.30}$$

We use this second approach. In this approach, the energy injected by the forcing term is $A\langle \rho \mathbf{u} \cdot \mathbf{u} \rangle = \langle \rho \rangle \varepsilon_{\text{target}}$. For single phase flow, this second approach corresponds to a constant rate of energy injection.

Extension to two-phase flows

To ensure that liner forcing does not affect the behavior of bubbles in turbulence, it can be useful to only apply this forcing outside of regions of air (where $f = 1$) (Rivière *et al.*, 2021). For the simulations in Chapter 5, the forcing term from (2.24) is multiplied by f (shifted to the appropriate velocity grid by linear interpolation) and $\langle \cdot \rangle$ is replaced with averaging in the water,

$$\langle \cdot \rangle_w = \frac{\int \cdot \, f \, \mathrm{d}V}{\int f \, \mathrm{d}V} \, . \tag{2.31}$$

Extension to vertically varying forcing

It is often useful to do forcing in only part of the domain rather than the entirety. Following Guo & Shen (2009) we will consider forcing weighted by a function $\mathcal{F}(z)$:

$$\mathbf{f} = A \mathbf{u}'' \mathcal{F}(z) \, . \tag{2.32}$$

The averaging $\langle \cdot \rangle$ used in (2.27) and (2.29) for calculating A is similarly weighted,

$$\langle \cdot \rangle = \frac{\int \cdot \, \mathcal{F}(z) \, \mathrm{d}V}{\int \mathcal{F}(z) \, \mathrm{d}V} \, . \tag{2.33}$$

For (2.25), we need to be a bit more careful with averaging. If we used (2.33), large-scale structures in $\langle u \rangle_H(z)$ and $\langle v \rangle_H(z)$ would grow quickly, as viscosity would be insufficient to

dissipate them. Instead, we use a planar average which is a function of z,

$$\langle \cdot \rangle_H(z) = \frac{\iint \cdot \, \mathrm{d}x \, \mathrm{d}y}{\iint \mathrm{d}x \, \mathrm{d}y} \,. \tag{2.34}$$

This avoids forcing structures of any scale in $\langle u \rangle_H(z)$ or $\langle v \rangle_H(z)$. While one could use more advanced filtering to avoid only large-scale structures, when we use this vertically varying forcing in simulations (see §4.3) our interest is turbulence outside the forcing region ($\mathcal{F}(z) = 0$). Because we are not too interested in the higher-order turbulence statistics within the forcing region ($\mathcal{F}(z) \neq 0$), we find this simple method sufficient.

2.2 Conservative volume of fluid method

This section describes the cVOF method developed by Weymouth & Yue (2010) to advance the VOF field in time, as used by (2.9d) and (2.10d). We start with (2.8b), the governing equation for the transport of the color function field. Because c is not smooth, $\boldsymbol{\nabla} c$ is not well defined; however, we can integrate over each cell Ω_{ijk} in the density grid and apply divergence theorem to obtain the well-defined equation

$$\frac{\partial}{\partial t} \int_{\Omega_{ijk}} c \, \mathrm{d}V = - \oint_{\partial \Omega_{ijk}} c u_n \, \mathrm{d}S + \int_{\Omega_{ijk}} c \boldsymbol{\nabla} \cdot \mathbf{u} \, \mathrm{d}V \,, \tag{2.35}$$

where $\partial \Omega_{ijk}$ is the surface of cell Ω_{ijk} and u_n is the velocity normal to the surface (using the convention that $u_n > 0$ indicates flow out of the cell). For conciseness, we consider a single cell and drop the "ijk" subscripts. Recalling (2.14), we can write the previous equation in terms of the VOF field,

$$\Delta \Omega \frac{\partial f}{\partial t} = -F_{net} + \int_{\Omega} c \boldsymbol{\nabla} \cdot \mathbf{u} \, \mathrm{d}V \,, \tag{2.36}$$

where F_{net} describes the net flux of c out of Ω.

Weymouth & Yue (2010) propose an operator-split advection scheme to solve (2.36). To go from the VOF field f^k at time t^k to f^{k+1} at $t^{k+1} = t^k + \Delta t$ in an $\mathcal{N}$ dimensional domain,[3] an operator-split method calculates intermediate fields $f^{(d)}$ for $d = 1 \ldots \mathcal{N}$ where $f^{(0)} = f^k$ and $f^{(\mathcal{N})} = f^{k+1}$. Each intermediate step corresponds to advection in one direction,

$$\frac{\Delta \Omega}{\Delta t} \left(f^{(d)} - f^{(d-1)} \right) = F_{d+1/2} - F_{d-1/2} + \int_{\Omega} c \frac{\partial u_d}{\partial x_d} \, \mathrm{d}V \quad \text{for} \quad d \in 1 \ldots \mathcal{N} \,, \tag{2.37}$$

Thus, for each direction, an operator split scheme requires calculating the flux on the positive face ($F_{d+1/2}$), the flux on the negative face ($F_{d-1/2}$), and the dilation term. Section 2.2.1 describes the explicit second-order interface reconstruction method used by cVOF to calculate the flux terms and section 2.2.2 describes the treatment of the dilation term.

Weymouth & Yue (2010) identify three requirements for a volume-conservative (to machine precision) operator-split advection scheme:

[3] Here, for simplicity, we consider single-stage time-stepping. For two-stage Runge-Kutta, the only difference is the velocity used for the second stage is the average of u^k and $u^{k+1/2}$, as shown in (2.10d).

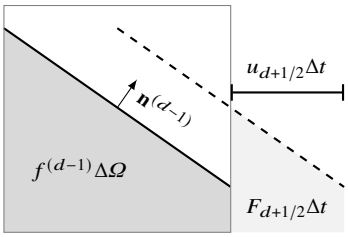


Figure 2-2: Illustration of the geometric calculation of the flux term $F_{d+1/2}$ based on the reconstructed interface (——) after being transported by a face velocity $u_{d+1/2} > 0$ (- - - -).

1. flux terms are conservative,
2. the dilation terms sum to zero, and
3. there is clipping due to over or under filling at any step.

Over or under filling refers to violation of

$$0 \leq f^{(d)} \leq 1\,, \tag{2.38}$$

which would clearly be incompatible with the definition of the VOF field as the average of a binary color function, (2.14). We note that, due to the arbitrary choice of which fluid is 'dark' and which is 'light', the difference between under filling and over filling is simply a matter of convention. A conservative advection scheme should have neither.

2.2.1 Interface reconstruction-based flux calculation

Within each grid cell, cVOF represents the interface as a linear plane, described by an interface normal vector $\mathbf{n}$ and a scalar intercept α. In this representation, the interface is located at locations $\mathbf{x}$ such that

$$\mathbf{n} \cdot \mathbf{x} = \alpha\,. \tag{2.39}$$

There is a known analytic function to determine the void fraction f of a cell given α and $\mathbf{n}$, and Scardovelli & Zaleski (2000) provide an explicit inversion to find α given f and $\mathbf{n}$. The challenge of interface reconstruction is determining $\mathbf{n}$.

Weymouth & Yue (2010) propose a second-order no-inversion VOF interface reconstruction algorithm (NIVIRA) which determines the interface normal $\mathbf{n}$ based on a $3 \times 3 \times 3$ region surrounding the grid cell. No inversion contrasts with other methods which require repeated generation and inversion of trial interfaces. The first step of NIVIRA is to determine the dominant direction of the interface. A first estimate of the interface normal, $\tilde{\mathbf{n}} = -\boldsymbol{\nabla} f$, is calculated using second-order central difference and the largest component of $\tilde{\mathbf{n}}$ determines the dominant direction of the interface. Once the dominate direction is identified, the void fraction is summed over groups of three cells in that direction to go from a $3 \times 3 \times 3$ array of void fractions to a 3×3 array of heights in the dominate direction. By selecting forward or backward differencing based on the height function in the center cell, Weymouth & Yue (2010) show that one is always able to obtain a second-order accurate interface normal $\mathbf{n}$, meaning any linear interface would be reconstructed exactly.

This interface reconstruction is repeated each operator-split step of cVOF based on the previous $f^{(d-1)}$ to determine $\mathbf{n}^{(d-1)}$ and $\alpha^{(d-1)}$. For each flux term, the upwind

interface is transported with the velocity on the face of the cell, from which the flux of dark fluid can be easily calculated geometrically (e.g., figure 2-2). The upwind nature of this approach guarantees flux terms are conservative, satisfying the first requirement for volume-conservation.

2.2.2 Treatment of the dilation term

In addition to NIVIRA to efficiently calculate interface normals to obtain the flux terms, the second development by Weymouth & Yue (2010) is a treatment of the dilation term which satisfies requirements 2 and 3 of a volume-conservative operator-split advection scheme. To start, Weymouth & Yue (2010) propose that $c(\mathbf{x})$ within a cell will be approximated by a single value $\tilde{c}$. In this case, (2.37) becomes

$$\frac{\Delta\Omega}{\Delta t}\left(f^{(d)} - f^{(d-1)}\right) = F_{d+1/2} - F_{d-1/2} + \tilde{c}\frac{\partial u_d}{\partial x_d}\Delta\Omega \quad \text{for} \quad d \in 1\ldots\mathcal{N}\,, \tag{2.40}$$

where $\partial u_d/\partial x_d = (u_{d+1/2} - u_{d-1/2})/\Delta x$. Because $\partial u_d/\partial x_d$ is calculated using the same scheme MAC uses for $\nabla\cdot\mathbf{u}$ in (2.9b) and (2.10b), we guarantee that (at least to the precision of the Poisson solver),

$$\sum_{d=1}^{\mathcal{N}}\frac{\partial u_d}{\partial x_d} = 0\,. \tag{2.41}$$

The conclusion is that only a $\tilde{c}$ that remains constant throughout the operator-split steps will lead to a dilation term that will sum to zero, satisfying requirement 2.

If $\tilde{c}$ must remain constant throughout the operator-split steps, for an explicit scheme it should only be a function of $f^{(0)}$ available at the start of the operator-split steps. Weymouth & Yue (2010) show how naive approaches like $\tilde{c} = f^{(0)}$, 0, or 1 will lead to over or under filling. Instead, Weymouth & Yue (2010) propose using a cell center value,

$$\tilde{c} = \begin{cases} 1 & \text{if} \quad f^{(0)} > 1/2 \\ 0 & \text{otherwise} \end{cases}\,, \tag{2.42}$$

which they prove, along with the Courant-like timestep restriction

$$\Delta t\sum_{d=1}^{\mathcal{N}}\left|\frac{u_d}{\Delta x_d}\right| < C \tag{2.13}$$

where $C = 1/2$, guarantees that no over or under filling can take place.

2.2.3 Suppression of spurious wisps

The cVOF method described so far conserves volume to machine precision; however, the calculation of fluxes using interface reconstruction often does not fully fill or empty a cell due to the finite precision of the geometric calculations (Baraldi *et al.*, 2014). This creates cells with f very close but not equal to 0 or 1, referred to as wisps. To suppress these

wisps, we introduce a zero-threshold value ϵ_f on VOF field (typically, $\epsilon_f = 10^{-12}$). After performing the $\mathcal{N}$ operator-split steps of (2.40) to obtain $f^{(\mathcal{N})}$, the following filter is applied to obtain the final new VOF field,

$$f^{k+1} = \begin{cases} 1 & \text{if} \quad 1 - f^{(\mathcal{N})} < \epsilon_f \\ 0 & \text{if} \quad f^{(\mathcal{N})} < \epsilon_f \\ f^{(\mathcal{N})} & \text{otherwise} \end{cases} . \tag{2.43}$$

In addition to suppressing wisps, this filter also addresses a practical challenge that arises from using a floating-point representation of f. Due to the relative resolution of floating-point arithmetic, operations on nearly full cells ($f \approx 1$) will have larger precision errors than nearly empty cells ($f \approx 0$). This creates a discrepancy between operations on f and $1 - f$, despite the choice of light/dark fluid being arbitrary. ϵ_f larger than machine epsilon ($O(10^{-16})$ for double precision) ensures that the nature of machine-precision-related errors are similar regardless of the choice of light/dark fluid.

2.3 Bubble identification using connected component labeling

With the DNS described above, we have access to the void fraction f^k_{ijk} of each cell at each simulation time step t^k, which is a discrete representation of the color function $c(\mathbf{x}, t^k)$. This is an Eulerian field describing the location of air and water, but for analysis we want to group the air into contiguous regions, i.e., bubbles. At some snapshot in time t^n corresponding to some simulation time step t^k, we want to identify the set of bubbles $\mathscr{B}^n$ and mark the air associated with each with a label $l \in \{1 \dots M^n\}$, where M^n is the number of bubbles at t^n. Note that we have switched from a time index k corresponding to increments of the simulation by Δt to a time index n corresponding to increments of time intervals between snapshots Δt_s, i.e., $t^{n+1} = t^n + \Delta t_s$. This notation will be useful in Chapter 3 when we introduce tracking bubbles through time between snapshots. This section addresses the prerequisite step of identifying bubbles at a single snapshot.

To develop a bubble labeling scheme, we first assume that the air in each cell can only be associated with a single bubble. This allows bubble labeling to be posed as connected-component labeling (CCL), which is ubiquitous in the field of computer vision. For CCL on large three-dimensional domains distributed across multiple nodes and processors, we develop a C++ library which uses equivalent label sets (He *et al.*, 2007) to store connection information on a single processor, and then the methods described by Harrison *et al.* (2011) and Iverson *et al.* (2015) to determine connections across parallel processors and assign globally unique labels. We also develop a simpler Python implementation, `blobid-python`[4], for post processing when massive parallelization is not necessary. We find the computational cost of CCL is generally trivial compared to the DNS flow solver. Thus, rather than the details of the CCL algorithm, the focus of this section is on how CCL can be adapted from computer vision to bubble labeling.

[4] `https://github.com/dgaylo/blobid-python`

To apply CCL to bubble labeling, we must specify two types of criteria. First, we must define which grid cells are "object" cells, i.e., cells that will eventually be part of a bubble. Second, we must define a way to determine if two adjacent object cells are connected. While the accuracy of bubble labeling is fundamentally limited by the flow solver's discrete representation of the air-water flow (i.e., grid size Δ), recent work has shown that the choice of objectivity and connectivity criteria also affect accuracy (Hendrickson *et al.*, 2020; Chan *et al.*, 2021*a*).

2.3.1 Objectivity criteria

The first step of bubble labeling is to determine which grid cells are to be part of a bubble. Chan *et al.* (2021*a*) focus on modifying objectivity criteria as a way to improve the accuracy of bubble labeling. Using the convention that $f^n_{ijk} = 1$ corresponds to a cell filled with air at time t^n and $f^n_{ijk} = 0$ corresponds to a cell filled with water, the simplest criteria is that the amount of air in the cell must exceed a threshold value ϕ_c:

$$f^n_{ijk} > \phi_c \,. \tag{2.44}$$

Recalling the zero-threshold filter (2.43), any $\phi_c \leq \epsilon_f$ will identify all air in the simulation as part of a bubble. A challenge with a small ϕ_c (e.g., $\phi_c \sim 10^{-12}$) is that it causes the wisps discussed in section 2.2.3 to be identified as part of bubbles. While a wisp connected to a real bubble has a negligible effect on the total volume, Chan *et al.* (2021*a*) show wisps can cause multiple bubbles to be artificially linked, significantly affecting the bubble size distribution. One option to suppress wisp bridging is to increase ϕ_c (e.g., $\phi_c = 0.5$); however, this causes cells on the surfaces of a real bubble to be excluded, reducing the total volume. A large ϕ_c means that a significant portion of air will not be assigned to a bubble. This is a particular issue for volume-based bubble tracking (see Chapter 3), as it means we will get an incomplete description of the evolution of the bubble population.

Chan *et al.* (2021*a*) propose a different objectivity criteria to allow suppression of wisp bridging while reducing unassigned air compared to large ϕ_c. In addition to (2.44), at least one of the cell's six neighbor cells must have a void fraction that exceeds a second threshold value $\phi_{c,m}$:

$$\left[f^n_{ijk} > \phi_c\right] \wedge \left\{\left[f^n_{i+1jk} > \phi_{c,m}\right] \vee \left[f^n_{i-1jk} > \phi_{c,m}\right] \vee \left[f^n_{ij+1k} > \phi_{c,m}\right] \vee \cdots\right\} . \tag{2.45}$$

Chan *et al.* (2021*a*) propose $\phi_c = 0$ and $\phi_{c,m} = 0.5$. This reduces the volume lost on the surfaces of bubbles, as any cell with air that is within one grid cell of a mostly full cell will be added to a bubble.

2.3.2 Connectivity criteria

In addition to whether a grid cell is an object cell or not, CCL needs a criteria for whether pairs of object cells are connected. To start, we specify that only adjacent cells can be connected. Depending on whether diagonal cells are included, a cell on a three-dimensional Cartesian grid can have either 6 or 26 adjacent cells. Figure 2-3 illustrates this for two

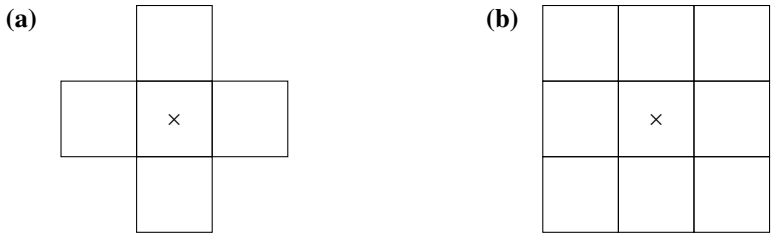
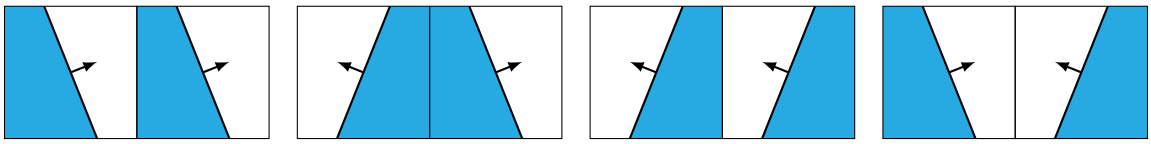


Figure 2-3: Illustration of (a) 4-connectivity and (b) 8-connectivity for a two-dimensional grid, where "×" is the object cell and the other cells are those that would be considered adjacent.



(a) $n_x > 0; n_x > 0$ **(b)** $n_x < 0; n_x > 0$ **(c)** $n_x < 0; n_x < 0$ **(d)** $n_x > 0; n_x < 0$

Figure 2-4: The four different configurations ICL considers for the sign of interface normals in two adjacent grid cells. In (a)–(c) ICL considers the two cells connected. In (d) the normals are opposed and ICL does not consider the two cells connected.

dimensions, where a cell can have either 4 or 8 adjacent cells. For bubble identification we do not include diagonal cells and use 6-connectedness. Without any additional criteria, an object cell will be considered connected to any object cell among its 6 neighbors.

A challenge with the simple connectivity criteria is that bubbles are often merged together if they have interfaces separated by $< 2\Delta$. This artificial merging can be reduced by sharpening the VOF field. The Informed Component Labeling algorithm (ICL) proposed by Hendrickson *et al.* (2020) uses NIVIRA developed by Weymouth & Yue (2010) (see description in section 2.2.1) to obtain piecewise linear interface reconstruction in each grid cell. As illustrated in figure 2-4, ICL adds an additional connectivity criteria: adjacent object cells are only considered connected if their interface normals are not opposed. Given ICL only uses the sign of the interface normal, one could be tempted to use a less accurate method to approximate the interface normals, such as $\tilde{\mathbf{n}} = -\boldsymbol{\nabla} f$ calculated using central differencing. However, we will show this notably reduces the accuracy of ICL.

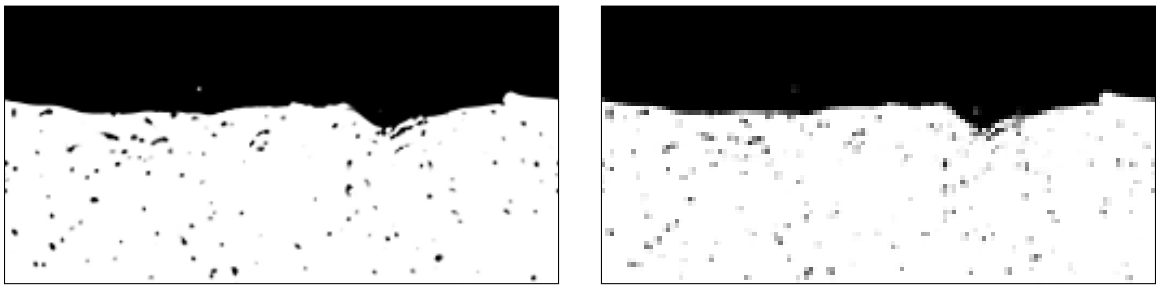
2.3.3 Comparison of bubble labeling schemes

To evaluate different bubble labeling schemes, we consider 400 snapshots of the VOF field near the free surface from an illustrative simulation[5] which has air bubbles in water near a free surface. We consider the original VOF field with uniform grid size Δ, as well as the field downsampled by averaging eight grid cells into one to give a larger grid size $\Delta' = 2\Delta$. Figure 2-5 shows an example of the two different fields.

We consider three different bubble labeling schemes:

- The method proposed by Chan *et al.* (2021*a*), where (2.45) is used ($\phi_c = 0$ and $\phi_{c,m} = 0.5$) for the object criteria and no additional connectivity criteria is imposed (THRESH).

[5]DNS of forced FST simulation at $Fr^2 = 1.2$, $We = \infty$, as described in section 4.3



(a) Original, $256^2 \times 128$ **(b)** Downsampled, $128^2 \times 64$

Figure 2-5: Slice of the VOF field from a snapshot of bubbles near an air entraining free surface at two different resolutions. Air is white and water is black.

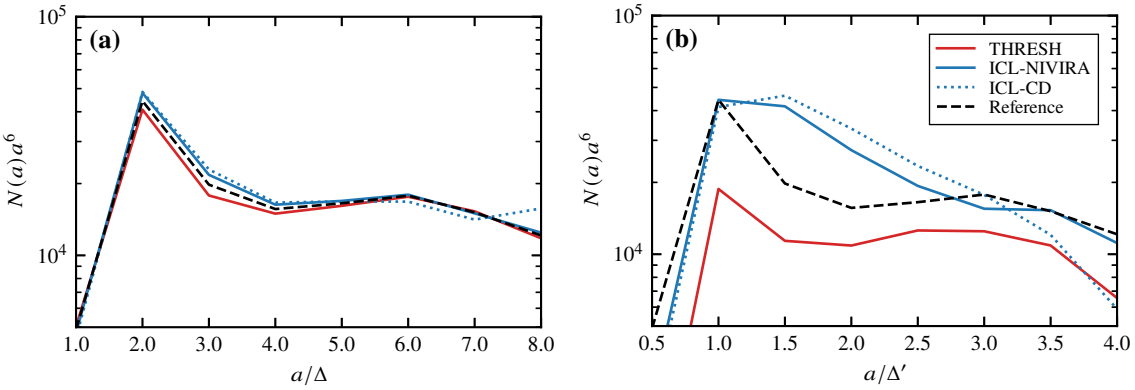


Figure 2-6: Bubble size distributions $N(a)$ calculated using different bubble identification methods for (a) the original grid and (b) the downsampled grid. As reference, (- - - -) shows $N_{\text{ref}}(a)$, the average result of THRESH and ICL-NIVIRA on the original grid.

- The method proposed by Hendrickson *et al.* (2020), where (2.44) is used ($\phi_c = 0$) for the object criteria and interface normals, calculated with NIVIRA, are used for an additional connectivity criteria (ICL-NIVIRA).
- The same method proposed by Hendrickson *et al.* (2020), but with central differencing rather than NIVIRA used to calculate the normals (ICL-CD).

Using each of these methods we identify bubbles in each of the 400 snapshots using the `bolbid-python` library. From a bubble's volume v, we calculate an effective radius $a = (3v/4\pi)^{1/3}$ and then bin the results to determine the average bubble size distribution $N(a)$, shown in figure 2-6.

Suppression of wisp bridging

First, we compare THRESH and ICL-NIVIRA at the original resolution. Especially for $a < 5\Delta$ we see that ICL-NIVIRA identifies more bubbles of a given radius than THRESH. In this simulation, there is a lot of air near the free surface. ICL-NIVIRA's stricter connectivity criteria implies more of this air is in bubbles, disconnected from the bulk region of air above the free surface (the 'sky'). THRESH's less strict connectivity criteria implies more of the

air is still connected to the sky. Chan *et al.* (2021*a*) suggest that excluding wisps using a large ϕ_c or $\phi_{c,m}$ in the objectivity criteria is necessary to suppress wisp bridging and robustly separate bubbles from the sky; however, our results show that ICL (which uses $\phi_c = 0$) separates more bubbles from the sky. The orientation of the interface normal for wisps is essentially random, so rather than a single large bridging structure, ICL identifies wisps as many small bubbles. Thus, wisps with ICL lead to many small bubbles of radius $a \ll \Delta$ which are easily excluded from analysis.

Effect of grid resolution

We now consider the accuracy of ICL-NIVIRA versus THRESH. Because this is real simulation data, we do not have a given bubble size distribution to use as truth. For the original grid, we note that both ICL-NIVIRA and THRESH give fairly similar values for $N(a)$ across the a/Δ range we consider, and almost identical values for $a > 5\Delta$. Based on this, we treat $N_{\text{ref}}(a)$, the average of the ICL-NIVIRA and THRESH $N(a)$s, as the reference "truth" for comparison. Against this reference, we consider the accuracy of ICL-NIVIRA versus THRESH on the down-sampled courser grid ($\Delta' = 2\Delta$). Figure 2-6 shows that THRESH consistently under predicts the number of bubbles by roughly a factor of two. This means that as the resolution goes down, more bubbles are incorrectly connected to the sky. The ratio of $N(a)$ predicted by THRESH and $N_{\text{ref}}(a)$ is consistent across bubble sizes, meaning the probability that THRESH will erroneously link a bubble to the sky is independent of the bubble's size.

For ICL-NIVIRA, we see that there is very little error on the down-sampled grid for bubbles $a > 3\Delta'$. Recall that NIVIRA produces a second-order piece-wise linear interface based on a $3 \times 3 \times 3$ stencil. For roughly spherical bubbles with radii larger than three times the grid size, this will produce very accurate interface orientations. Figure 2-6 confirms that, as intended by Hendrickson *et al.* (2020), ICL-NIVIRA can accurately distinguish sufficiently sized closely spaced bubbles using interface reconstruction. For bubbles $a < 3\Delta'$ we see ICL-NIVIRA over predicts the number of bubbles, though by a factor similar to how THRESH under predicts bubbles of all sizes.

Effect of less accurate interface reconstruction on ICL

Finally, we highlight that the accuracy we see here for ICL-NIVIRA is closely tied to the accuracy of the NIVIRA reconstruction method. In Figure 2-6 we also consider ICL-CD, where simple central differencing is used instead. Weymouth & Yue (2010) discuss how central differencing in this case is less than second order accurate, and figure 2-6 shows that using this less than second order accurate method significantly reduces the overall accuracy of ICL. In addition to a reversion to the THRESH behavior of linking bubbles to the sky as resolution decreases, we also see some error in the large bubbles at the full resolution. When using ICL, one should make sure the interface reconstruction method is strictly second order accurate, as is the case with NIVIRA. Additionally, some caution should be taken for bubbles with radii near the grid size, as the piece-wise linear representation of the interface may be too course.

Chapter 3

Eulerian Label Advection Method for Volume-Conservative Bubble Tracking

As discussed in Chapter 1, multiple bubble evolution mechanisms are present and may be relevant to the bubble population beneath an entraining free surface. Measurements of the bubble population alone only provide the net effect of all these mechanisms, limiting what can be elucidated about any individual mechanism. To study these evolution mechanisms, we require methods to quantify individual evolution mechanisms within the bubble population. This is achieved with bubble tracking, which identifies the individual events related to each evolution mechanisms. While our interest in this work is bubbles, we note that the methods derived in this chapter are equally applicable to droplets.

Key results from this chapter are summarized in "An Eulerian label advection method for conservative volume-based tracking of bubbles/droplets" by Gaylo, Hendrickson & Yue (2022). Here we provide more details, particularly on how the method is used to quantify individual evolution mechanisms. An implementation of the method is available in the `flexELA`[1] library.

3.1 Introduction

For bubble tracking, simulations have an advantage over experiments in that all (resolved) properties of the flow are readily available, relevant here, the velocity field $\mathbf{u}(\mathbf{x}, t)$ and the color function field $c(\mathbf{x}, t)$. Still, bubble tracking is a challenge in simulations. Formally, a bubble can be defined as a volume enclosed by a continuous surface representing the interface between air ($c = 1$) and water ($c = 0$).[2] Neglecting dissolution, this surface is material, so its evolution (and thus bubble evolution) can be obtained from $\mathbf{u}(\mathbf{x}, t)$. However, the evolution of the surface is a challenge because numerical methods such as volume of fluid or level set do not explicitly define this interface. This lack of an explicit interface creates two challenges in regard to bubble tracking. The first challenge is the need to identify individual bubbles at an instant in time (snapshot) based on the discretized representation

[1] `https://github.com/dgaylo/flexELA`

[2] This chapter uses the convention that air is 'dark' fluid, the opposite of the convention used in section 2.1.

of the color function field. The second challenge is to track how these individual bubbles evolve between two adjacent snapshots. The first challenge is, more broadly, the process of connected-component labeling (CCL), as discussed in section 2.3. In this chapter, we will focus on the second challenge of tracking the evolution of individual bubbles. To this end, we will take the bubble identification provided by CCL as a given and focus on the accuracy of the tracking of these identified bubbles. Although the ultimate accuracy of measured bubble evolution mechanisms depends on both the accuracy of CCL and tracking, this approach avoids entangling these two challenges.

Recently, two methods for bubble tracking in numerical simulations have been proposed (Chan *et al.*, 2021*a*; Gao *et al.*, 2021). Both methods can be classified as Lagrangian methods, as their inputs from the simulation are Lagrangian integral quantities of bubbles, e.g., volume, centroid, and total momentum. A limitation of both these methods is that they assume that all events are binary, meaning events involve at most two bubbles from one snapshot and one bubble from the other snapshot. Applied to air entraining flows, the Lagrangian methods have had some success describing fragmentation away from the free surface (Gao *et al.*, 2021; Chan *et al.*, 2021*c*); however, the near the free surface the Lagrangian methods struggle. They are unable to accurately describe the high-arity evolution of large complex air structures near the free surface. Application of the binary assumption to non-binary events introduces erroneous creation/extinction events (Chan *et al.*, 2021*a*). This is a serious problem, as such error is inseparable from entrainment and degassing statistics. Although a Lagrangian method could theoretically be extended to capture higher-arity events, it is not practical computationally.

We pursue a different, Eulerian approach to bubble tracking. Unlike Lagrangian methods, the available velocity field $\mathbf{u}(\mathbf{x}, t)$ is used to determine the evolution of bubbles. Because an Eulerian approach focuses on grid-level detail to describe bubble evolution, the complexity of the formulation can be independent of the arity of the bubble-level events. In section 3.2 we develop a volume-based framework to describe the evolution of bubbles in terms of the movement of air volume between bubbles. This volume-based tracking approach uniquely describes the evolution of the bubble population through a volume-tracking matrix (VTM), which can describe evolution regardless of the complexity. From this more general description of bubble evolution, individual events (entrainment, degassing, fragmentation, etc.) can be extracted. In section 3.3 we demonstrate that Lagrangian methods are unable to uniquely provide the VTM. This is true even if a Lagrangian method were able to avoid the binary assumption. In section 3.4 we build upon the volume-fraction fluxes provided by the cVOF method (Weymouth & Yue, 2010) to create the Eulerian Label Advection method (ELA), a volume-conservative numerical implementation of volume-based tracking. By leveraging fluxes already calculated by cVOF, ELA minimizes additional computational cost. In section 3.5 these results are validated using the canonical problem of bubbles fragmenting in homogeneous isotropic turbulence (HIT).

3.2 Volume-based bubble tracking

Before going further, it is useful to define a mathematical framework to describe bubble evolution. Consider that given a numerical approximation of the color function from a

snapshot at time t^n, $c(\mathbf{x}, t^n)$, a CCL method provides a set of non-overlapping bubbles $\mathscr{B}^n = \{1 \dots M^n\}$, where M^n is the number of bubbles. The CCL method identifies these bubbles by labeling the air withing each with a corresponding label $l \in 1 \dots M^n$. Some CCL methods do not identify all air as being part of a bubble (see §2.3.1); so we define an additional "bubble" with label $l = 0$ containing all the (not necessarily contiguous) air not otherwise assigned to a bubble. To represent the results of CCL, we define a vector color function $\mathbf{c}^n(\mathbf{x}, t)$ with elements initially defined at time t^n by

$$c_l^n(\mathbf{x}, t^n) = \begin{cases} 1 & \text{if} \quad \mathbf{x} \in \text{bubble } l \\ 0 & \text{else} \end{cases} \qquad \text{for} \quad l \in 0 \dots M^n \,. \tag{3.1}$$

Rather than a single binary fluid color function describing if there is air or water at a location $\mathbf{x}$ and time t, we have split the air into separate immiscible fluids based on which bubble it was in at time t^n. Equivalent to (2.8b), the evolution of $\mathbf{c}^n$ must satisfy

$$\frac{\partial \mathbf{c}^n}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla} \mathbf{c}^n = \mathbf{0} \,. \tag{3.2}$$

With inflow boundary conditions, not all air will have necessarily been in an identified bubble at t^n, but this can be accounted for by adding a bubble (or bubbles) to $\mathscr{B}^n$ and setting the corresponding element(s) of $\mathbf{c}^n$ at the boundaries. Because bubbles are defined to be non-overlapping at t^n, all air is associated with a bubble, and $\mathbf{c}^n$ is advected with the same $\mathbf{u}$ as c, the following consistency requirement is true at all times and locations:

$$\sum_l c_l^n(\mathbf{x}, t) = c(\mathbf{x}, t) \,. \tag{3.3}$$

For the air at any location $\mathbf{x}$ at a time t, $\mathbf{c}^n(\mathbf{x}, t)$ provides the bubble $l \in \mathscr{B}^n$ that contained the particle at time t^n. Thus, $\mathbf{c}^n$ provides a complete Eulerian description of the flow of air.

3.2.1 A volume-tracking matrix description of bubble evolution

While $\mathbf{c}^n$ provides a complete Eulerian description of the evolution of air over arbitrary time, we are interested in the evolution of individual bubbles over snapshot intervals Δt_s. Describing the evolution in terms of individual bubbles implies a Lagrangian description, where we follow individual bubbles rather than the entire field of air. Here we will show how a matrix-based Lagrangian description of bubble evolution is obtainable from the previous Eulerian description.

Integrating (3.2) over time (advection), the vector color function $\mathbf{c}^n(\mathbf{x}, t)$ originally defined by (3.1) at time t^n, i.e., $\mathbf{c}^n(\mathbf{x}, t^n)$, can be advanced in time to the next snapshot $t^{n+1} = t^n + \Delta t_s$ to give $\mathbf{c}^n(\mathbf{x}, t^{n+1})$. At the next snapshot ($t^{n+1}$), a CCL method provides a new set $\mathscr{B}^{n+1} = \{0 \dots M^{n+1}\}$ of bubbles, from which we use (3.1) again to define a new vector color function $\mathbf{c}^{n+1}(\mathbf{x}, t)$, with the initial value defined at t^{n+1} by

$$c_m^{n+1}(\mathbf{x}, t^{n+1}) = \begin{cases} 1 & \text{if} \quad \mathbf{x} \in \text{bubble } m \\ 0 & \text{else} \end{cases} \qquad \text{for} \quad m \in 0 \dots M^{n+1} \,. \tag{3.4}$$

Based on $\mathbf{c}^n$ and $\mathbf{c}^{n+1}$, both available at time t^{n+1}, the volume of dark fluid from a bubble $l \in \mathscr{B}^n$ that ends up in a bubble $m \in \mathscr{B}^{n+1}$ is

$$q_{ml} = \int_{\forall} c_m^{n+1}\left(\mathbf{x}, t^{n+1}\right)\ c_l^n\left(\mathbf{x}, t^{n+1}\right) \mathrm{d}V\,, \tag{3.5}$$

where $\forall$ is the whole domain. Applying (3.4), this is equivalent to

$$q_{ml} = \int_{\mathbf{x}\in\text{bubble } m} c_l^n\left(\mathbf{x}, t^{n+1}\right) \mathrm{d}V\,. \tag{3.6}$$

For $l = 0 \ldots M^n$ and $m = 0 \ldots M^{n+1}$, we define the matrix $\mathbf{Q}^{(n\to n+1)} = \{q_{ml}\}$, which provides a complete description of the flow of air from bubbles $\mathscr{B}^n$ to bubbles $\mathscr{B}^{n+1}$. Each element q_{ml} in the matrix $\mathbf{Q}^{(n\to n+1)}$ provides the (absolute) volume of air that transfers from bubble l to bubble m over the interval t^n to t^{n+1}.

We note that, based on $\mathbf{c}^n$, we can express the volume of all the bubbles at time t^n as a vector $\mathbf{v}^n$ of length M^n,

$$\mathbf{v}^n = \int_{\forall} \mathbf{c}^n(\mathbf{x}, t)\, \mathrm{d}V\,. \tag{3.7}$$

Recalling (3.3), the column and row sums of $\mathbf{Q}^{(n\to n+1)}$ give the bubble volumes at t^n and t^{n+1} respectively:

$$\sum_m q_{ml} = v_l^n \quad , \tag{3.8a}$$

$$\sum_l q_{ml} = v_m^{n+1}\,. \tag{3.8b}$$

Normalizing the columns of $\mathbf{Q}^{(n\to n+1)}$ by $\mathbf{v}^n$,

$$a_{ml} = q_{ml} \big/ v_l^n\,, \tag{3.9}$$

we define the volume-tracking matrix (VTM), $\mathbf{A}^{(n\to n+1)} = \{a_{ml}\}$. The VTM is a left stochastic matrix,

$$\sum_m a_{ml} = 1\,, \tag{3.10}$$

which describes the evolution of volume from bubbles $\mathscr{B}^n$ to bubbles $\mathscr{B}^{n+1}$ as

$$\mathbf{v}^{n+1} = \mathbf{A}^{(n\to n+1)}\mathbf{v}^n\,. \tag{3.11}$$

We note that if instead the rows of $\mathbf{Q}^{(n\to n+1)}$ are normalized by $\mathbf{v}^{n+1}$, one gets a functionally equivalent right stochastic matrix description of the reverse evolution.

It is important to note that, through the integration in (3.5), we have lost some information about the flow of air on the particle level. Rather than a deterministic description of the flow of air particles, each entry a_{ml} of the VTM can be interpreted as the probability a particle of dark fluid is in bubble $m \in \mathscr{B}^{n+1}$ at t^{n+1} given that it was in bubble $l \in \mathscr{B}^n$ at t^n. This means that, unlike $\mathbf{c}^n$, the VTM $\mathbf{A}^{(n\to n+1)}$ does not provide a complete Eulerian description of the evolution of air. However, on the bubble level, each entry a_{ml} provides the (deterministic)

proportion of the volume of bubble $l \in \mathscr{B}^n$ that ends up in bubble $m \in \mathscr{B}^{n+1}$. Thus, the VTM provides a complete Lagrangian description of the evolution of bubbles.

Given that the VTM is a left stochastic matrix, we can also consider the evolution of bubbles over N snapshot intervals,

$$\mathbf{v}^{n+N} = \left[\prod_{m=n+N-1}^{n} \mathbf{A}^{(m \to m+1)} \right] \mathbf{v}^n \,, \tag{3.12}$$

This implies an effective VTM

$$\tilde{\mathbf{A}}^{(n \to n+N)} \equiv \prod_{m=n+N-1}^{n} \mathbf{A}^{(m \to m+1)} \,, \tag{3.13}$$

with an effective snapshot interval $\Delta t_{s,\text{eff.}} \equiv N \Delta t_s$. However, as the VTM does not provide a complete *Eulerian* description, the effective VTM $\tilde{\mathbf{A}}$ is generally not equal to the true VTM $\mathbf{A}$ with Δt_s increased to equal $\Delta t_{s,\text{eff.}}$. The computational cost of the tracking method we develop in section 3.4 increases with Δt_s so matrix multiplication provides a cheaper way to obtain long snapshot intervals; however, we find obtaining the true VTM with sufficient Δt_s is not cost prohibitive. Although not used in this work, Appendix B discusses the tradeoff between accuracy and cost when using VTM multiplication.

3.2.2 Extracting bubble evolution mechanism from the VTM

We now look at how bubble evolution mechanism can be identified by the VTM. Figure 3-1a illustrates the VTM when one bubble fragments into m daughter bubbles. In general, whenever there are multiple non-zero entries in a row of the VTM this means the volume from one bubble is now in multiple bubbles, meaning fragmentation occurred. Figure 3-1b illustrates the VTM when m bubbles coalesce into a single daughter bubble, the reverse of fragmentation. Coalescence is identified when there are multiple non-zero entries in a column of the VTM.

For entrainment and degassing, we first note that from the perspective of the VTM the bulk region of air above the free surface (the "sky") is treated just like a bubble. Thus, entrainment is simply fragmentation of the "sky" bubble and degassing is coalescence of with the "sky" bubble. This is illustrated in figure 3-1c. The VTM entry a_{11} represents air that stayed in the sky. Each non-zero entry in the same column represents an entrainment event, and each non-zero entry in the same row represents a degassing event.

Figure 3-2 shows a real example from the simulation in §3.5. Based on the first column of $\mathbf{A}^{(n \to n+1)}$, we see that the majority of the volume from the parent with label 1 at time t^n went to the child with label 4 at time t^{n+1}, apart from a small portion that went to child 3. For clarity, under-resolved bubbles have been excluded, causing the first column not sum to 1, cf. (3.10). Based on the second column, the volume from parent 2 went to three different children, 1, 2, and 3. Based on the third column, all the volume from parent 3 went to child 5. The events described by this VTM are complex in two ways: there is a non-binary fragmentation event, and both fragmentation and coalescence occurred to form child 3.

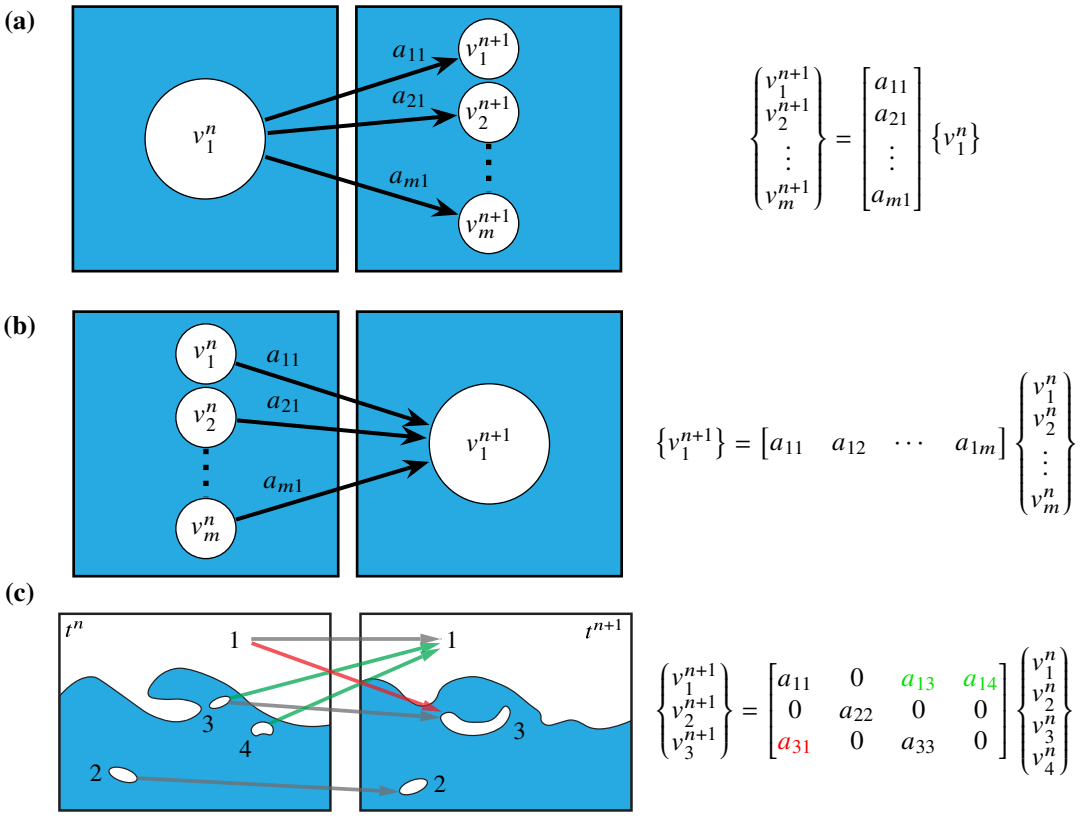
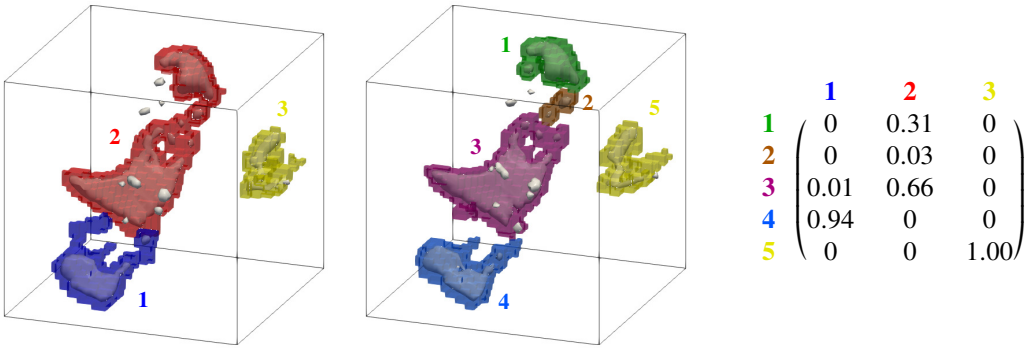


Figure 3-1: Illustration of bubble evolutions and the associated VTM equation $\mathbf{v}^{n+1} = \mathbf{A}\mathbf{v}^n$, expanded to show each term. (a) One bubble fragments into m bubbles. (b) m bubbles coalesce into a single bubble. (c) One bubble is entrained (red) and two bubbles are degassed (green), along with other processes.



(a) $t^n/t_b = 0.88$ **(b)** $t^{n+1}/t_b = 0.95$ **(c)** Subset of $\mathbf{A}^{(n\to n+1)}$

Figure 3-2: The $f = 0.5$ iso-surface from two snapshots of a subset of a simulation (see §3.5 for details) and the corresponding part of the VTM extracted using ELA (described in §3.4). Grid cells are highlighted corresponding to the label assigned by the CCL method. Note that an iso-surface is itself a CCL method, which does not necessarily align with the chosen method. 3 resolved parent bubbles (a) have their volume distributed among 5 resolved child bubbles (b). ELA gives the associated VTM (c).

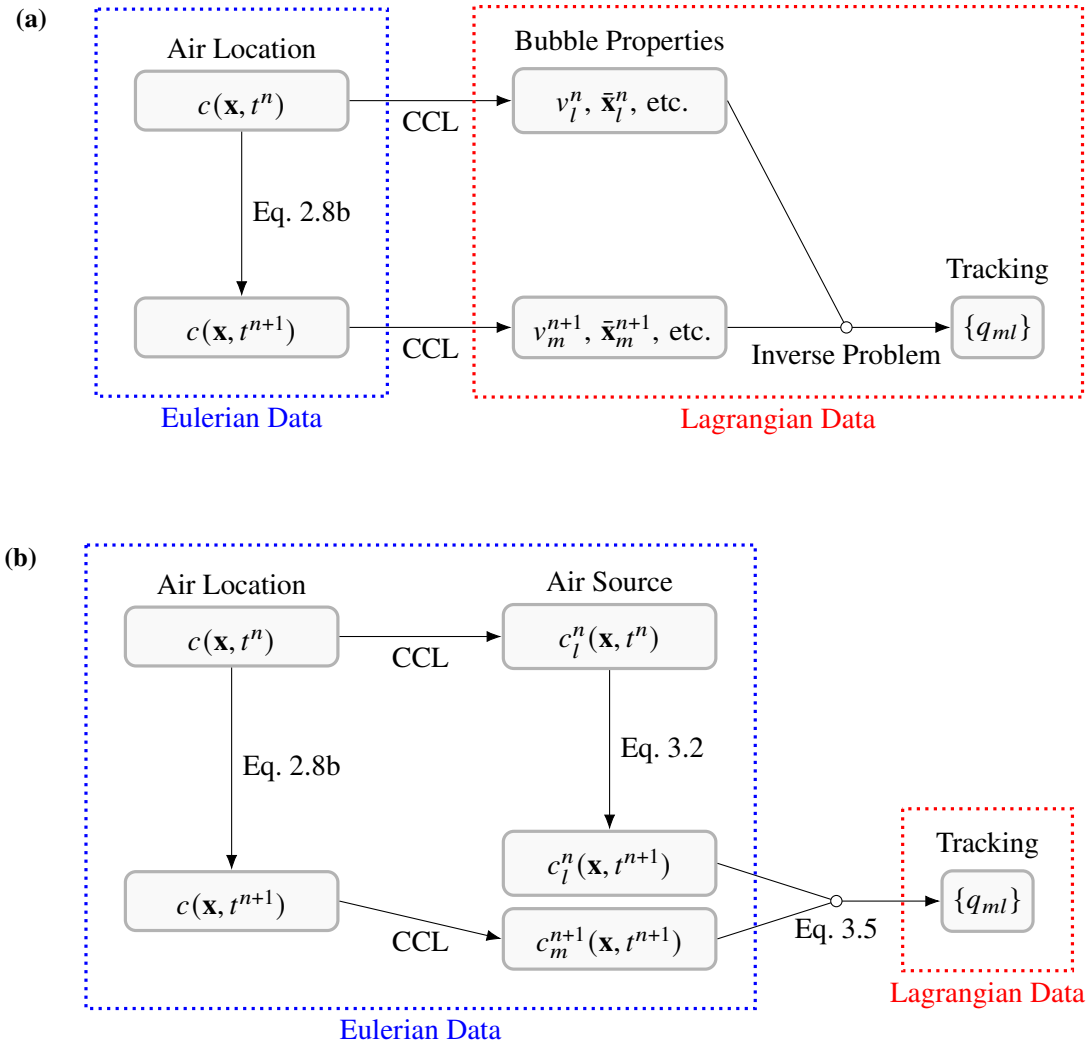


Figure 3-3: Sketch of (a) Lagrangian-based tracking and (b) Eulerian-based tracking methods to answer how much volume from a bubble labeled l at t^n ended up in a bubble labeled m at $t^{n+1} = t^n + T_s$.

3.3 Limitations of Lagrangian-based tracking

Previous methods for bubble tracking (Chan *et al.*, 2021*a*; Gao *et al.*, 2021) can be categorized as Lagrangian-based methods. Figure 3-3a illustrates how Lagrangian-based tracking works. At each snapshot interval t^n, they use the results from CCL to calculate Lagrangian properties of bubbles. In addition to bubble volume (3.7), this could include bubble centroid,

$$\bar{\mathbf{x}}_l^n = \frac{1}{v_l^n} \int_\forall c_l^n(\mathbf{x}, t^n)\, \mathbf{x}\, \mathrm{d}V\,, \tag{3.14}$$

or more generally any integral quantity over a bubble calculated at the snapshot interval t^n. From such bubble properties at two successive snapshots, the methods seek to solve an inverse problem to find a possible evolutionary path between the two snapshots. This inverse problem can be solved through physical and numerical constraints (Chan *et al.*, 2021*a*)

or by minimizing a prescribed error function (Gao *et al.*, 2021). Because the Lagrangian data for bubbles is typically very small compared to any Eulerian field, Lagrangian-based tracking is often done as part of post-processing. During the simulation the Lagrangian bubble properties are recorded then later post-processed. The two limitations of Lagrangian methods come from the fact that they are trying to solve an inverse problem.

3.3.1 Computational complexity

The first problem with the inverse problem is how the cost of the methods scales with the complexity of the evolution of the bubbles. For a simple example, first consider bubble evolution where bubbles only translate and there is no coalescence or fragmentation. For M initial bubbles, there are $M!$ possible solutions to the inverse problem. If instead each bubble at the first snapshot fragments into two bubbles, there are now $(2M)!$ possible solutions. Generalizing, if all the bubbles fragment into n daughter bubbles, the possible solutions are $(nM)!$. This simplified example illustrates the super-polynomial complexity of Lagrangian methods as the complexity of the bubble evolution (illustrated by n) increases. Lagrangian methods typically assume that all events are binary ($n = 2$), where either one bubble fragments into two or two bubbles coalesce into one. Although both Gao *et al.* (2021) and Chan *et al.* (2021*a*) both note that their methods are theoretically extendable to identify non-binary events, the underlying scaling makes it challenging.

Recently Basak *et al.* (2026) made some progress extending the Lagrangian method by Gao *et al.* (2021) to non-binary events, but this is done by assuming only simple non-binary events, where one bubble fragments into many (figure 3-1a) or many coalesce into one (figure 3-1b). Their method cannot handle complex events involving both fragmentation and coalescence (e.g., figure 3-2).

3.3.2 Cycle generation leads to non-unique solutions

In addition to computational complexity, a more fundamental problem with Lagrangian-based tracking methods is that they cannot always provide enough information to determine a unique solution to the VTM. This is because, rather than tracking air volume, Lagrangian-based methods track bubbles. To make this distinction clear, consider the case where two large bubbles of volume $v_1^n = v_2^n$ exchange two smaller bubbles of volume $0.05v_1^n$ each over a time T, as shown in figure 3-4. In the form of (3.11), the correct VTM for this evolution is:

$$\begin{Bmatrix} v_1^{n+1} \\ v_2^{n+1} \end{Bmatrix} = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix} \begin{Bmatrix} v_1^n \\ v_2^n \end{Bmatrix} . \tag{3.15}$$

Lagrangian-based methods seek to identify a solution to which bubbles at t^n contributed to a bubble at t^{n+1} based only on the information available at t^n and t^{n+1}. In terms of our VTM description, they seek to identify which entries of $\mathbf{Q}^{(n\to n+1)}$ are non-zero. However, identifying non-zero elements is insufficient to obtain a unique solution. Suppose that a Lagrangian-based method correctly identifies the non-zero elements. Applying volume

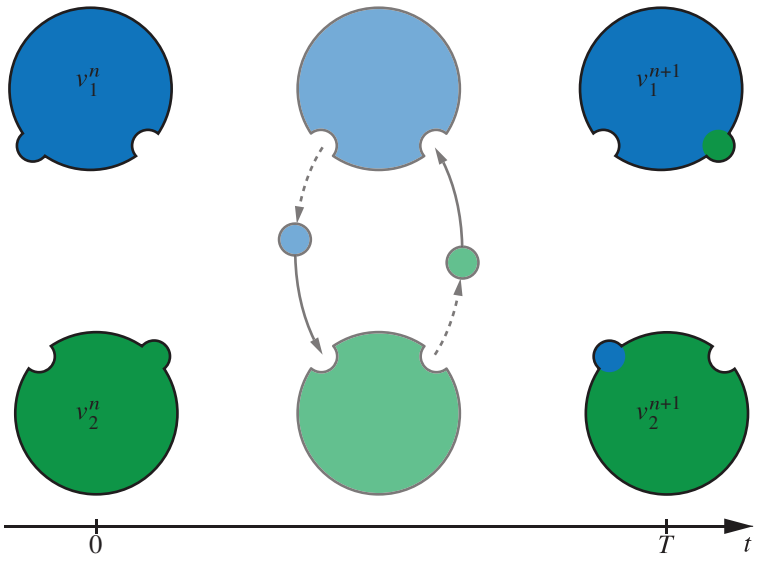


Figure 3-4: Illustration of two bubbles of equal volume exchanging 5% of their volume over a time T less than the snapshot interval Δt_s. The vector color function $\mathbf{c}^n(\mathbf{x}, t)$ is shown as blue where $c_1^n(\mathbf{x}, t) = 1$ and green where $c_2^n(\mathbf{x}, t) = 1$.

conservation using (3.10) gives the set of equations,

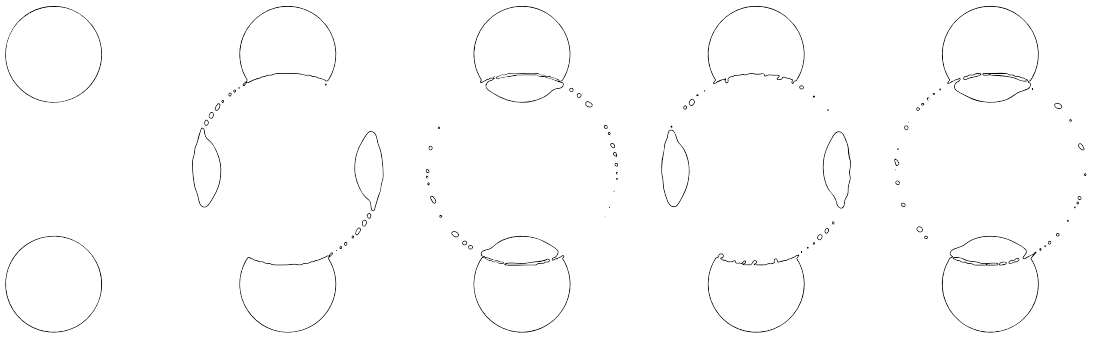
$$\begin{Bmatrix} v_1^{n+1} \\ v_2^{n+1} \end{Bmatrix} = \begin{bmatrix} a_{11} & (1 - a_{22}) \\ (1 - a_{11}) & a_{22} \end{bmatrix} \begin{Bmatrix} v_1^n \\ v_2^n \end{Bmatrix}, \tag{3.16}$$

which does not have a unique solution.

For a generalized explanation of why some VTMs cannot be solved based only on their non-zero entries and volume conservation, it is useful to use graph theory. The evolution of bubbles over a snapshot interval can be described as a graph where the nodes are all bubbles present at t^n and t^{n+1}, $\{\mathscr{B}^n, \mathscr{B}^{n+1}\}$, and the edges are the non-zero elements of $\mathbf{Q}^{(n \to n+1)}$. Each element value q_{ij} represents the flow of volume (i.e., current) along the edge. Any cycle in the graph allows a loop current, introducing a null space in the solution for the currents q_{ij}. If these cycles exist in the non-zero elements of $\mathbf{Q}^{(n \to n+1)}$, a unique solution is not obtainable without additional information. In practice, a Lagrangian-based method used for identifying non-zero elements my implicitly disallow cycles, however this only masks the problem. For (3.16), such a method would pick the solution $a_{11} = a_{22} = 1$ (or 0), missing the exchange of volume.

3.3.3 Quantifying cycle generation in two-dimensional vortical exchange

To illustrate the challenge of bubble tracking when complex interactions are present, we consider a flow similar to the conceptual one shown in figure 3-4. Consider two circular bubbles of radius a whose centers are both a distance L from the center of a forced vortex of radius R and rotation Ω, giving an angular velocity field $v_\theta(r)$ as a function of the radial



(a) $t/T = 0$ **(b)** $t/T = 1.5$ **(c)** $t/T = 3$ **(d)** $t/T = 4.5$ **(e)** $t/T = 6$

Figure 3-5: Evolution of the $f = 0.5$ iso-surface in the two-dimensional vortical exchange simulation.

distance r from the vortex center,

$$v_\theta(r) = \Omega \pi r \, \mathcal{H}(R - r) \,, \tag{3.17}$$

where $\mathcal{H}$ is the Heaviside step function. The vortex rips volume from one bubble, creating smaller bubbles which are then transferred to the other bubble. The time it takes for this exchange of volume gives the characteristic time $T = 1/\Omega$. As discussed in §3.3.2, if $\Delta t_s > T$, cycles are formed which prevent a unique solution using previous methods. Here, we perform a two-dimensional simulation of $L/R = 6/5$ and $a/R = 1/2$ with resolution $\Delta x = R/32$ over $0 < t/T < 8$ and study the effect of $\Delta t_s/T$ on cycle production. To obtain the VTM we use the Eulerian-based tracking method that will be introduced in §3.4, which is robust to the presence of cycles. Figure (3-5) shows the evolution of the $f = 0.5$ isosurface.

To avoid inflating the count of cycles by including under-resolved events, we remove columns of the VTM $\mathbf{A}^{(n \to n+1)}$ relating to under-resolved parent bubbles, $v_i^n < v_{\text{res, 2D}}$, as well as rows relating to under-resolved child bubbles, $v_j^{n+1} < v_{\text{res, 2D}}$, where $v_{\text{res, 2D}} = \pi(2\Delta x)^2$. After removing under-resolved events, we generate a matrix $\mathbf{B}^{(n \to n+1)}$ by removing all elements of the VTM $\mathbf{A}^{(n \to n+1)}$ that are not involved in a cycle. This is achieved by iteratively setting any element to zero if it is the only non-zero entry in a row or column, until no such cases exist. The proportion of the (resolved) volume of dark fluid involved in cycles can then be written

$$C^n = \frac{\sum \left[\mathbf{B}^{(n \to n+1)} \mathbf{v}^n \right]}{\sum \left[\mathbf{A}^{(n \to n+1)} \mathbf{v}^n \right]} \,. \tag{3.18}$$

Figure 3-6 shows there is no cycle production for $\Delta t_s \ll T$, a jump in cycle production at $\Delta t_s \sim T$, and all volume is involved in a cycle for $\Delta t_s > T$. Thus, unless $\Delta t_s \ll T$, previous tracking methods would be unable to provide unique solutions to the flow of volume between bubbles.

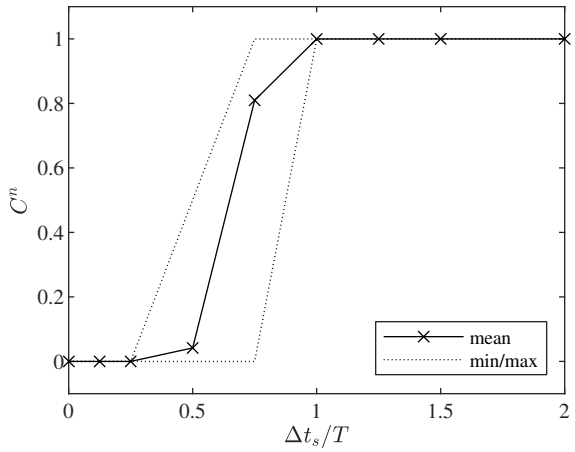


Figure 3-6: Statistics for C^n, the proportion of total volume involved in cycles, using different snapshot intervals for the vortical exchange simulation over $0 < t/T < 6$.

3.4 Volume-conservative Eulerian-based tracking using Eulerian label advection

As section 3.3 shows, to reliably obtain the VTM, we need more information than just the Lagrangian properties of bubbles at t^n and t^{n+1}. Based on the evolution equation (3.2) from which the VTM is derived, an obvious source for additional information is the Eulerian velocity field $\mathbf{u}$ during $t \in (t^n, t^{n+1})$. While some have proposed modifications to the Lagrangian approach which use $\mathbf{u}$ to explicitly provide the non-zero entries of the VTM (Rubel & Owkes, 2019; Langlois *et al.*, 2016), there is still the uniqueness problem.

Here we propose a novel approach where we solve (3.2) directly. As opposed to Lagrangian methods, the use of the velocity field $\mathbf{u}$ through (3.2) gives a fully Eulerian-based tracking method which, along with (3.5), allows direct calculation of the VTM $\mathbf{A}^{(n\to n+1)}$, as outlined in Figure 3-3b. For a numerical advection scheme of $\mathbf{c}^n$ to provide an accurate VTM, we identify two core requirements. The first is that, for the VTM to describe the movement of air, the velocity used to advect the scalar color function c in (2.2) must be exactly equal to the velocity used to advect $\mathbf{c}^n$. As discussed previously, this leads to the consistency requirement (3.3). The second requirement is that, as described by (3.2), each component c_l^n of $\mathbf{c}^n$ must be conserved, ideally to machine precision.

In this section we describe such an advection scheme, the Eulerian label advection (ELA) method. The numerical representation of the vector color function is based on the Volume-of-Fluid (VOF) representation of the scalar color function. The advection scheme itself is based on the conservative Volume-of-Fluid (cVOF) method of Weymouth & Yue (2010). Through its close link to cVOF, ELA provides a consistent advection scheme, which we will show maintains the volume-conservative nature of that approach.

3.4.1 Discretizing the vector color function

Before deriving ELA to solve (3.2), we define a discretization of the vector color function $\mathbf{c}^n(\mathbf{x}, t)$ by integrating it over each cell Ω_{ijk},

$$\mathbf{s}^n_{ijk}(t) \equiv \frac{\int_{\Omega_{ijk}} \mathbf{c}^n(\mathbf{x}, t)\ \mathrm{d}V}{\Delta\Omega_{ijk}}\,. \tag{3.19}$$

We will call this the vector source fraction field. It is a vector equivalent to f_{ijk} defined by (2.14). Recall that the vector color function $\mathbf{c}^n(\mathbf{x}, t)$ is initially defined based on CCL at time t^n, as described by (3.1). Noting that point $\mathbf{x}$ is only in a bubble l if the point $\mathbf{x}$ is also in air ($c(\mathbf{x}, t^n) = 1$), we can integrate (3.1) in the same way as (3.19) to obtain

$$(s^n_l)_{ijk}(t^n) = \frac{1}{\Delta\Omega_{ijk}} \int_{\mathbf{x}\in\left[\Omega_{ijk}\,\cap\,\text{bubble } l\right]} c(\mathbf{x}, t^n)\,\mathrm{d}V \quad \text{for} \quad l \in 0\ldots M^n\,. \tag{3.20}$$

We can simplify this definition by making an assumption about the CCL method used to determine if $\mathbf{x} \in$ bubble l. Typical CCL algorithms assign all of the air in a cell to a single bubble. Thus, the integral in the previous equation would be equal to either zero or the volume of air in the cell. Recalling (2.14), this gives

$$(s^n_l)_{ijk}(t^n) = \begin{cases} f_{ijk}(t^n) & \text{if} \quad \Omega_{ijk} \in \text{bubble } l \\ 0 & \text{otherwise} \end{cases} \quad \text{for} \quad l \in 0\ldots M^n\,, \tag{3.21}$$

where $\Omega_{ijk} \in$ bubble l simply means the CCL algorithm identified the cell ijk as part of bubble l. Because bubbles are defined to be non-overlapping, we note that at $t = t^n$ no more than one element of $\mathbf{s}^n_{ijk}(t^n)$ can be non-zero.

With (3.21), at $t = t^n$ we are able to initialize the vector source fraction field $\mathbf{s}^n_{ijk}(t^n)$ based on the VOF field $f_{ijk}(t^n)$ and the results of a CCL algorithm. This definition of $\mathbf{s}^n_{ijk}(t^n)$ can also be shown to satisfy the consistency requirement. Integrating (3.3) in the same way as (3.19) we obtain a discretized consistency requirement,

$$\sum_l (s^n_l)_{ijk}(t) = f_{ijk}(t)\,. \tag{3.22}$$

Recalling bubbles are non-overlapping and that all air is associated with a bubble, it is clear that (3.21) gives an initial $\mathbf{s}^n_{ijk}(t^n)$ which satisfies (3.22).

3.4.2 ELA method for evolving the vector source fraction

We now seek a consistent and volume-conservative numerical advection scheme to model (3.2). The derivation will closely follow the cVOF derivation (Weymouth & Yue, 2010) presented in section 2.2. First, we integrate (3.2) over a cell Ω_{ijk} and apply divergence

theorem to obtain

$$\frac{\partial}{\partial t}\int_{\Omega_{ijk}} \mathbf{c}^n \, \mathrm{d}V + \oint_{\partial\Omega_{ijk}} \mathbf{c}^n u_n \, \mathrm{d}S = \int_{\Omega_{ijk}} \mathbf{c}^n \left(\boldsymbol{\nabla}\cdot\mathbf{u}\right) \mathrm{d}V \,. \tag{3.23}$$

As in section 2.2, $\partial\Omega_{ijk}$ is the surface of the cell and u_n is the velocity normal to the surface (using the convention that $u_n > 0$ indicates flow out of the cell). For conciseness, we consider a single cell and drop the "ijk" subscripts. Writing in terms of $\mathbf{s}^n$ using (3.19) and rearranging,

$$\Delta\Omega\frac{\partial \mathbf{s}^n}{\partial t} = -\mathbf{F}_{net} + \int_{\Omega} \mathbf{c}^n \left(\boldsymbol{\nabla}\cdot\mathbf{u}\right) \mathrm{d}V \,, \tag{3.24}$$

where the vector $\mathbf{F}_{net}$ describes the net flux of $\mathbf{s}^n$ out of Ω. As we discretize (3.24) in time, to avoid confusion of the snapshot index n based on Δt_s and the fluid-solver time index k based on Δt, we will omit n for the rest of this section, e.g., $\mathbf{s}^k \equiv \mathbf{s}^n(t^k)$.

As discussed in section 2.2, (Weymouth & Yue, 2010) provide an operator-split method to solve the evolution of c (represented by the VOF field f). For convince, the equation to go from the VOF field f^k at time t^k to f^{k+1} at $t^{k+1} = t^k + \Delta t$ in an $\mathcal{N}$ dimensional domain is repeated here,

$$\frac{\Delta\Omega}{\Delta t}\left(f^{(d)} - f^{(d-1)}\right) = F_{d+1/2} - F_{d-1/2} + \tilde{c}\frac{\partial u_d}{\partial x_d}\Delta\Omega \quad \text{for} \quad d \in 1\ldots\mathcal{N} \,. \tag{2.40}$$

Recall that the scalar flux on the positive face ($F_{d+1/2}$) and negative face ($F_{d-1/2}$) are calculated using a second-order interface reconstruction based on $f^{(d-1)}$ (section 2.2.1) and that the dilation term is approximated based on $f^{(0)}$ (section 2.2.2). For ELA, we solve (3.24) using a similar operator-split equation,

$$\frac{\Delta\Omega}{\Delta t}\left(\mathbf{s}^{(d)} - \mathbf{s}^{(d-1)}\right) = \mathbf{F}_{d+1/2} - \mathbf{F}_{d-1/2} + \tilde{\mathbf{c}}\frac{\partial u_d}{\partial x_d}\Delta\Omega \quad \text{for} \quad d \in 1\ldots\mathcal{N} \,, \tag{3.25}$$

where $\mathbf{F}_{d+1/2}$ and $\mathbf{F}_{d-1/2}$ are vector flux terms on the positive and negative faces and $\tilde{\mathbf{c}}$ is a vector dilation term. As discussed in section 3.2, consistency arises from the fact that $\mathbf{c}^n$ is advected with the same $\mathbf{u}$ as c. From this, it is natural that each vector flux term should itself be consistent with the associated scalar flux terms in (2.40), e.g.,

$$\sum_l (F_l)_{d+1/2} = F_{d+1/2} \,. \tag{3.26}$$

If we further require that the vector dilation term is consistent with the scalar dilation term,

$$\sum_l (\tilde{c}_l) = \tilde{c} \,, \tag{3.27}$$

summing (3.25) shows that after each operator-split step the vector source field is consistent with the VOF field,

$$\sum_l (s_l)^{(d)} = f^{(d)} \,. \tag{3.28}$$

By extension, (3.22) will always be satisfied. Thus, any operator-split advection scheme for $\mathbf{s}$ of the form (3.25) which satisfies (3.26) and (3.27) will be consistent with cVOF, or any VOF method that can be written in the same form as (2.40).

In addition to consistency and conservation, it is important that ELA is computationally efficient. (3.25) is a vector equation which implies solving as many advection equations as there are bubbles. While perhaps theoretically possible, applying approaches like interface reconstruction for each individual equation would be computationally infeasible. Instead, we design ELA to use as much information as possible from cVOF to calculate $\mathbf{F}_{d+1/2}$ and $\tilde{\mathbf{c}}$. For convenience, we define the normalized vector source fraction $\hat{\mathbf{s}}$ as

$$\hat{s}_l = \frac{s_l}{\sum_i s_i}\,, \tag{3.29}$$

which has the property $\sum_l \hat{s}_l = 1$. An explicit conservative upwind scheme is used to determine the composition of the air flux based on the previous operator-split step's $\mathbf{s}^{(d-1)}$ and the scalar flux F from cVOF:

$$\mathbf{F}_{d+1/2} = F_{d+1/2} \cdot \begin{cases} \hat{\mathbf{s}}_{d+1}^{(d-1)} & \text{if} \quad F_{d+1/2} > 0 \\ \hat{\mathbf{s}}_{d}^{(d-1)} & \text{if} \quad F_{d+1/2} < 0 \end{cases}, \tag{3.30a}$$

$$\mathbf{F}_{d-1/2} = F_{d-1/2} \cdot \begin{cases} \hat{\mathbf{s}}_{d}^{(d-1)} & \text{if} \quad F_{d-1/2} > 0 \\ \hat{\mathbf{s}}_{d-1}^{(d-1)} & \text{if} \quad F_{d-1/2} < 0 \end{cases}. \tag{3.30b}$$

To describe the vector dilation term based on the scalar dilation term $\tilde{c}$ from cVOF,

$$\tilde{\mathbf{c}} = \tilde{c}\,\hat{\mathbf{s}}^{(0)}\,. \tag{3.31}$$

For volume conservation, it is critical that $\tilde{\mathbf{c}}$ remain the same throughout the operator-split steps, so (similar to cVOF) we base it on the initial $\mathbf{s}^{(0)}$.

ELA volume conservation

Recall the three requirements Weymouth & Yue (2010) identify for a volume-conservative (to machine precision) operator-split advection scheme:

1. flux terms are conservative,
2. the dilation terms sum to zero, and
3. there is clipping due to over or under filling at any step.

For the flux terms, ELA satisfies (3.26), so, by extension from cVOF, the total flux of $\mathbf{s}$ is conservative. Additionally, the use of upwinding in (3.30) guarantees that the flux of each individual component of $\mathbf{s}$ is conservative. For the dilation term, a similar argument is true. ELA satisfies (3.27) so, by extension from cVOF, the total of the dilation terms sums to zero. Because each component of $\tilde{\mathbf{c}}$ remains constant, each individual element of the dilation terms also sums to zero. What remains is to prove requirement 3.

Because ELA is consistent with cVOF, (3.28) is true at any step. As proven by Weymouth

& Yue (2010), cVOF with the Courant restriction

$$\Delta t \sum_{d=1}^{\mathcal{N}} \left| \frac{u_d}{\Delta x_d} \right| < C \tag{2.13}$$

guarantees $0 \leq f^{(d)} \leq 1$ at any step. By extension, ELA satisfies

$$0 \leq \sum_l (s_l)^{(d)} \leq 1\,, \tag{3.32}$$

establishing that the sum of $\mathbf{s}^{(d)}$ cannot over or under fill. However, to show ELA is volume conservative, we must show that individual components cannot over or under fill, i.e.,

$$0 \leq (s_l)^{(d)} \leq 1 \tag{3.33}$$

for all l. First, we note that, given (3.32) is true, proving $0 \leq (s_l)^{(d)}$ proves (3.33). In Appendix C, by considering all possible combinations of the sign of the velocity on either face, we prove that $0 \leq (s_l)^{(d)}$ for all l provided the Courant restriction (2.13) is true.

While based on the proof by Weymouth & Yue (2010), the proof in Appendix C is not a trivial extension because, unlike with cVOF, it is not guaranteed the flux in on one side of a cell replaces the flux out on the other side, as they could be composed of different elements of $\mathbf{s}$. Surprisingly, we still find that the Courant restriction required for ELA volume conservation is no more restrictive than that for cVOF volume conservation. Meaning that no change in Δt is needed to use ELA with cVOF versus cVOF alone.

Finite precision considerations

The ELA algorithm we have described so far is exactly volume conservative and consistent with cVOF. In practice, floating point calculations have finite precision, so a few modifications are useful. First, finite precision leads to rounding errors, which may cause an element $(s_l)^{(d)}$ that should go to zero to be either slightly above or below zero. To deal with this, after calculating each step in (3.25), the following filter is applied,

$$(s_l)^{(d)} \leftarrow \begin{cases} 0 & \text{if} \quad (\hat{s}_l)^{(d)} < \epsilon_{\text{mach.}} \\ (s_l)^{(d)} & \text{otherwise} \end{cases}\,. \tag{3.34}$$

Here, $\epsilon_{\text{mach.}}$ is defined to be the smallest value such that with finite precision $1 + \epsilon_{\text{mach.}} \neq 1$. With this filter, can say that the consistency condition (3.28) is true *to machine precision*.

Even without the above filter, finite precision means that (3.28) will not be exactly true. While the error is on the order of machine precision, it can be useful to further minimize this error as much of the volume conservation proof relies on (3.28). To this end, we can apply the following filter after (3.34),

$$\mathbf{s}^{(d)} \leftarrow f^{(d)} \hat{\mathbf{s}}^{(d)}\,. \tag{3.35}$$

The performance of ELA with and without this filter are presented in §3.5.2.

Finally, we recall that cVOF implementations often include a filter (2.43) designed to

suppress spurious whips and ensure symmetry between operations on f and $1-f$. This introduces a zero-threshold value ϵ_f and a volume conservation error of the same order of magnitude. To ensure (3.28) is always true, as required for ELA volume conservation, $\mathbf{s}$ must be similarly adjusted by the same ϵ_f at the end of the operator split steps:

$$\mathbf{s}^{k+1} = \begin{cases} \hat{\mathbf{s}}^{(\mathcal{N})} & \text{if} \quad 1 - f^{(\mathcal{N})} < \epsilon_f \\ \mathbf{0} & \text{if} \quad f^{(\mathcal{N})} < \epsilon_f \\ \mathbf{s}^{(\mathcal{N})} & \text{otherwise} \end{cases} . \tag{3.36}$$

While neither cVOF nor ELA are strictly conservative if $\epsilon_f \neq 0$, the total volume addition/loss in $\mathbf{s}$ is equal to that in f. Therefore, ELA continues to satisfy the consistency requirement and satisfies the necessarily weakened volume-conservation requirement in the sense that it tracks all the air, including that artificially added/subtracted by (2.43).

3.4.3 Extracting the VTM from the vector source fraction

Using the method described in section 3.4.2 we are able to evolve the vector source fraction $\mathbf{s}^n_{ijk}(t)$ in time from its original definition at $t = t^n$. We perform (3.25) at every simulation time step Δt until we reach the next snapshot interval $t^{n+1} = t^n + \Delta t_s$. From this, we have the discretized representation

$$\mathbf{s}^n_{ijk}(t^{n+1}) = \frac{\int_{\Omega_{ijk}} \mathbf{c}^n(\mathbf{x}, t^{n+1}) \ \mathrm{d}V}{\Delta\Omega_{ijk}} . \tag{3.37}$$

To relate this to the VTM, we start by expanding (3.6) into a summation over all (non-overlapping) cells in the domain,

$$q_{ml} = \sum_{ijk} \int_{\mathbf{x} \in \left[\Omega_{ijk} \cap \text{bubble } m\right]} c^n_l(\mathbf{x}, t^{n+1}) \, \mathrm{d}V . \tag{3.38}$$

As done for (3.21), we assume the CCL algorithm assign all of the air in a cell to a single bubble, giving

$$q_{ml} = \sum_{\Omega_{ijk} \in \text{bubble m}} \int_{\Omega_{ijk}} c^n_l(\mathbf{x}, t^{n+1}) \, \mathrm{d}V . \tag{3.39}$$

Using (3.37), we have

$$q_{ml} = \sum_{\Omega_{ijk} \in \text{bubble m}} \Delta\Omega_{ijk} \ (s_l)^n_{ijk}(t^{n+1}) . \tag{3.40}$$

After normalizing using (3.9), we obtain the VTM $\mathbf{A}^{n \to n+1}$.

After obtaining $\mathbf{A}^{n \to n+1}$, we can use the same CCL results at t^{n+1} used to determine $\Omega_{ijk} \in$ bubble m in (3.40) to define a new $\mathbf{s}^{n+1}_{ijk}$ with (3.19), and repeat the process of evolving it using ELA and then extracting $\mathbf{A}^{n+1 \to n+2}$. Thus, this method provides tracking information throughout a simulation.

3.5 Validation for three-dimensional bubble fragmentation in homogeneous isotropic turbulence

To verify and demonstrate the properties of ELA, we consider the canonical problem of a low void-fraction distribution of air-bubbles in water (density ratio $\rho_w/\rho_a = 1000$) fragmented by strong homogeneous isotropic turbulence (HIT). We choose HIT as it is spatially homogeneous, quasi-steady, and well understood, allowing simple measurement of averaged turbulence properties, particularly the turbulent dissipation rate ε. Additionally, multi-phase HIT has been well studied experimentally (e.g., Qi *et al.*, 2020; Martínez-Bazán *et al.*, 1999*a*; Vejražka *et al.*, 2018) and serves as a building block for understanding bubbly-flow near an air entraining free surface. Using this canonical problem, we verify that ELA is volume conservative and quantify the abundance of loop currents. In Appendix B this same simulation is used to demonstrate matrix multiplication and the associated trade-off between cost and accuracy.

3.5.1 Simulation setup

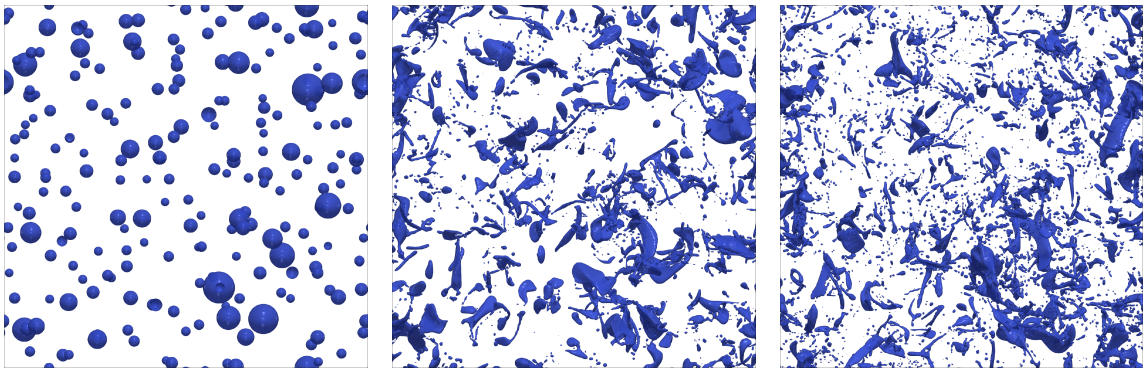
We perform three-dimensional DNS on a triply periodic grid of 256^3. For simplicity, we consider the case where surface tension is negligible compared to the strength of turbulence and set $We = \infty$. HIT is created and maintained using a linear forcing method (Lundgren, 2003; Rosales & Meneveau, 2005) resulting in a turbulent Reynolds number $Re = u_{\text{rms}}^4/\varepsilon\nu = 200$ and Kolmogorov scale $\eta \approx \Delta x$. After turbulence is initialized for one phase, a population of spherical bubbles with radii between $r_{min} = 3\Delta x$ and $r_{max} = 15\Delta x$ following a $r^{-10/3}$ power law is randomly distributed without overlap at $t = 0$ such that the void fraction is 1%. The simulations are performed using the cVOF method described in section 2.2, with $\epsilon_f = 10^{-12}$ used for (2.43) and (3.36). We choose the Informed Component Labeling (ICL) algorithm developed by Hendrickson *et al.* (2020) as the CCL method to provide ELA with contiguous bubbles because it is volume conservative, i.e., all dark fluid is marked as part of a bubble.[3] This allows us to validate that ELA is volume conservative.

For fragmentation of bubbles in HIT, a characteristic time is $t_b = 0.42\,\varepsilon^{-1/3} {r_{max}}^{2/3}$, corresponding to the typical lifetime of the largest bubbles (Martínez-Bazán *et al.*, 1999*a*). We run our simulations over $0 < t/t_b < 2$, over which we observe an increase from 215 to 588 resolved bubbles (defined as bubbles with a volume larger than $v_{\text{res}} = 4/3\pi(1.5\Delta x)^3$). The evolution of the bubble field is shown in figure 3-7 and the bubble size distribution is shown in figure 3-8. We perform a series of otherwise identical runs using different $\tau \equiv \Delta t_s/t_b$ (see table 3-1). For the largest, we choose $\tau = 0.1$, consistent with recommendations by Chan *et al.* (2021*a*). Figure 3-2 provides an event observed by ELA over this snapshot interval and the corresponding tracking matrix from a subset of the domain.

3.5.2 Validating ELA volume conservation

Using ELA with the longest snapshot interval ($\tau = 0.1$), we first seek to verify the conservative nature of ELA. Following Weymouth & Yue (2010), a measure of the typical change in the

[3] For details on ICL including comparison to other CCL methods, see §2.3



(a) $t/t_b = 0$ **(b)** $t/t_b = 1$ **(c)** $t/t_b = 2$

Figure 3-7: Evolution of the $f = 0.5$ iso-surface in the three-dimensional HIT simulation (without surface tension).

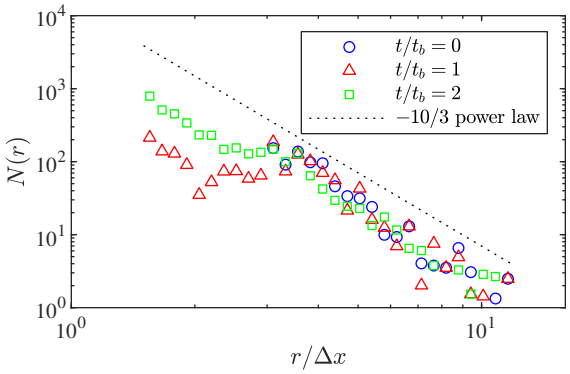


Figure 3-8: Evolution of the bubble size distribution $N(r)$ in the HIT simulation versus the expected $-10/3$ power law. Note, only resolved bubbles ($r/\Delta x \geq 1.5$) are considered.

	$\tau = \Delta t_s/t_b$	$K = \langle \Delta t_s/\Delta t \rangle$	Normalized with (3.35)?
Case La	0.1000	109.9	no
Case Lb	0.1000	109.9	yes
Case S1	0.0500	54.96	yes
Case S2	0.0250	27.50	yes
Case S3	0.0125	13.75	yes
Case S4	0.0063	6.870	yes
Case S5	0.0031	3.434	yes
Case S6	0.0016	1.717	yes

Table 3-1: Summary of different runs performed using the same flow but different ELA settings. Note, the simulation time step Δt is chosen dynamically as described in §2.1.2, so the average value based on all snapshot intervals is reported for K.

	$(L_1)_{\text{ELA}}/V^0$	$(\text{Change})_{\text{ELA}}$	$(L_1)_{\text{cVOF}}/V^0$	$(\text{Change})_{\text{cVOF}}$
Case La	8.2×10^{-16}	-1.9×10^{-12}	5.7×10^{-13}	1.9×10^{-9}
Case Lb	2.2×10^{-17}	4.8×10^{-14}	"	"

Table 3-2: The L_1 and relative change metrics for volume conservation error separated into ELA contribution and cVOF contribution for HIT simulations over $0 < t/t_b < 2$ corresponding to $T = 30$ snapshot intervals. Note that the L_1 errors are per fluid solver step while relative change errors are over the entire simulation (~ 3000 steps).

volume of dark fluid per simulation time step Δt is

$$(L_1)_{\text{cVOF}} = \frac{1}{TK} \sum_{n=0}^{T-1} |V^{n+1} - V^n| \,, \tag{3.41}$$

Here, TK is the total number of simulation time steps. The total relative change over the entire simulation is

$$(\text{Change})_{\text{cVOF}} = \frac{V^T - V^0}{V^0} \,. \tag{3.42}$$

We define equivalent metrics for ELA based on the tracking matrices:

$$(L_1)_{\text{ELA}} = \frac{1}{TK} \sum_{n=0}^{T-1} \left\| \mathbf{v}^{n+1} - \mathbf{A}^{(n \to n+1)} \mathbf{v}^n \right\|_1 \,, \tag{3.43}$$

$$(\text{Change})_{\text{ELA}} = \frac{\sum \left[\mathbf{v}^T - \tilde{\mathbf{A}}^{(0 \to T)} \mathbf{v}^0 \right]}{\sum \left[\mathbf{v}^T \right]} \,. \tag{3.44}$$

By comparing the new volume predicted by the VTM to the new volume calculated from the void fraction field, we measure any volume error in ELA separate from that related to cVOF.

The results for the simulations are shown in table 3-2. We first run ELA without the additional normalization of (3.35) (Case La). As expected, $(L_1)_{\text{cVOF}}/V^0 \sim O(\epsilon_f)$, indicating that the small growth in cVOF volume error is due to (2.43). Therefore, if ϵ_f is changed we expect $(L_1)_{\text{cVOF}}$ and $(\text{Change})_{\text{cVOF}}$ to change proportionally. As $\mathbf{s}$ remains consistent with f through (3.36), ϵ_f does not affect the ELA metrics in table 3-2. The growth of the ELA volume conservation error per step, $(L_1)_{\text{ELA}}/V^0$, is approximately machine precision, validating that the ELA method is volume conservative to machine precision.

Case La validates that, as expected, ELA is volume conservative to machine precision. To reduce the accumulation of error related to machine precision, we can use the additional normalization step of (3.35). Case Lb in table 3-2 shows that this achieves a $1/40$ reduction in the (already near machine precision) error. All further use of the ELA method includes (3.35) unless stated otherwise.

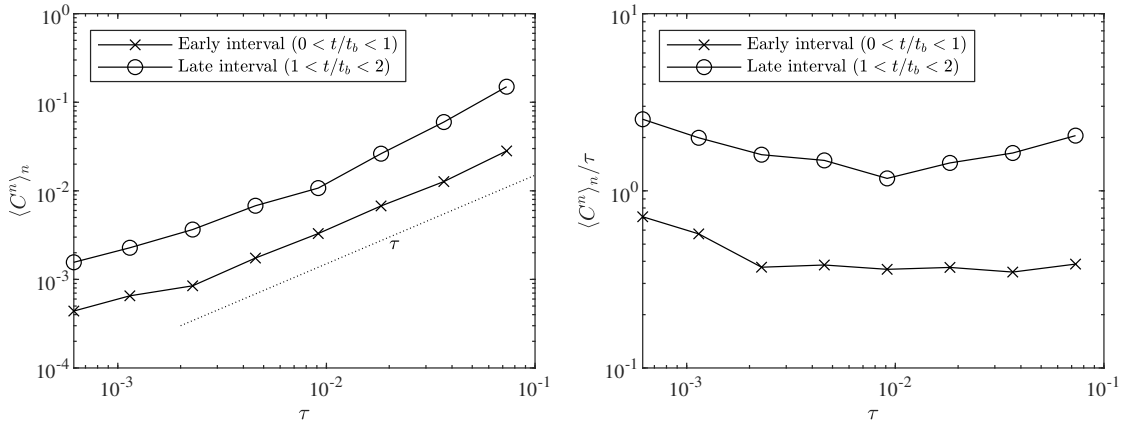


Figure 3-9: The average proportion of the volume of resolved bubbles involved in cycles C^n (a), and the average rate of change C^n/τ (b), for HIT simulations with different snapshot intervals $\tau = \Delta t_s/t_b$ (see table 3-1), compared to a linear relationship.

3.5.3 Relationship between cycle generation and snapshot interval

For the HIT simulations, we now evaluate cycle production using the same analysis method described in §3.3.3. To evaluate the effect of the number of bubbles and events, two different time periods are considered: an early period $0 < t/t_b < 1$ covering 15 snapshot intervals with fewer bubbles and events, and a late period $1 < t/t_b < 2$ covering 15 snapshot intervals with more bubbles and events. Noting that we track all bubbles, not just those that are resolved ($v > v_{\text{res}}$), the largest value of M^n is 9×10^4 for the early and 2×10^5 for the late period. To avoid inflating the count of cycles by including under-resolved events, we first remove columns of $\mathbf{A}^{(n \to n+1)}$ relating to under-resolved parent bubbles, $v_i^n < v_{\text{res}}$, and rows relating to under-resolved child bubbles, $v_j^{n+1} < v_{\text{res}}$. By repeating the simulations with different Δt_s (see table 3-1), we obtain a range of $\tau = \Delta t_s/t_b$. The results are shown in figure 3-9.

Previous work on numerical tracking (Chan *et al.*, 2021*a*; Gao *et al.*, 2021; Rubel & Owkes, 2019) has identified that over small time intervals, CCL causes chains of spurious fragmentation and coalescence, as CCL methods struggle to consistently identify distinct regions of dark fluid separated by lengths on the order of the grid. Generally, interfaces can be arbitrarily close, making this a fundamental limitation of CCL (Herrmann, 2010). When tracking bubbles identified by CCL, imposing a minimum Δt_s typically mitigates the inclusion of spurious events (Chan *et al.*, 2021*a*; Gao *et al.*, 2021). As this strategy is adopted to improve the CCL information provided to tracking, it applies to all tracking methods, including ELA. Note that the characteristic time period of spurious events, and thus appropriate minimum Δt_s, is likely sensitive to *both* the fluid solver and the CCL method.

Cycle production depends on the number of events and number of bubbles, as illustrated by the difference between the early and late time periods; however, over both time periods an approximately linear scaling $C^n \propto \tau$ is observed. Considering the magnitude of C^n, we observe that, for $\tau = O(0.1)$ proposed by Chan *et al.* (2021*a*), 5% of the resolved volume is

involved in cycles for the early period and 20% of the resolved volume is involved in cycles for the late period. This means previous tracking methods would be unable to provide unique solutions to the tracking matrix, and the differences between solutions could be significant.

Given its Eulerian nature, ELA itself will not introduce cycles, however it is not necessarily clear which cycles are the result of either physical exchanges of volume (as illustrated in figure 3-4 and figure 3-5) or spurious events caused by CCL. Because tracking considers CCL given, ELA cannot directly quantify what portion of these cycles fall into each category; however, the results in figure 3-9 do suggest that physical cycles are abundant in this simulation. As C^n provides a measure of the total cycle production over an interval τ, C^n/τ provides a measure of the rate of cycle production. For small $\tau \lesssim 10^{-2}$, C^n/τ decreases with increasing τ, consistent with the decreasing probability of spurious events. For the late interval and $\tau > 10^{-2}$, C^n/τ increases with increasing τ, consistent with the increasing probability of physical cycles (illustrated in §3.3.3). This suggests that physical cycles dominate at large τ in the late interval of this simulation.

Previous Lagrangian tracking methods were often limited to small Δt_s because their binary assumption precludes identifying events over long time periods (Chan *et al.*, 2021*a*) and their identification of advection is inaccurate over large displacements (Gao *et al.*, 2021). Due to its Eulerian volume-tracking nature, ELA does not suffer these limitations and is accurate for large Δt_s. With ELA making large Δt_s possible, it is now easier to select a Δt_s sufficiently large to negate the effect of spurious events caused by CCL. On selecting Δt_s in practice, there is also physical motivation, as Δt_s defines the distinction between a single event and multiple events, and thus directly affects measured event statistics (Vejražka *et al.*, 2018; Solsvik *et al.*, 2016). The effect of Δt_s on the physical meaning of measured fragmentation statistics will be addressed in Chapter 5.

Chapter 4

Characterizing the Surface Layer of Strong Free-Surface Turbulence

In subsequent chapters we will show how each of the bubble evolution mechanisms (fragmentation in Chapter 5, entrainment in Chapter 6, and degassing in Chapter 7) depends on the near-surface turbulence, in particular the turbulent dissipation rate ε and the characteristic turbulent velocity u_{rms}. To apply any of the subsequent models we develop for these evolution mechanisms, one needs an accurate method to predict the near-surface turbulence. Therefore, before we address bubble evolution, this chapter focuses on characterizing air entraining free-surface turbulence, particularly the surface layer where air and water are highly mixed and turbulence modeling is the most challenging.

4.1 Introduction

For free-surface turbulence, the characteristic velocity is the turbulence fluctuations u_{rms} and the characteristic length scale is $L_T = u_{\text{rms}}^3/\varepsilon$, where ε and u_{rms} are measured near the surface. As discussed in §1.1 the primary quantity for characterizing FST is the turbulent Froude number (squared),

$$Fr_T^2 = \frac{u_{\text{rms}}^2}{L_T \text{g}}\,, \tag{1.1}$$

or equivalently $Fr_T^2 = \varepsilon/u_{\text{rms}}\text{g}$. For $Fr_T^2 \ll 1$ the free surface is only slightly deformed and vertical turbulent fluctuations near the surface are suppressed, referred to as the blockage effect. The blockage effect's suppression of vertical fluctuations and corresponding transfer of energy to horizontal fluctuations leads to highly anisotropic, quasi-two-dimensional turbulence (Shen *et al.*, 1999; Guo & Shen, 2010; Ruth & Coletti, 2024). The strength of the blockage effect decreases with increasing Fr_T^2. Recently, Yu *et al.* (2019) observed that for very large Fr_T^2 the near-surface turbulence is nearly isotropic. They named this regime strong FST and theorized it may correspond to the onset of entrainment; however, the critical Fr_T^2 above which the strong FST regime is obtained was not determined.

The most observable effect of large Fr_T^2 FST (isotropic or not) is that the free surface becomes highly distorted, to the point where it breaks apart forming bubbles (and droplets) (Brocchini & Peregrine, 2001*a*). Brocchini & Peregrine (2001*b*) provide a notational sketch

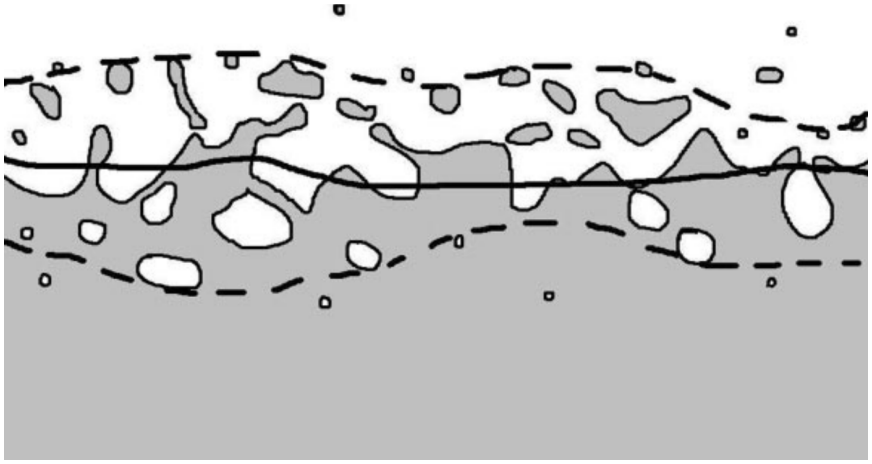


Figure 4-1: Notational sketch of the surface layer from 'The dynamics of strong turbulence at free surfaces. Part 2. Free-surface boundary conditions' by Brocchini & Peregrine (2001*b*). © 2001 Cambridge University Press.

of the surface layer (reprinted in figure 4-1), where the mixing of the air and water means there is no clearly distinguishable free surface. In place of the free surface is a region where water and air are highly mixed, called the surface layer. While the creation of bubbles and subsequent evolution is our ultimate interest, these processes are driven by the near-surface turbulence, especially the turbulence within the surface layer. Due to the highly variable density in this surface layer, turbulence modeling is a particular challenge (Brocchini & Peregrine, 2001*a*; Hendrickson & Yue, 2019).

One air entraining turbulent flow which is often studied is supercritical open-channel flow (Chanson, 1996). One approach to modeling the boundary layer turbulence in this flow is to separately model the surface layer (referred to as turbulent wavy layer) and the turbulent boundary layer beneath (Killen, 1968; Kramer & Valero, 2023). We note that in these flows there is typically not a large separation of scales: the height of the surface layer is similar to the total depth of the flow. This implies a close relationship between the dynamics in the surface layer and the turbulent boundary layer beneath, making it a challenge to generalize models for open-channel flow to general free-surface flow. Looking at free-surface wakes, Hendrickson *et al.* (2019) noted that the resulting air entraining FST bears little similarity to open-channel flow. For describing surface layers in general free-surface flow, Brocchini & Peregrine (2001*b*) provide a theoretical framework based on phase-weighted averaging. For the case where the surface layer is thin relative to the large-scale features of the flow, Brocchini (2002) shows theoretically how the surface layer could be treated as a modified boundary condition in k-ε type Reynolds-averaged Navier Stokes (RANS) simulations.

We consider air entraining FST where the mechanisms that generate turbulence in the water are sufficiently separate (in either distance or scale) from the surface layer. We study this using direct numerical simulation (DNS) of a canonical example of such a flow, where homogeneous, statistically steady turbulence is generated in a region deep beneath the free surface (Guo & Shen, 2009). In §4.2 we develop a quantitative definition of the surface layer thickness δ_s, which unlike previous definitions (Brocchini & Peregrine, 2001*b*) is independent of the (Fr_T^2-dependent) quantity of bubbles deep beneath the surface and makes

no assumptions about the distribution of intermittency. Using measurements from DNS across a wide range of Fr_T^2 (described in §4.3), in §4.4 we quantify the distribution of intermittency, confirming that it collapses after scaling by δ_s and showing that it is well described by a logistic distribution. In §4.5 we study the turbulence in the surface layer. We show that $Fr_T^2 > 0.1$ is the criteria for strong FST, when blockage effects are small and the turbulence in the surface layer is isotropic. For strong FST, we find that the turbulence within the surface layer only depends on δ_s and the turbulence quantities at the bottom of the surface layer. These results confirm the possibility of modeling the surface layer as a modified boundary condition in RANS, especially for $Fr_T^2 > 0.1$. In §4.6 we show how the two necessary quantities for such a boundary condition, the surface layer thickness δ_s and the work done on the surface layer W, scale with Fr_T^2.

4.2 Defining the surface layer in free-surface turbulence

The defining property of the surface layer is the heterogeneous mixture of air and water. As introduced in Chapter 2, the mixture of any two immiscible phases can be described by a binary color function (2.2). For air and water,

$$c(\mathbf{x}, t) \equiv \begin{cases} 0 & \text{if} \quad \mathbf{x} \in \text{air} \\ 1 & \text{if} \quad \mathbf{x} \in \text{water} \end{cases} . \tag{4.1}$$

Our interest is the average behavior of the surface layer, and we consider the simplest case where FST is horizontally and temporally homogeneous. Averaging along the homogeneous dimensions,

$$\bar{\cdot} \equiv \frac{\iiint \cdot \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t}{\iiint \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t} , \tag{4.2}$$

we define the intermittency factor (Brocchini & Peregrine, 2001*b*)

$$\gamma(z) \equiv \overline{c(\mathbf{x}, t)} , \tag{4.3}$$

which here is only a function of depth z. This intermittency factor gives the average portion of volume occupied by water, with $\gamma = 1$ corresponding to all water and $\gamma = 0$ all air. The opposite, $1 - \gamma$, is often referred to as the void fraction.

Because the free surface is highly distorted, we first need a definition of the mean free surface. While alternative definitions may be more appropriate for wave-driven flows (Brocchini & Peregrine, 1996, 2001*b*), for FST we find a simple definition is sufficient. We define the mean free-surface height $\bar{\eta}$ as the height where air and water are evenly mixed:

$$\gamma(z = \bar{\eta}) = 0.5 . \tag{4.4}$$

We note that, because FST entrains air bubbles deep beneath the free surface, $\bar{\eta}$ will be higher than the location of the quiescent free surface. In this way, $\bar{\eta}$ is a measure of the entire bubble population, which is the focus of subsequent chapters. In this chapter, our interest is the turbulence near the free surface rather than bubbles deep beneath, so we will

focus on depths relative to the mean free surface, $z - \bar{\eta}$.

Recalling that the defining property of the surface layer is the mixture of air and water, a common way to define the surface layer is to choose a $\gamma_{\max}$ and $\gamma_{\min}$, and define the surface layer as the depths z where $\gamma_{\max} > \gamma(z) > \gamma_{\min}$ (Brocchini & Peregrine, 2001*b*). Others express this in terms of density $\rho_{\max} > \overline{\rho}(z) > \rho_{\min}$ (Hendrickson & Yue, 2019), which is equivalent. The first challenge is that the choice of these limits is somewhat arbitrary. To justify a choice, Brocchini & Peregrine (2001*b*) start by considering a single-valued free surface where the instantaneous free-surface height $\eta(x, y, t)$ follows a Gaussian distribution with mean $\bar{\eta}$ and variance σ^2. It can be shown that this gives an intermittency factor

$$\gamma_{\text{Gaussian}}(z) = \frac{1}{2}\left[1 - \text{erf}\left(\frac{z - \bar{\eta}}{\sqrt{2}\sigma}\right)\right] . \tag{4.5}$$

Using $z - \bar{\eta} \in [-3\sigma, 3\sigma]$ to define the surface layer gives $\gamma_{\min} \approx 0.001$ and $\gamma_{\max} \approx 0.999$. In §4.4 we will show that FST produces a non-negligible volume of bubbles deep beneath the free surface, causing γ to approach 1 very slowly. Due to these bubbles, we find $\gamma_{\max} \approx 0.999$ overestimates the depth of the surface layer. More generally, the challenge here is that the definition of the surface layer depends on the tail behavior of γ, which may be only indirectly linked to the behavior closest to the mean free surface.

As a more robust definition of the surface layer, we use the behavior of γ at the mean free surface rather than the tail behavior far from the mean free surface. To do this, we define a free surface thickness

$$\delta_s \equiv \frac{6}{\sqrt{2\pi}}\left(\left.\frac{\mathrm{d}\gamma}{\mathrm{d}z}\right|_{z=\bar{\eta}}\right)^{-1} . \tag{4.6}$$

and define the surface layer as $z - \bar{\eta} \in [-\delta_s/2, \delta_s/2]$. We choose the constant $6/\sqrt{2\pi}$ so that in the case where $\gamma = \gamma_{\text{Gaussian}}$, $\delta_s = 6\sigma$ and our definition of the surface layer is equivalent to $z - \bar{\eta} \in [-3\sigma, 3\sigma]$.

Asserting that δ_s characterizes the surface layer, it is useful to define an aptly nondimensionalized depth,

$$z^* \equiv \frac{z - \bar{\eta}}{\delta_s} , \tag{4.7}$$

and the surface layer is $z^* \in [-0.5, 0.5]$. Using DNS of FST at a wide range of Fr_T^2, we will show in section 4.4 and section 4.5 that this scaling indeed characterizes the surface layer. Specifically, scaling by z^* and turbulence properties measured at $z^* = -0.5$ collapses measurements in the surface layer across Fr_T^2; and $z^* = -0.5$ separates the surface layer where turbulence is directly affected by the free surface from the region beneath only indirectly affected by the free surface. These two insights suggest the possibility of distinct turbulence closure modeling within the surface layer, and in section 4.6 we explore this and show how δ_s can be predicted *a priori*.

4.3 Direct numerical simulation of statistically steady forced free-surface turbulence

To study FST, we pick a flow which isolates FST from other free surface mechanisms (e.g., waves) and is homogeneous in the horizontal and temporal dimensions. This is obtained with forced free-surface turbulence (forced FST), where isotropic turbulence is continuously forced deep beneath the surface to obtain statistically steady turbulence at the surface. This simulation setup was first described by Guo & Shen (2009) for small Fr_T^2 and has recently been extended to moderate and large Fr_T^2 (Calado & Balaras, 2025; Gaylo & Yue, 2025).

We use the forcing method described in section 2.1.5 to maintain isotropic turbulence in a bulk region deep beneath the free surface. Recall, this adds a linear forcing term

$$\mathbf{f} = A\mathbf{u}''\mathcal{F}(z)\,. \tag{2.32}$$

to the Navier-Stokes equation, (2.5). We choose A dynamically using (2.30), where we set the target dissipation rate $\varepsilon_{\text{target}} = 1$. The horizontal domain length is 2π in both x and y, which gives $u_{\text{rms}} \approx 1$ in the forcing region (Rosales & Meneveau, 2005). In this way, we choose the simulation setup to give characteristic turbulent scales of unity within the forcing region. An equivalent interpretation is that we use the characteristic turbulent scales within the forcing region to nondimensionalize all values in the simulation, i.e., to go from (2.1b) to (2.5).

Following Guo & Shen (2009), we use

$$\mathcal{F}(z) = \begin{cases} 0 & z_c > l_b + l_d \\ \frac{1}{2}\left(1 - \cos\left[\frac{\pi}{l_d}(z_c - l_b - l_d)\right]\right) & z_c \le l_b + l_d \\ 1 & z_c \le l_b \end{cases} ; \quad z_c \equiv |z + l_f + l_d + l_b| \tag{4.8}$$

to define the region deep beneath the free surface where forcing is applied. The forcing is primarily contained to a forcing region of height $2l_b = 3\pi$ centered around $z_c = 0$. There is damping region $l_d = \pi/2$ on either side of the forcing region, then a free length $l_f = \pi/2$ before the bottom of the domain or the quiescent free surface at $z = 0$. For the majority of the simulations there is an air gap of π above the quiescent free surface, to give a total domain height 6π. For $Fr^2 > 2$, the air gap is increased to 1.75π to avoid the top boundary affecting the turbulence.

For all simulations, we set $Re = 200$ and use a constant grid $256^2 \times 768$ or $256^2 \times 864$ depending on the air gap. This gives a grid spacing $\Delta \approx 0.025$. Comparing the grid spacing to the Kolmogorov microscale η_T, in the bulk region where turbulence is strongest $\eta_T/\Delta \approx 0.78$, and near the free surface where we are interested $\eta_T/\Delta \approx 1.8$. As discussed in section 2.1.1, this is sufficient resolution for DNS of air entraining FST. To obtain different conditions, we change Fr^2 and We in (2.5), with $We = \infty$ corresponding to cases where surface tension is not modeled. Table 4-1 shows forced FST simulations performed for the analysis of turbulence here. In this chapter our focus is the scaling of turbulence with Fr_T^2 so all simulations use $We = \infty$. In Chapter 6 we will address the effects of (weak) surface tension.

Fr^2	We	$\bar{\eta}$	δ_s	u_{rms}	ε	L_T	Fr_T^2	Re_T
0.3	∞	0.010	0.41	0.27	0.025	0.81	0.03	44
0.6	∞	0.030	0.77	0.30	0.030	0.89	0.06	53
0.9	∞	0.080	1.08	0.34	0.030	1.32	0.08	90
1.2	∞	0.124	1.29	0.27	0.030	0.69	0.13	38
1.5	∞	0.180	1.40	0.27	0.028	0.73	0.15	40
1.8	∞	0.268	1.60	0.30	0.030	0.85	0.18	50
2.1	∞	0.345	1.73	0.27	0.027	0.70	0.21	37
2.4	∞	0.363	1.78	0.26	0.024	0.74	0.22	39
2.7	∞	0.543	1.76	0.26	0.024	0.69	0.26	35

Table 4-1: List of forced FST simulations used for turbulence analysis. Turbulence properties u_{rms} and ε are measured at the bottom of the surface layer ($z^* = -0.5$). The characteristic length scale $L_T = u_{\text{rms}}^3/\varepsilon$ is used to calculate the near-surface turbulent Froude number (squared) $Fr_T^2 = (u_{\text{rms}}^2/L_T)Fr^2$ and the turbulent Reynolds number $Re_T = (u_{\text{rms}}L_T)Re$.

Simulations are initialized at $t = 0$ from a random velocity in the bulk and a quiescent free surface at $z = 0$ and then run for at least 400 bulk eddy turnover times. We consider total gravitational potential energy, turbulence statistics, the bubble population, and entrainment statistics to assess convergence to a steady state. We find that total gravitational potential energy,

$$PE_{\text{g}} = \frac{1}{Fr^2} \iiint \rho z \, \mathrm{d}V - C \tag{4.9}$$

where the constant C is chosen such that $PE_{\text{g}} = 0$ at $t = 0$, is the most useful for assessing steady state convergence. One concern for these horizontally periodic simulations is that a standing wave will form. Recalling the horizontal domain length is 2π, the standing wave would have wave number $\kappa \approx 1$ and (by the deep-water dispersion relation) frequency $\omega \approx 1/Fr$. The characteristic timescale of the turbulence in the bulk $T = L/U$ is unity, and we could expect some coupling if T/ω is close to an integer multiple. Indeed, we had some trouble with $Fr^2 = 0.9$ (where $T/\omega = 0.95 \approx 1$), but were able to results by considering an earlier time before the standing wave developed. We were unable to obtain a steady state for $Fr^2 = 3.3$ (where $T/\omega = 1.82 \approx 2$).

Figure 4-2 shows a representative rendering near the free surface of a simulation once its reached steady state. After a statistically steady state is reached at $t = t_0$, steady-state statistics are obtained over $t \in [t_0, t_0 + T_{\text{sim}}]$. For the turbulence analysis in this chapter, we fix $T_{\text{sim}} = 128$. Over this period, we use 400 evenly spaced samples for temporal averaging in (4.2).

While the simulation setup allows us to prescribe the turbulence levels in the bulk region, our interest is the turbulence levels near the surface. Guo & Shen (2009) provide an empirical fit for $Fr^2 \ll 1$; however, the near surface turbulence is not known *a priori*. Instead, we measure the near surface turbulence *a posteriori* and calculate new non-dimensional numbers to properly characterize the near-surface turbulence. At $z^* = -0.5$, we measure the turbulent

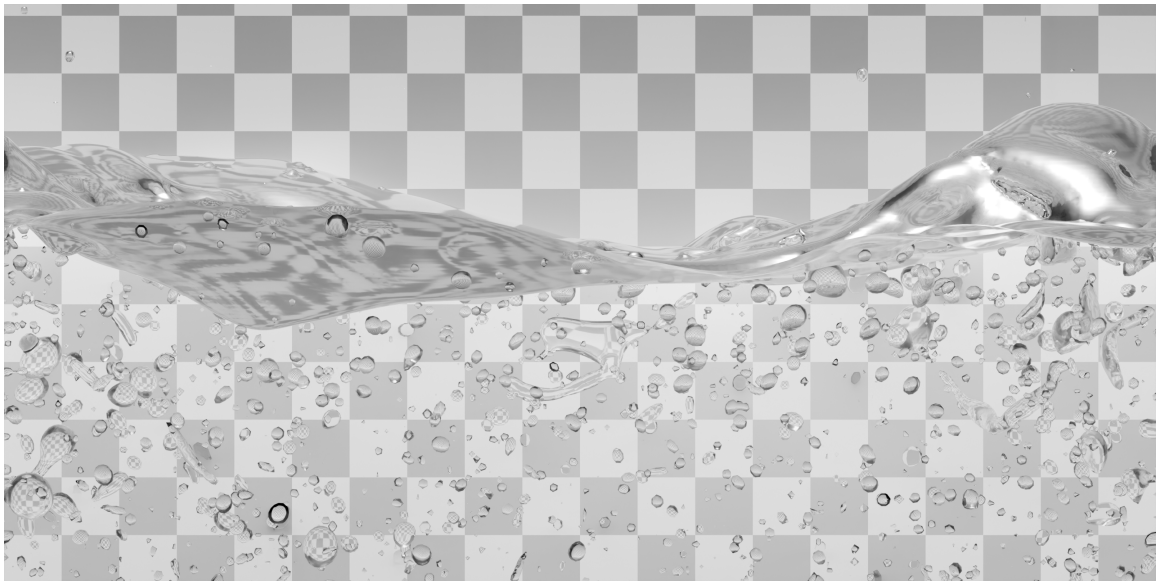


Figure 4-2: Rendering of a forced FST simulation ($Fr^2 = 1.2$ and $We = 200$) showing the free surface and bubbly flow beneath. Scaling to match Earth gravity and air-water surface tension and density, each square on the background grid is 1.09cm across.

kinetic energy

$$\overline{\rho}\tilde{k} \equiv \tfrac{1}{2}\overline{\rho \mathbf{u}\cdot\mathbf{u}}\,, \tag{4.10}$$

which gives a characteristic velocity $u_{\text{rms}} = \sqrt{2\tilde{k}/3}$, and the dissipation rate

$$\overline{\rho}\varepsilon \equiv \overline{\tau_{ij}\,\partial u_i/\partial x_j}\,. \tag{4.11}$$

Both are shown in table 4-1. These quantities define near-surface nondimensional numbers, most relevant here is the turbulent Froude number (squared) $Fr_T^2 = \varepsilon/u_{\text{rms}}\text{g}$. In §4.5 we will show that the turbulence levels at $z^* = -0.5$ best characterize the turbulence within the surface layer.

4.4 Intermittency in the surface layer

From the DNS results, we start by analyzing the intermittency factor $\gamma(z)$. Figure 4-3 shows $\gamma(z)$ for the range of Fr_T^2, before and after scaling relative depth $z - \bar{\eta}$ by surface layer thickness δ_s. Given δ_s is defined by the derivative of γ at $z = \bar{\eta}$, the distributions necessarily collapse at the mean free surface. However, we see that scaling by δ_s also does a good job collapsing the distribution of $\gamma(z)$ throughout the surface layer. For $z^* < -0.5$ we do see some small Fr_T^2-dependent differences, and in section 4.4.1 we will confirm that this is the result of Fr_T^2-dependent bubble entrainment rates, rather than any difference in the nature of the surface layer. In section 4.4.2 we will investigate the shape of $\gamma(z)$ and show that it is closer to a logistic distribution rather than the Gaussian distribution discussed by Brocchini & Peregrine (2001*b*).

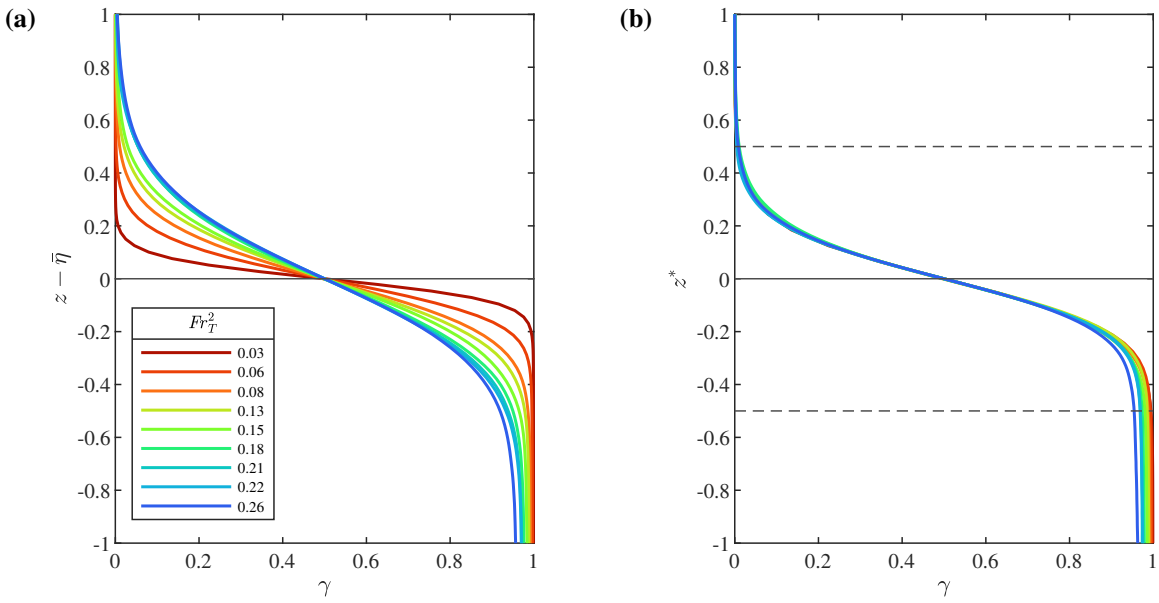


Figure 4-3: Intermittency factor γ as a function of (a) depth relative to the mean free surface and (b) depth scaled by surface layer thickness. (——) indicated the mean free surface $\bar{\eta}$ and in (b) (- - - -) shows the extent of the surface layer.

4.4.1 Separating out the effects of bubbles and droplets

The intermittency factor γ includes the effects of three different types of features: the free surface, bubbles, and droplets. To separate these, we start by applying ICL (Hendrickson *et al.*, 2020, see also §2.3) to identify contiguous regions of air that are disconnected from the bulk region of air above the free surface (bubbles) and contiguous regions of water that are disconnected from the bulk region of water below the free surface (droplets). From this post-processing, we can define two new binary color functions,

$$c_D(\mathbf{x},t) \equiv \begin{cases} 1 & \text{if} \quad \mathbf{x} \in \text{droplet of water} \\ 0 & \text{otherwise} \end{cases}, \tag{4.12a}$$

$$c_B(\mathbf{x},t) \equiv \begin{cases} -1 & \text{if} \quad \mathbf{x} \in \text{bubble of air} \\ 0 & \text{otherwise} \end{cases}. \tag{4.12b}$$

Subtracting these from the original color function, we obtain a third binary color function

$$c_0 \equiv c - (c_D + c_B)\,, \tag{4.12c}$$

which describes the mixture of air and water if all bubbles were filled with water and all droplets replaced by air. This c_0 isolates free-surface effects. Figure 4-4 shows a representative example of splinting the original color function c into these three color functions.

Qualitatively, we note some key differences between figure 4-4 from DNS and figure 4-1, the sketch by Brocchini & Peregrine (2001*b*). Despite the high Fr_T^2, DNS shows an intact free

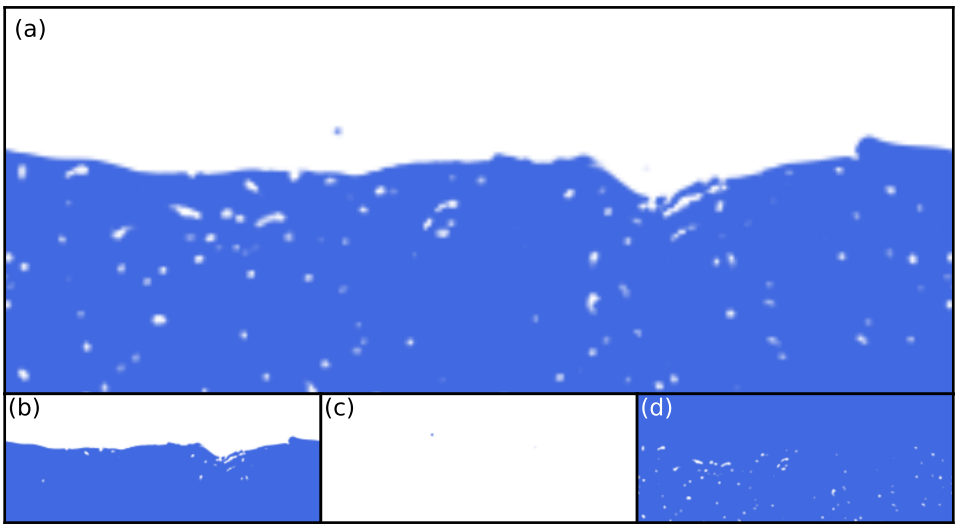


Figure 4-4: Two-dimensional slice over $z^* \in [-1, +1]$ from the $Fr_T^2 = 0.13$ simulation showing: (a) the original color function c; (b) the free-surface color function c_0; (c) the droplet color function c_D; and (d) the bubble color function $1 + c_B$. Color scale from 0 (white) to 1 (blue).

surface is still easily identifiable. The easily identifiable free surface is also apparent in the three-dimensional rendering in figure 4-2. This observation is consistent with experimental work on open-channel flow. While the surface layer appears to the human eye like a uniform mixture of air and water, high-speed photography shows the free surface, while highly distorted, is still mostly intact (Killen, 1968; Wilhelms & Gulliver, 2005).

To quantify the air-water mixture in the surface layer, we apply averaging like in (4.3) to each of the three new color functions to obtain γ_0, γ_B, and γ_D, shown in in figure 4-5. These split the intermittency factor into the contributions from each of the three phenomenon,

$$\gamma = \gamma_0 + \gamma_B + \gamma_D \, . \tag{4.13}$$

By rough order-of-magnitude, we see γ_0 is 100 times larger than γ_B, which is 100 times larger than γ_D. For droplets, we see a peak in γ_D within the upper surface layer ($z^* \approx 0.2$) and then a rapid decay with increasing height. The behavior for bubbles is very different: $-\gamma_B$ increases with increasing depth throughout the surface layer, and then only very slowly decays with further increasing depth. Because $\rho_a/\rho_w \ll 1$, the mass of a bubble is negligible compared to the water surrounding it, allowing the turbulence in the water to advect the bubble deep beneath the surface. As expected, $-\gamma_B$ increases with Fr_T^2, as bubble entrainment increases as Fr_T^2 increases (see Chapter 6).

Focusing on the free surface effects, we see in figure 4-5 that the distribution of γ_0 collapses very well when scaling by z^*. Because γ_D is so small, droplet effects are negligible. The small Fr_T^2-dependence of γ for $z^* < 0$ in figure 4-3b is confirmed to be because of bubbles. For significantly larger bubble void fractions (i.e., larger Fr_T^2) one may need to calculate δ_s using γ_0 rather than γ in using (4.6); however, for these simulations the difference is negligible.

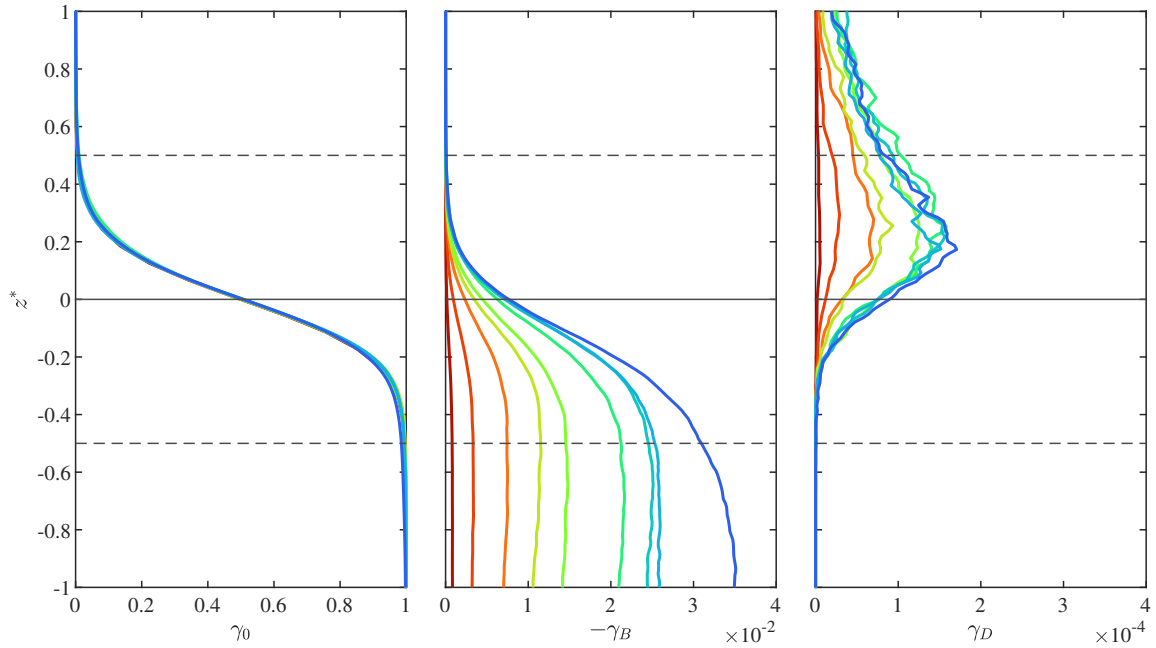


Figure 4-5: The three sub-components of intermittency factor γ across a range of Fr_T^2 (see figure 4-3a for color legend). γ_0 describes the free surface, γ_B bubbles, and γ_D droplets. Note the difference in magnitude of the horizontal axes.

4.4.2 Describing the distribution of intermittency

We now consider the distribution of the intermittency factor. To highlight the tail-behavior of the distribution, figure 4-6 shows the derivative of intermittency with z^*. In general, we see tail behavior like

$$-\mathrm{d}\gamma/\mathrm{d}z \propto \exp\left[|z|\right] , \tag{4.14}$$

especially for the upper tail ($z^* > 0$). This is opposed to $-\mathrm{d}\gamma/\mathrm{d}z \propto \exp[z^2]$ predicted by (4.5). Instead of Gaussian, this tail behavior is characteristic of a logistic distribution,

$$\gamma_{\text{Logistic}}(z) = \frac{1}{2}\left[1 - \tanh\left(\frac{z-\bar{\eta}}{s}\right)\right] . \tag{4.15}$$

The behavior of γ_0 (figure 4-6b) is generally the same as γ (figure 4-6a), demonstrating that this is logistic behavior is the result of free-surface phenomenon rather than droplets. We can compare this result to experiments by Ruth & Coletti (2024), who studied a similar horizontally and temporally steady FST forced from below, but at $Fr_T^2 \ll 1$ where the free surface is not broken. For such a single-valued surface, $-\mathrm{d}\gamma/\mathrm{d}z$ is equivalent to the probability distribution function of free surface deformations. Their results also appear to follow $\propto \exp[|z|]$ (Ruth & Coletti, 2024, Figure. 5a). An interpretation of the logistic distribution of γ comes from Wacławczyk (2021), who noted analogies between the effects of macroscopic turbulence and mesoscopic thermal fluctuations on intermittency at the respective scales.

While the logistic distribution does a much better job than the Gaussian distribution

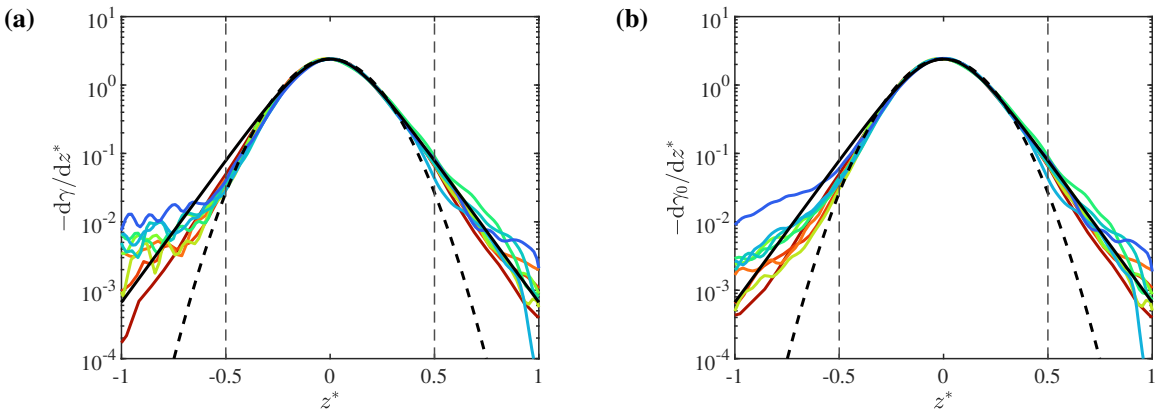


Figure 4-6: Distribution of (a) total intermittency and (b) intermittency excluding droplets and bubbles across a range of Fr_T^2 (see figure 4-3a for color legend). Distributions are compared to (– – –) γ_{Gaussian} from (4.5) and (——) γ_{Logistic} from (4.15). Vertical (- - - -) show the extent of the surface layer.

at describing intermittency γ, we do note that for the lower tail ($z^* < 0$) the distribution of γ is more complicated. Beneath the surface layer ($z^* < -0.5$) there is a larger tail than even predicted by the logistic distribution (4.15). While removing the effect of bubbles with γ_0 reduces the tail, there is still a larger tail. This can be understood by examining c_0 in figure 4-4b. Even after removing bubbles, there are still pockets of air which, in the two-dimensional slice of the domain, appear separate from the bulk air above the free surface. In three dimensions it is clear that these are long filaments of air that can reach deep beneath the free surface while still being connected to the bulk air above. These filaments are presumably the result of bubbles that recently reconnected to the surface (degassing) or are soon to disconnect from the surface to form a bubble (entrainment), making it a bit unclear if filaments should be treated like free-surface phenomenon (as done here) or bubble phenomenon.

4.5 Turbulence in the surface layer

We now examine the turbulence in and around the surface layer. First, section 4.5.1 examines the isotropy, showing that the transition to strong FST, where near-surface turbulence is nearly isotropic (Yu *et al.*, 2019), happens at the critical turbulent Froude number $Fr_T^2 = 0.1$. In section 4.5.2 we examine the terms in the averaged vertical momentum ($\overline{\rho w}$) conservation equation, and in section 4.5.3 we examine the terms in the averaged kinetic energy ($\overline{\rho}\tilde{k}$) conservation equation. Throughout this analysis we will show that relevant properties collapse across a wide range of Fr_T^2 when properly scaled by z^*, and that, partially for strong FST ($Fr_T^2 > 0.1$), the free surface is not felt by the turbulence outside of the surface layer, with $z^* = -0.5$ being the transition point. In section 4.6 we discuss how these insights can be applied to RANS turbulence modeling.

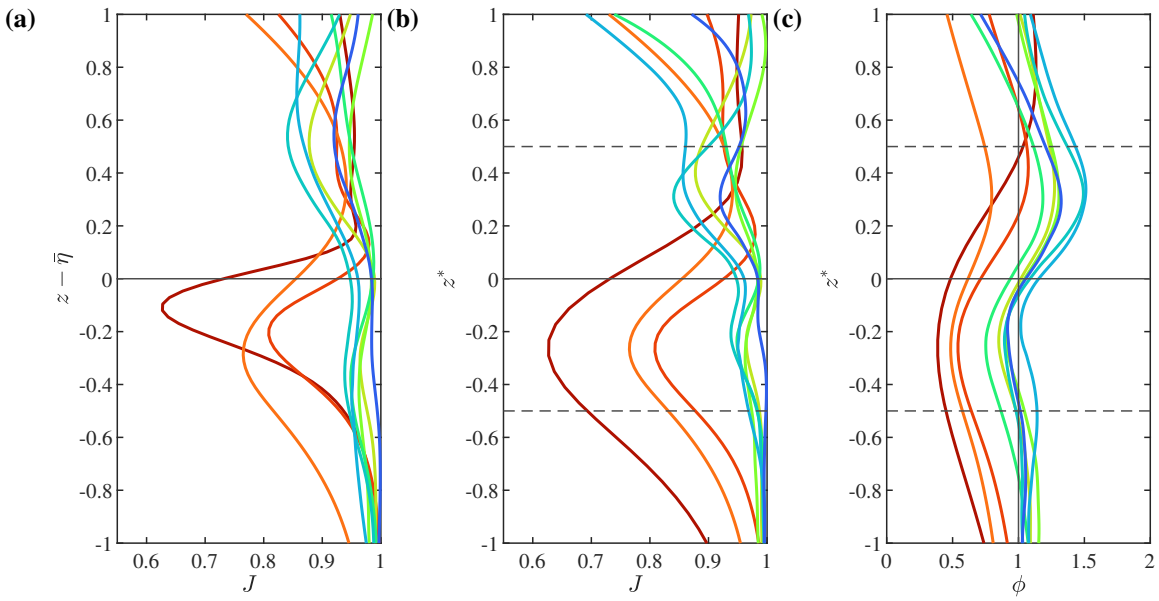


Figure 4-7: (a, b) Isotropy metric J and (c) relative contributions of vertical fluctuations ϕ across a range of Fr_T^2 (see figure 4-3a for color legend).

4.5.1 Isotropy

For weak FST ($Fr_T^2 \ll 1$), the restoring force of gravity prevents large deformations of the free surface, suppressing vertical fluctuations. This blockage effect creates highly anisotropic turbulence (Shen *et al.*, 1999; Guo & Shen, 2010; Ruth & Coletti, 2024). However, as Fr_T^2 increases the strength of the turbulence increases relative to gravity. Yu *et al.* (2019) show that for sufficiently large Fr_T^2 the blockage effectively disappears and the near-surface turbulence is nearly isotropic. Calado & Balaras (2025) show that for more moderate Fr_T^2 vertical fluctuations are partially suppressed. It is now known what Fr_T^2 is the transition between moderate FST where vertical fluctuations are partially suppressed and strong FST where the turbulence is nearly isotropic.

A formal way to quantify isotropy is the isotropy metric

$$J \equiv 1 - 9\left(\tfrac{1}{2} b_{ij} b_{ij} - b_{ij} b_{jk} b_{ki}\right), \tag{4.16}$$

where $b_{ij} \equiv \overline{u_i u_j}/\overline{u_k u_k} - \delta_{ij}/3$ is the anisotropy tensor. $J = 1$ corresponds to perfectly isotropic turbulence, and $J = 0$ corresponds to either one- or two-component turbulence (Pope, 2000). Figure 4-7a and 4-7b show J as a function of depth, unscaled and scaled by δ_s respectively. First, we observe that scaling by δ_s does a very good job collapsing the behavior of J with depth. Across all Fr_T^2, there is a local minimum in the isotropy metric at $z^* \approx -0.25$ suggesting this is where the blockage effect is strongest.

Transition from moderate to strong FST

Figure 4-8 shows the values of J at the local minimum. This result shows a clear transition from moderate FST to strong FST at $Fr_T^2 \approx 0.1$. For $Fr_T^2 > 0.1$ we observe nearly isotropic

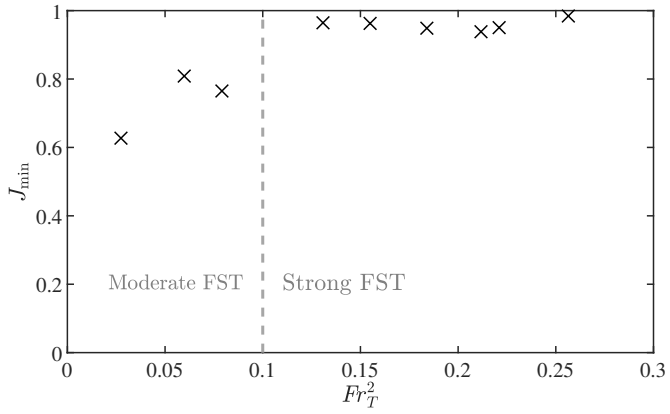
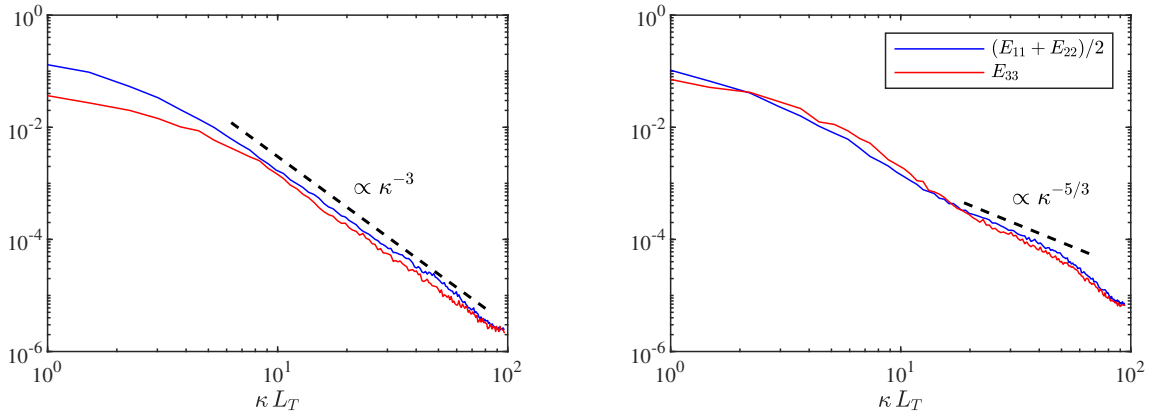


Figure 4-8: Minimum value of the isotropy metric J in the lower surface layer ($z^* \in [-0.5, 0]$) as a function of Fr_T^2, with the approximate transition from moderate to strong FST indicated by the vertical dashed line at $Fr_T^2 = 0.1$.



(a) Moderate FST, $Fr_T^2 = 0.03$ **(b)** Strong FST, $Fr_T^2 = 0.13$

Figure 4-9: Comparison of one-dimensional turbulent energy spectra at $z = 0$ between moderate and strong FST. $E_{11} \approx E_{22}$, so for clarity the horizontal components are averaged.

turbulence ($J \approx 0.95$), as expected for strong FST. For $Fr_T^2 < 0.1$ we see a moderate FST regime where turbulence becomes more anisotropic as Fr_T^2 decreases, although J is still far from $J \approx 0$ characteristic of weak FST.

In addition to measuring isotropy, we examine the turbulence energy spectrum at $z = 0$. From the momentum $\mathbf{p} \equiv \rho\mathbf{u}$ measured at $z = 0$, we apply Fourier analysis and average over $t \in [t_0, t_0 + T_{\text{sim}}]$ to obtain the one-dimensional (in the x-direction) energy spectrum E_{ii}, normalized such that $\int (E_{11} + E_{22} + E_{33})\, \mathrm{d}\kappa = 1$. Figure 4-9 compares the spectra of moderate and strong FST. For moderate FST, we see that the horizontal fluctuations are stronger than the vertical fluctuations at all scales and that $E \propto \kappa^{-3}$, characteristic of two-dimensional turbulence (Kraichnan, 1967). For strong FST, we see that the turbulence is isotropic at all scales ($E_{11} \approx E_{22} \approx E_{33}$ for all κ) and that there is a range of κ where $E \propto \kappa^{-5/3}$, consistent with the Kolmogorov inertial sub range of isotropic turbulence.

Behavior of isotropy in and around the surface layer

Having established that $Fr_T^2 = 0.1$ separates moderate from strong FST, for each we now examine the behavior of isotropy in and around the surface layer. Returning to figure 4-7b, the first and most important observation is that for strong FST there is no effect of the free surface on isotropy beneath the surface layer (i.e., $J \approx 1$ for $z^* < -0.5$). This is the first indication that, for strong FST, turbulence beneath the surface layer does not directly feel the presence of the free surface. This observation will form the basis of the surface layer model we develop in section 4.6.

We now investigate isotropy in the surface layer. The cause of anisotropy in FST is the effect of gravity on vertical fluctuations, so it is also useful to consider the relative magnitude of vertical fluctuations,

$$\phi \equiv 3\overline{ww}/\overline{u_k u_k}\,, \tag{4.17}$$

shown in figure 4-7c. For moderate FST, we see $\phi < 1$ below and throughout the surface layer, indicating that vertical fluctuations are suppressed. Focusing on strong FST, we see that in the lower surface layer ($z^* \in [-0.5, 0]$) vertical fluctuations are suppressed, but in the upper surface layer ($z^* \in [0, 0.5]$) $\phi > 1$, indicating that vertical fluctuations are in fact amplified. This helps explain the shape of J in strong FST. By symmetry, our FST flow will satisfy $\overline{uu} \approx \overline{vv}$ and $\overline{uv} \approx 0$, from which one can show

$$J \approx \tfrac{1}{4}\,\phi\,(3-\phi)^2 \tag{4.18}$$

For ϕ to go from < 1 in the lower surface layer to > 1 in the upper surface layer, it must pass through $\phi = 1$ around the mean free surface. This explains why near perfect isotropy ($J \approx 1$) is observed at the mean free surface ($z^* = 0$) for strong FST (see figure 4-7b).

4.5.2 Momentum

We now examine the vertical momentum equation. We define density fluctuations $\rho' \equiv \rho - \overline{\rho}$, in which case the vertical component of (2.1), neglecting surface tension, can be written in conservative form as

$$\frac{\partial \rho w}{\partial t} = -\boldsymbol{\nabla}\cdot(\mathbf{u}\rho w) - \partial p_d/\partial z + \partial \tau_{i3}/\partial x_i + \rho' \mathrm{g}\,. \tag{4.19}$$

For pressure we consider dynamic pressure $p_d = p - p_h$, where the true hydrostatic pressure[1] is defined

$$p_h(z) \equiv p_0 + \mathrm{g}\int_z^{Z_{max}} \overline{\rho}(z')\,\mathrm{d}z'\,, \tag{4.20}$$

with p_0 as some reference pressure. Averaging (4.19) using (4.2) and splitting the Reynolds stress into two terms, $\overline{\rho w w} = \overline{\rho}\,\overline{ww} + \overline{\rho' ww}$, we are left with four momentum flux terms,

$$0 = -\frac{\partial}{\partial z}\left(\overline{\rho}\,\overline{ww} + \overline{\rho' ww} + \overline{p_d} - \overline{\tau_{33}}\right)\,. \tag{4.21}$$

[1]This true hydrostatic pressure is time-independent, unlike the pseudo hydrostatic pressure (2.7) used internally by the DNS solver.

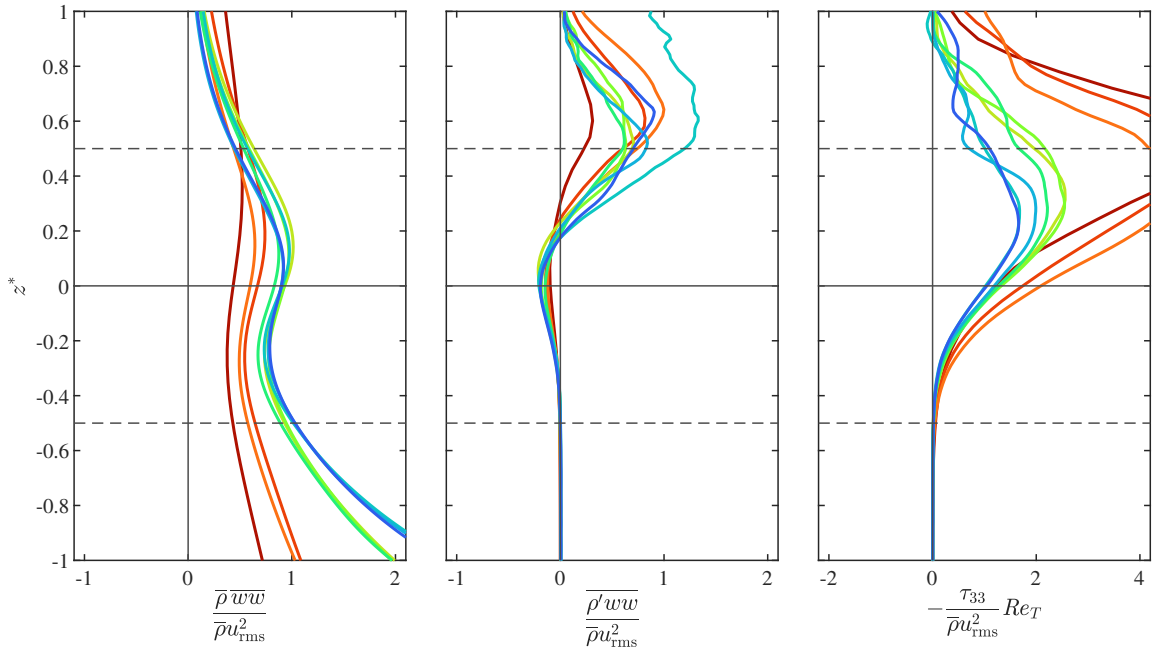


Figure 4-10: Momentum flux normalized by mean density $\overline{\rho}$ measured for each z^* and characteristic velocity u_{rms} measured at $z^* = -0.5$ (see Table 4-1) across a range of Fr_T^2 (see figure 4-3a for color legend). For the viscous term, we further multiply by Re_T to bring it to a similar scale.

The appropriate choice for p_0 makes the sum of momentum flux terms (i.e., stresses) zero:

$$\overline{\rho}\,\overline{ww} + \overline{\rho' ww} + \overline{p_d} - \overline{\tau_{33}} = 0 \tag{4.22}$$

In figure 4-10 we plot (per unit mass) the mean-density component of Reynolds stress $\overline{ww}$, the fluctuating-density component of Reynolds stress $\overline{\rho' ww}/\overline{\rho}$, and the viscous stress $-\overline{\tau_{33}}/\overline{\rho}$. As shown in (4.22), pressure stress $\overline{p_d}/\overline{\rho}$ cancels the sum of these other three stresses. We see that plotting against z^* and normalizing the stresses by u_{rms}^2 measured at the bottom of the surface layer ($z^* = -0.5$) does a good job collapsing the results, especially for strong FST. Beneath the surface layer ($z^* < -0.5$), there is no contribution of fluctuating-density Reynolds stress or viscous stress. Within the surface layer ($z^* \in [-0.5, 0.5]$), viscous stress is of order Re_T^{-1}, meaning it is negligible in this DNS and certainty negligible at the larger Re_T typical of real-world FST. For strong FST, we see that $\overline{ww}$ slightly decreases in the lower surface layer as vertical fluctuations are suppressed, then slightly increases in the upper surface layer as vertical fluctuations are amplified, consistent with ϕ in figure 4-7c.

We now focus on the fluctuating-density component of Reynolds stress, $\overline{\rho' ww}$. Within the lower surface layer this term becomes slightly negative, then increases through the upper surface layer. As with intermittency in section 4.4.1, we can split this term into separate contributions from bubbles, droplets, and the free surface. Recalling (2.4), we can write density fluctuations as

$$\rho' = \Delta\rho(c - \gamma)\,, \tag{4.23}$$

where $\Delta\rho = \rho_w - \rho_a$ is the difference in density between water and air. As in section 4.4.1, we can split the color function into the contributions from each of the three phenomenon

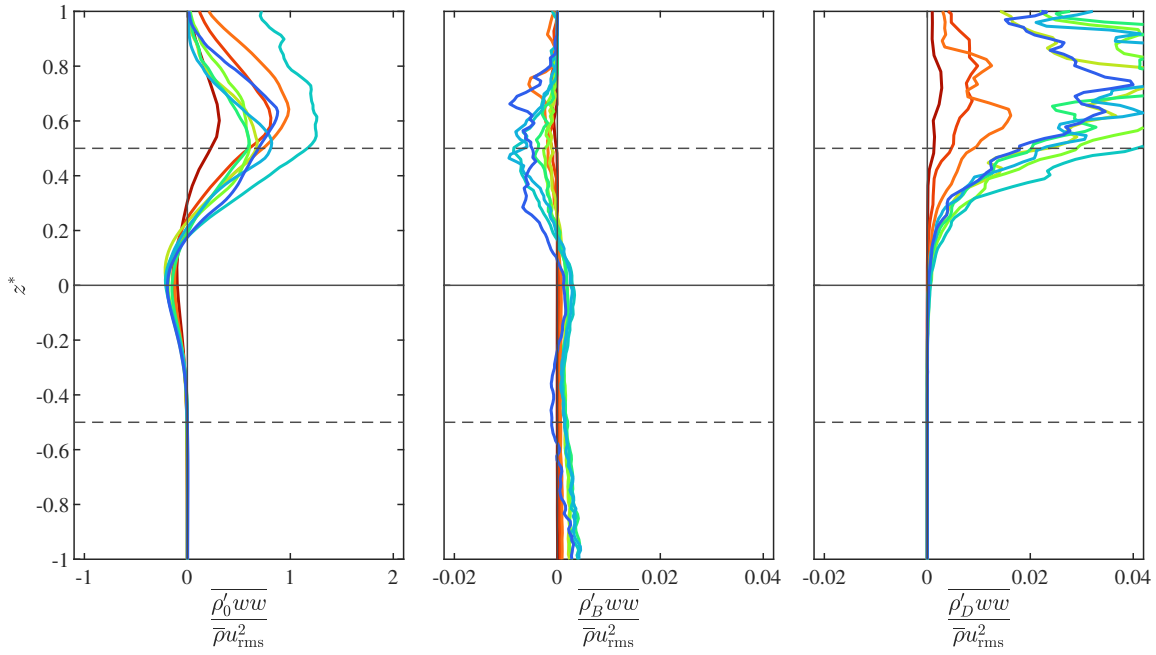


Figure 4-11: Components of $\overline{\rho' ww}$ from (4.25), normalized by mean density $\overline{\rho}(z^*)$ and characteristic velocity u_{rms} measured at $z^* = -0.5$ (see Table 4-1) across a range of Fr_T^2 (see figure 4-3a for color legend). Note the difference in magnitude of the horizontal axes.

($c = c_0 + c_D + c_B$) and the same for the intermittency factor ($\gamma = \gamma_0 + \gamma_D + \gamma_B$). Through (4.23), this allows us to split the density fluctuations ($\rho' = \rho_0' + \rho_D' + \rho_B'$). The density fluctuations caused by the free surface are

$$\rho_0' \equiv \Delta\rho(c_0 - \gamma_0)\,, \tag{4.24}$$

and similar equations for ρ_D' and ρ_B' describing fluctuations caused by droplets and bubbles respectively. This allows us to decompose $\overline{\rho' ww}$ into three terms,

$$\overline{\rho' ww} = \overline{\rho_0' ww} + \overline{\rho_B' ww} + \overline{\rho_D' ww}\,, \tag{4.25}$$

each shown in figure 4-11. We see that, by at least two orders of magnitude, free-surface effects ($\overline{\rho_0' ww}$) dominate bubble or droplet effects. As expected, droplets only have a (very small) effect above the mean free surface ($z^* > 0$). Bubbles on the other hand never have a significant effect through $\overline{\rho_B' ww}$, even for large Fr_T^2 where the volume fraction of bubbles is non-negligible. To explain this result, we note that $\overline{\rho_B' ww}$ can be interpreted as the correlation between the presence of a bubble and large vertical velocities. Appendix D contains an analysis of all the $\overline{\rho' ww}$ terms interpreted as correlation coefficients. In summary, zero correlation is what one would obtain if bubbles were treated as passive particles, showing that bubbles in FST are primarily advected by the turbulence, with buoyancy having a smaller contribution.

4.5.3 Turbulent kinetic energy

We now examine the turbulent kinetic energy (TKE) equation. Taking the inner product of $\mathbf{u}$ and (2.1) written in conservative form (again, neglecting surface tension),

$$\frac{\partial}{\partial t}\left(\tfrac{1}{2}\rho\mathbf{u}\cdot\mathbf{u}\right) = -\boldsymbol{\nabla}\cdot\left[\mathbf{u}\left(\tfrac{1}{2}\rho\mathbf{u}\cdot\mathbf{u}\right) + \mathbf{u}p - \mathbf{u}\cdot\boldsymbol{\tau}\right] - \boldsymbol{\tau} : \boldsymbol{\nabla}\mathbf{u} - \mathrm{g}\rho w\,. \tag{4.26}$$

It is also useful to consider the mass conservation equation,

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\mathbf{u}) = 0\,. \tag{4.27}$$

Averaging (4.27) and noting that there is no flux at the bottom or top boundaries of the domain, we obtain $\overline{\rho w} = 0$. With this the gravity term cancels out when we average (4.26), and we obtain

$$0 = -\frac{\partial}{\partial z}\left(\tfrac{1}{2}\overline{\rho w u_i u_i} + \overline{wp} - \overline{u_i \tau_{i3}}\right) - \overline{\rho}\varepsilon\,. \tag{4.28}$$

In this section, we will: compare the evolution of $\tilde{k}$ and ε between the simulations with the free surface and single phase simulation to show, especially for strong FST, that free-surface effects are restricted to within the surface layer; examine $\tilde{k}$ and ε within the surface layer; and evaluate the three energy flux terms in (4.28). Predicting the flux of energy into the surface layer (or equivalently, the energy dissipation within the surface layer) will form an important part of the turbulence model discussed in section 4.6.

Extent of the free surface's effect on turbulence

Recall our simulation setup generates turbulence (i.e., injects turbulent kinetic energy $\tilde{k}$) in a region deep beneath the free surface, which then diffuses toward the surface while some is dissipated by ε. Because of this setup, turbulence levels vary with depth, even in the absence of free-surface effects (Guo & Shen, 2009). To elucidate the effect of the free surface on $\tilde{k}$ and ε, we first perform a new DNS without any free surface. This no free surface simulation has the same setup as those described in §2.1, except the entire domain is filled with water. We find the air gap of π is sufficient such that the upper boundary has negligible effect on the turbulence around $z = 0$ (the location of the quiescent free surface in the other simulations).

Figure 4-12 compares $\tilde{k}$ and ε from the simulations with and this new the simulation without a free surface. Turbulence values are normalized by their value at $z = -\pi/2$, the point above which there is no forcing and the flow evolves naturally (see (4.8)). Focusing first on figure 4-12a and 4-12c, in the absence of the free surface we see the expected exponential decay with depth for both $\tilde{k}$ and ε (Guo & Shen, 2009). This is also true for simulations with the free surface below a certain depth; however, at shallower depths both $\tilde{k}$ and ε depart and are much larger than in the simulation without a free surface.

To quantify the amplification of $\tilde{k}$ and ε due to free-surface effects, in figure 4-12b and 4-12d we divide the value of $\tilde{k}$ and ε in the simulations with a free surface to the value at the same z in the simulation without a free surface. We then normalize depth by δ_s to plot in terms of z^*. We see that, especially for strong FST, the amplification of $\tilde{k}$ and ε only happens above the lower limit of the surface layer ($z^* > -0.5$). These results demonstrate

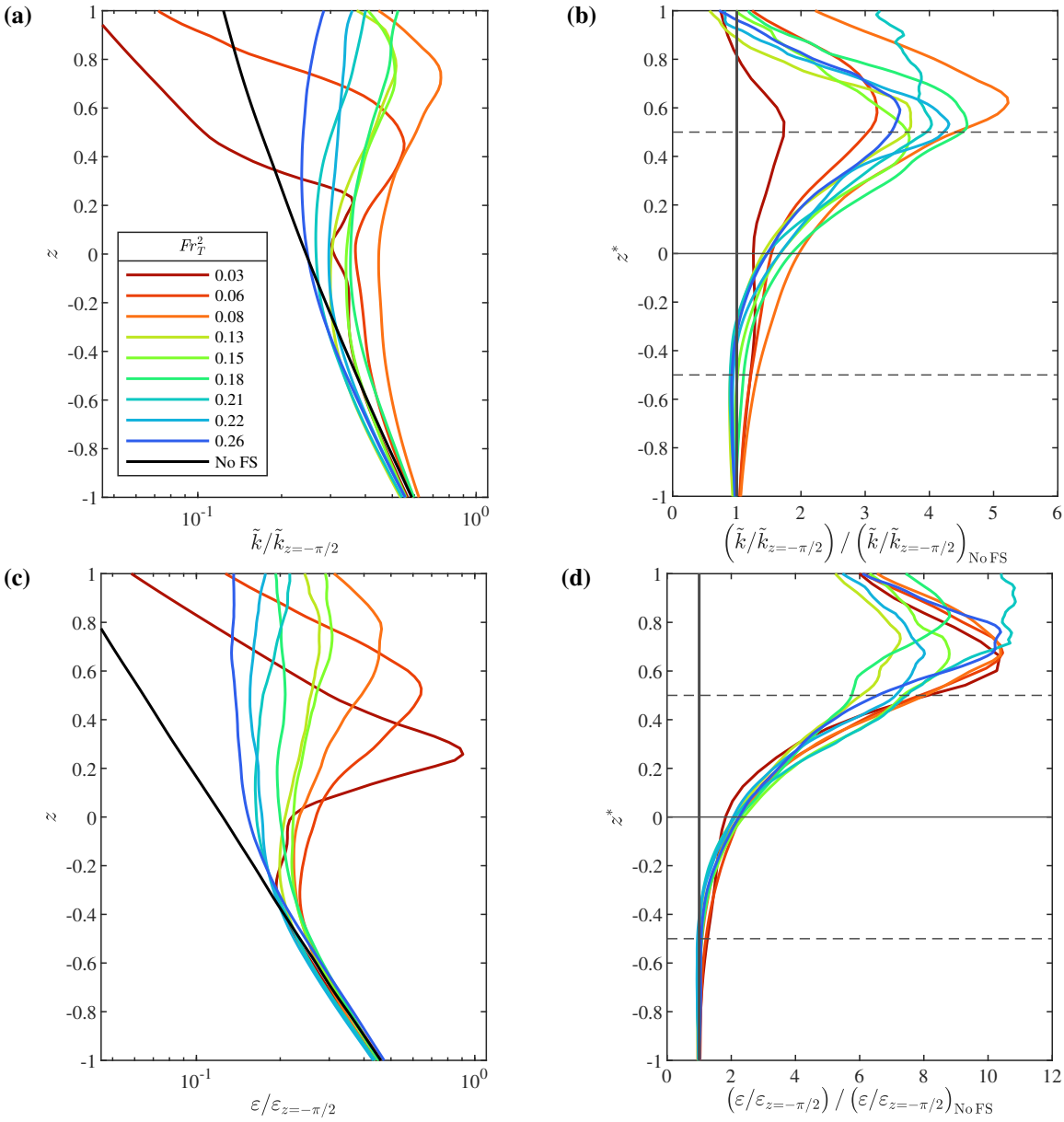


Figure 4-12: (a & b) Turbulent kinetic energy $\tilde{k}$ and (c & d) dissipation rate ε across a range of Fr_T^2 (see figure 4-3a for color legend), compared to a simulation without a free surface (——).

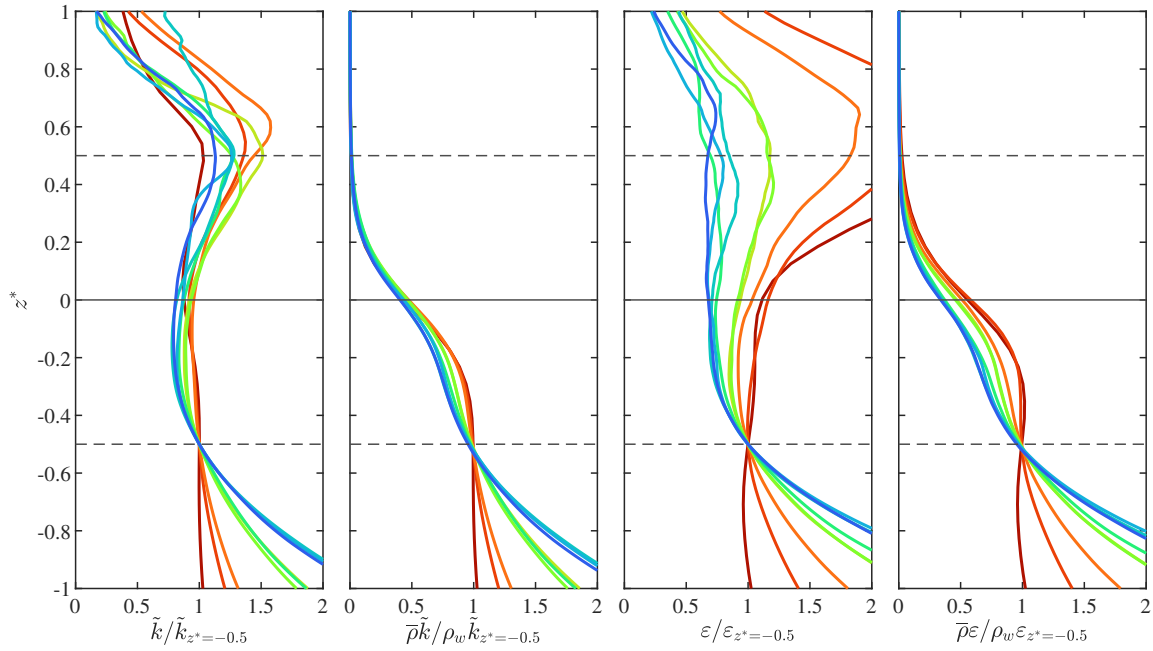


Figure 4-13: Turbulent kinetic energy $\tilde{k}$ and dissipation rate ε, per unit volume and per unit mass across a range of Fr_T^2 (see figure 4-3a for color legend).

that the free surface does not directly affect the turbulence beneath the surface layer, and that $z^* = -0.5$ is an appropriate definition of the lower limit of the surface layer. Notably, this suggests that beneath the surface layer standard turbulence closure models are applicable.

Turbulence within the surface layer

Unlike beneath the surface layer, the turbulence within the surface layer is directly affected by free surface effects. Figure 4-12b and 4-12d show that $\tilde{k}$ and ε are significantly increased relative to flow without a free surface. Figure 4-13 shows $\tilde{k}$ and ε normalized by the value measured at the bottom of the surface layer ($z^* = -0.5$). For the turbulent kinetic energy per unit mass $\tilde{k}$, we see that there is relatively little variation within the surface layer, especially for the lower surface layer ($z^* \in [-0.5, 0]$). In other words, throughout the surface layer, $\tilde{k}$ does not depart significantly from the value at the bottom of the surface layer ($z^* = -0.5$). Furthermore, because $\rho_a \ll \rho_w$, the small variation in the upper surface layer ($z^* \in [0, 0.5]$) is negligible when one considers the turbulent kinetic energy per unit volume, $\overline{\rho}\tilde{k}$.

For the turbulent dissipation rate per unit mass ε, the variations in the surface layer are slightly larger than for $\tilde{k}$, but still ε measured at the bottom of the surface layer does a good job characterizing dissipation rates within the surface layer, especially the lower surface layer ($z^* \in [-0.5, 0]$). Because $\rho_a \ll \rho_w$, the majority of the absolute dissipation of energy happens in the lower surface layer, and we see a reasonably strong collapse of $\overline{\rho}\varepsilon$.

As a first approximation, one could model the dissipation rate per unit mass in the surface layer as a constant equal to ε measured at $z^* = -0.5$, in which case the rate at which energy

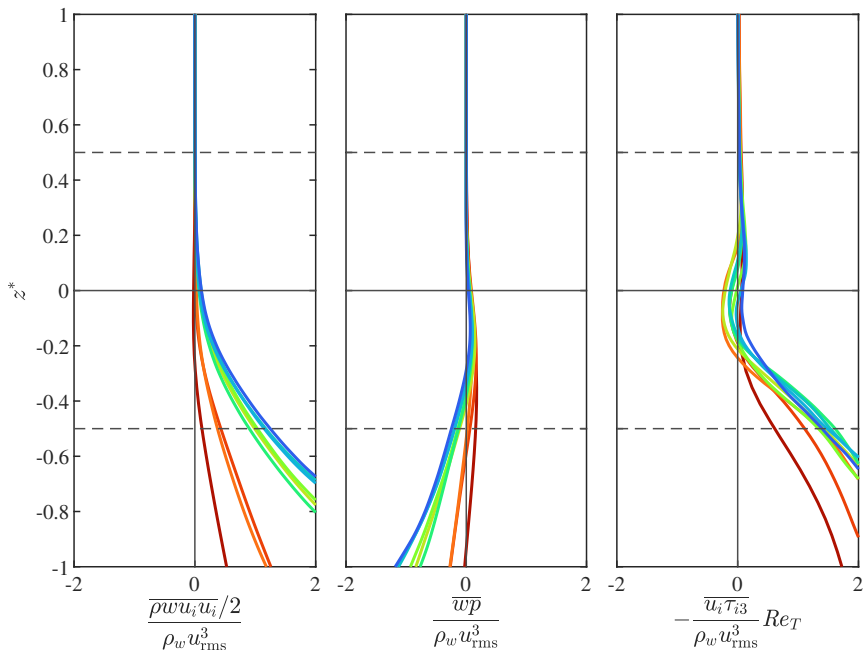


Figure 4-14: Turbulent kinetic energy flux terms across a range of Fr_T^2 (see figure 4-3a for color legend), normalized by water density ρ_w and characteristic velocity u_{rms} measured at $z^* = -0.5$ (see Table 4-1). For the viscous term, we further multiply by Re_T to bring it to a similar scale.

is dissipated within the surface layer (per unit mean free-surface area) could be approximated

$$W \approx \varepsilon\, \delta_s \left[\rho_w - (\rho_w - \rho_a) \int_{-0.5}^{0.5} \gamma \, \mathrm{d}z^* \right] . \tag{4.29}$$

Using a distribution of γ that is symmetric about $\bar{\eta}$ (this includes both (4.5) and (4.15)) and noting $\rho_a \ll \rho_w$, one obtains

$$W/\rho_w \approx \tfrac{1}{2}\, \varepsilon\, \delta_s \, . \tag{4.30}$$

This first approximation turns out to be very close to the fit we obtain later in section 4.6 based on measuring the flux of energy into the surface layer.

Energy flux into the surface layer

We now investigate the three energy flux terms in (4.28), shown in figure 4-14. As with the momentum flux terms, we plot against z^* and normalize the stresses by u_{rms} measured at the bottom of the surface layer ($z^* = -0.5$). First, we note that, as expected, the viscous diffusion term $\overline{u_i \tau_{i3}}$ is of order Re_T^{-1}, meaning it is negligible in this DNS and certainty negligible at the larger Re_T typical of real-world FST. For the remaining terms, we see turbulent diffusion $\overline{\rho w u_i u_i}$ moves energy upward into the surface layer, and pressure diffusion $\overline{wp}$ partially counteracts this, moving some energy out of the surface layer. Of note, despite scaling by z^* and u_{rms}, we still see Fr_T^2 dependence in these energy flux terms.

While we have shown that the turbulence beneath the surface layer does not directly feel the free surface, the increased energy dissipation with the surface layer is still important to

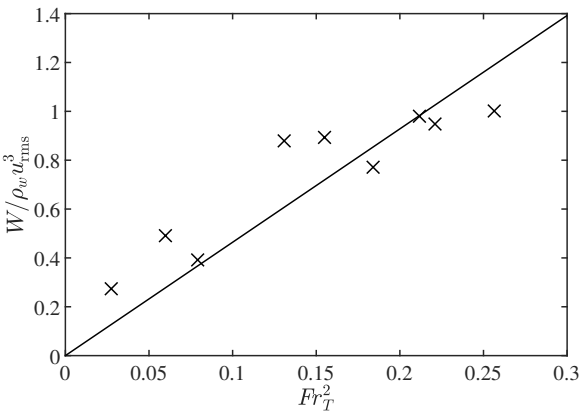


Figure 4-15: Work done on the surface layer by turbulence beneath W as a function of Fr_T^2, including (——) the linear fit to (4.32) ($R^2 = 0.641$).

the overall energy budget of the flow. Thus, the energy flux into the surface layer (i.e., the net work done on the surface layer by the turbulence beneath) is an important quantity to model. Ignoring the negligible viscous term, this net work is given by the sum of turbulent and pressure diffusion measured at the bottom of the surface layer,

$$W = \left[\tfrac{1}{2}\overline{\rho w u_i u_i} + \overline{wp}\right]_{z^*=-0.5} . \tag{4.31}$$

In figure 4-15 we show that (nondimensionalized by ρ_w and $u_{\rm rms}$ measured at $z^* = -0.5$), this net work scales like Fr_T^2. Linear regression gives

$$W/(\rho_w u_{\rm rms}^3) = C_W\, Fr_T^2 , \tag{4.32}$$

where the 95% confidence interval is $C_W = 4.6 \pm 0.8$. Dimensionalizing (4.32), we see how the work done on the surface layer depends on gravity and the turbulence (ε and $u_{\rm rms}$) measured at the bottom of the surface layer,

$$W/\rho_w = C_W\, u_{\rm rms}^2 \varepsilon\, \mathrm{g}^{-1} . \tag{4.33}$$

4.6 A surface layer model for RANS

We have shown that, particularly for strong FST, there is a clear difference between the turbulence beneath the surface layer and the turbulence within the surface layer, with $z^* = -0.5$ separating the two. Because the turbulence beneath the surface layer is not directly affected by the free surface, standard turbulence closure models, such as RANS, will likely still be appropriate. What is needed then is a model for the surface layer that acts like a boundary condition for RANS, applied at $z = \bar{\eta} - 0.5\delta_s$ from the mean free surface, analogous to wall models for boundary layers.

So far, we have obtained the surface layer thickness δ_s *a posteriori* using measurements of γ, as described by (4.6). For a predictive model appropriate to RANS, we need δ_s based on the turbulence levels ε and $u_{\rm rms} = \sqrt{2\tilde{k}/3}$. Figure 4-16 shows the scaling of δ_s, normalized

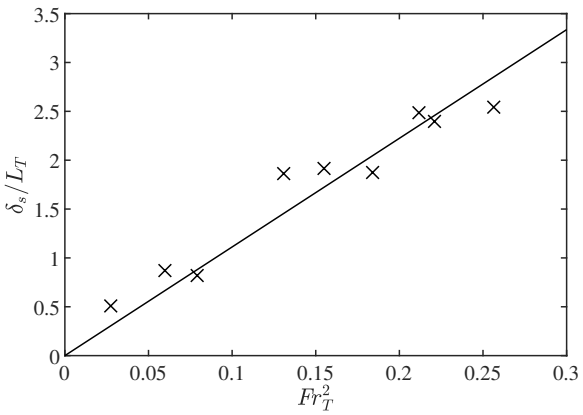


Figure 4-16: Surface layer thickness δ_s, normalized by the characteristic length scale $L_T = u_{\text{rms}}^3/\varepsilon$ measured at $z^* = -0.5$, as a function of Fr_T^2, including (——) the linear fit to (4.34) ($R^2 = 0.909$).

by the characteristic turbulence length scale $L_T = u_{\text{rms}}^3/\varepsilon$, as a function of Fr_T^2. We see a strong linear relationship

$$\delta_s/L_T = C_\delta \, Fr_T^2 \,, \tag{4.34}$$

where the 95% confidence interval is $C_\delta = 11.1 \pm 1.1$. Dimensionalizing,

$$\delta_s = C_\delta \, u_{\text{rms}}^2 \, \text{g}^{-1} \,. \tag{4.35}$$

This is the scaling one would expect from a simple wave argument: that energy is roughly evenly distributed between gravitational potential energy ($\propto \rho_w \text{g} \delta_s$) and kinetic energy ($\propto \rho_w u_{\text{rms}}^2$). We note that the depth that u_{rms} is measured at ($z^* = -0.5$) is a function of δ_s, so (4.35) is not strictly explicit; however, because u_{rms} does not change quickly with depth (see $\tilde{k}$ in figure 4-13), solving (4.35) should not be a challenge.

For a k-ε RANS model, we need a boundary condition which describes the flux of turbulent kinetic energy across the surface layer boundary at $z^* = -0.5$. Brocchini (2002) derive the form of this boundary condition for general FST, and here we study a subset where there is no turbulent energy production in the surface layer and energy flux is driven by only dissipation within the surface layer. This flux of $\tilde{k}$ into the surface layer (per unit area) is given by (4.33). Using (4.35), we can rewrite (4.33) as

$$W/\rho_w = (C_W/C_\delta) \, \varepsilon \, \delta_s \,. \tag{4.36}$$

While Brocchini (2002) consider $k = \overline{\mathbf{u} \cdot \mathbf{u}}/2$ rather than $\tilde{k} = \overline{\rho \mathbf{u} \cdot \mathbf{u}}/2\overline{\rho}$, beneath the surface layer $\gamma \approx 1$ and the difference is negligible. Therefore, this C_W/C_δ is equivalent to C_{diss} in the model derived by Brocchini (2002). We have that $C_W/C_\delta = 0.41 \approx \frac{1}{2}$, meaning that (4.36) is consistent with the first approximation (4.30) developed in section 4.5.3 based on the simplifying assumption of constant ε throughout the surface layer.

4.7 Conclusion

For sufficiently large Fr_T^2, turbulence is strong enough to overcome the restoring force of gravity and break apart the free surface, leading to air entrainment. Our interest in this chapter is characterizing the turbulence in the resulting surface layer, where air and water a highly mixed. The turbulence in this region drives bubble evolution mechanisms, but it is hard to predict using turbulence modeling, like RANS.

We develop a definition of the surface layer thickness δ_s based on the derivative of intermittency $\mathrm{d}\gamma/\mathrm{d}z$ at the mean free-surface height $z = \bar{\eta}$. This removes dependence of δ_s on the tail-behavior of γ, which can be affected by the (Fr_T^2-dependent) bubble void fraction deep beneath the $\bar{\eta}$ and makes no assumptions about the distribution of γ. It turns out that our more robust definition of the surface layer is appropriate to characterize the relevant properties of the near-surface turbulence. In particular, scaling depth by $z^* = (z - \bar{\eta})/\delta_s$ collapses results across broad ranges of Fr_T^2, and there is a clear difference in behavior between within the surface layer and beneath the surface layer, separated by $z^* = -0.5$.

We show the collapse with z^* for intermittency, turbulence isotropy, turbulent momentum flux, and turbulent energy flux. For intermittency, we show that even at large Fr_T^2 the surface layer intermittency is driven by the movement of a mostly intact free surface (rather than bubbles or droplets), and that the distribution is logistic rather than Gaussian as often assumed (Brocchini & Peregrine, 2001*b*). For isotropy, we show a clear transition to nearly isotropic turbulence in the surface layer around $Fr_T^2 \approx 0.1$. This defines the range ($Fr_T^2 > 0.1$) of the strong FST regime (Yu *et al.*, 2019). Our measurements of turbulence highlight that, particularly for strong FST, the direct effects of the free surface are constrained to the surface layer ($z^* \in [-0.5, 0.5]$), and that the turbulence within the surface layer follows a universal behavior characterized by δ_s as well as ε and u_{rms} measured at the bottom of the surface layer ($z^* = -0.5$).

In section 4.6 we discuss that, because the direct effects of the free surface are constrained to the surface layer, RANS turbulence modeling could be applied without modification beneath the surface layer, with a boundary condition at the bottom of the surface layer capturing the indirect effects though energy flux. This is similar to the thin-layer model theorized by Brocchini (2002). We determine the scaling of δ_s and the energy flux into the surface layer, W, with Fr_T^2, and linear regression to DNS data provides the scaling coefficients. These are the two values needed to implement a surface layer boundary condition in k-ε RANS.

Chapter 5

Bubble Fragmentation in Homogeneous Isotropic Turbulence

In this chapter we focus on bubble fragmentation, the $S_f(a)$ term in the population balance equation (1.2). We study fragmentation in homogeneous isotropic turbulence (HIT), which, as shown in Chapter 4, is a good approximation of the turbulence beneath the free surface. For strong FST ($Fr_T^2 > 0.1$), this is true even in close proximity to the free surface. Fragmentation is important to understanding the bubble population beneath entraining free surfaces because, as discussed in section 1.3, a fragmentation cascade is the basis for the $N(a > a_H) \propto a^{-10/3}$ bubble population predicted by Garrett *et al.* (2000). By elucidating three fundamental timescales of fragmentation, we validate some of the modeling assumptions used to obtain $-10/3$ and determine how long it takes for a bubble population to reach this equilibrium solution.

Key results from this chapter are summarized in "Fundamental timescales of bubble fragmentation in homogeneous isotropic turbulence" by Gaylo, Hendrickson & Yue (2023).

5.1 Introduction

For general turbulent flows, there is the large length scale of the flow driven by the geometry and the way in which turbulence is generated, and the small Kolmogorov scale η_T where turbulent kinetic energy is dissipated to heat. As shown in (2.18), the ratio between these two scales with $\sim Re^{3/4}$. For many different turbulent flows with $Re \gg 1$, the turbulence between these two scales (the Kolmogorov inertial sub-range) exhibits universal behavior like that of simple HIT. Hence, fragmentation of bubbles within the Kolmogorov inertial sub-range of HIT is often studied due to its wide applicability, including to free-surface turbulence (Yu *et al.* 2019; Gaylo & Yue 2025; see also Chapter 4).

In HIT, fragmentation is primarily governed by the disturbing effect of turbulent fluctuations and the restoring effect of surface tension. The ratio between the two is given by the bubble Weber number introduced in section 1.3,

$$We_B = \frac{2\varepsilon^{2/3}(2a)^{5/3}}{(\sigma/\rho_w)}, \tag{1.10}$$

where ε is the turbulent dissipation rate, a is the parent bubble's effective radius, and σ is the surface-tension coefficient. The Hinze scale is defined as the Weber number We_H below which surface tension largely prevents fragmentation (Hinze, 1955), corresponding to a bubble radius

$$a_H = 2^{-8/5}\, We_H^{3/5}\ (\sigma/\rho_w)^{3/5}\ \varepsilon^{-2/5}\,. \tag{5.1}$$

Thus, our focus here is moderate ($We_B \gtrsim We_H$) to large ($We_B \gg We_H$) Weber numbers where fragmentation is present. Especially for $We_B \gg We_H$, we expect the daughter bubbles of a fragmentation event to fragment themselves, leading to a fragmentation cascade. An assumption made by Garrett *et al.* (2000) to obtain this fragmentation cascade is that fragmentation is a local process, meaning it is uncommon for fragmentation to form daughter bubbles much smaller than the parent. Locality in fragmentation cascades is confirmed by Chan *et al.* (2021*b*,*c*).

Our interest is the statistical modeling of fragmentation for the population balance equations (PBE) (1.2). In principle, the necessary statistics can be derived from a (more) complete mechanistic description of fragmentation, which is a subject of active investigation (e.g., Liao & Lucas, 2009; Qi *et al.*, 2022; Rivière *et al.*, 2021, 2022). In addition to the challenge of disparate mechanistic descriptions, another challenge is that these often describe the behavior of a bubble as dependent on its history (for example, the two-step process presented by Rivière *et al.* (2022)). Contrarily, the PBE assumes that the statistical behavior of a bubble is independent of its history, i.e., no hysteresis. The present work complements mechanistic studies by focusing on the fundamental statistics of turbulent fragmentation, quantified through their characteristic timescales. While individual physical mechanisms can also be characterized by timescales, such as the timescale for a sufficiently strong eddy to fragment a bubble (Qi *et al.*, 2022) or the timescale for capillary-driven production of sub-Hinze bubbles (Rivière *et al.*, 2021, 2022), our focus is on the timescales that characterize the fundamental statistics of fragmentation. Understanding these timescales will directly support the statistical modeling of bubble size distributions through PBE. Additionally, the understanding provided by these statistical timescales will provide robust ways to compare disparate mechanistic models of fragmentation.

In this work, we elucidate and quantify three fundamental timescales of fragmentation for moderate- to large-*We* HIT. In order of magnitude from small to large, these are: the bubble relaxation time τ_r which characterizes the time from when a bubble is formed to when its subsequent dynamics (e.g., the rate of fragmentation) become statistically stationary, the (well-established) bubble lifetime τ_ℓ which characterizes the time from when a bubble is formed to when it undergoes fragmentation (Martínez-Bazán *et al.*, 1999*a*), and the convergence time τ_c which characterizes the time needed for the air to go from the scale of the largest bubble, radius a_{max}, to the Hinze scale, a_H. In section 5.2 we examine how these timescales relate to statistical modeling of bubble size distributions through PBE. In previous work, τ_c could not be described for realistic fragmentation statistics (Deike *et al.*, 2016; Qi *et al.*, 2020). In section 5.3 we develop a Lagrangian mathematical description which provides the speed at which volume moves through the fragmentation cascade. This volume-propagation speed allows us to derive τ_c for realistic fragmentation statistics at large *We*. We prove that, unlike typical fragmentation statistics, the volume-propagation speed can be obtained independent of the time interval used for measurement. Using DNS of

moderate- to large-*We* bubble fragmentation in HIT (described in section 5.4), in section 5.5 we use ELA to measure fragmentation to quantify the three fundamental timescales. In section 5.6 we discuss the new insights provided by the quantification of these timescales: τ_r shows hysteresis can be neglected in PBE, and τ_c provides a new constraint on large-*We* fragmentation models.

5.2 Three fundamental timescales of fragmentation

To define characteristic timescales of fragmentation, we start by reviewing the statistics typically used to describe fragmentation. As introduced in section 1.2, the fragmentation source term in the PBE can be split into a destruction and a creation term,

$$S_f(a) \equiv S_f^+(a) - S_f^-(a)\,, \tag{1.4}$$

where the destruction term (assuming a Poisson process) is

$$S_f^-(a) = \Omega(a)N(a)\,. \tag{1.5}$$

This introduces the first fragmentation statistic $\Omega(a)$, the fragmentation rate (dimensions $[1/T]$). The creation term $S_f^+(a)$ describes when any bubble of radius $a' > a$ fragments to form a bubble of radius a. This can be written as an integral,

$$S_f^+(a) = \int_a^\infty \bar{m}(a')\beta(a;a')S_f^-(a')\,\mathrm{d}a'\,, \tag{5.2}$$

where we have the other two fragmentation statistics: $\bar{m}(a')$ is the average number of daughter bubbles created by fragmentation of a parent of radius a'; and $\beta(a;a')$ is the daughter-size distribution, expressed as a probability distribution function of daughter radius a for a given parent radius a'. Fragmentation cannot create or destroy air, only move it between bubble sizes. Expressing this volume conservation requirement is easier if we write the daughter-size distribution in terms of a volume ratio $v^* = (a/a')^3$, giving a daughter-size distribution f_V^* related to β by

$$a'\beta(a;a') = 3(v^*)^{2/3}f_V^*(v^*;a')\,. \tag{5.3}$$

Volume conservation requires the distribution satisfy (Martínez-Bazán *et al.*, 2010)

$$\bar{m}(a')\int_0^1 v^* f_V^*(v^*;a')\,\mathrm{d}v^* = 1\,. \tag{5.4}$$

While there is great variety in models for $\bar{m}(a')$ and $\beta(a,a')$ (Liao & Lucas, 2009), models for $\Omega(a)$ generally follow

$$\Omega(a) = C_\Omega(We_B)\varepsilon^{1/3}a^{-2/3}\,, \tag{5.5}$$

where $C_\Omega(We_B)$ approaches a constant value $C_{\Omega,\infty}$ as $We \to \infty$. Dimensional analysis shows

C_Ω could also depend on Reynolds number and an additional parameter, such as the ratio between parent-bubble radius and the Kolmogorov scale, a/η_T, implied by Qi *et al.* (2022); however, the power-law scaling in (5.5) is robust at large We_B (Martínez-Bazán *et al.*, 2010). Assuming $We_B \sim \infty$ to neglect surface tension, this scaling can be arrived at mechanistically by dividing the characteristic velocity of the turbulent fluctuations on the scale of a bubble ($\sim \varepsilon^{1/3} a^{1/3}$) by the characteristic length a bubble must deform to fragment ($\sim 2a$) (Garrett *et al.*, 2000). A model for moderate to large *We* based on the assumption that the rate of fragmentation is proportional to the difference between the deforming force of turbulent fluctuations and the restoring force of surface tension is

$$C_\Omega(We_B) = C_{\Omega,\infty}\sqrt{1 - We_H/We_B}\,, \tag{5.6}$$

with $C_{\Omega,\infty} \approx 0.42$ from experiments (Martínez-Bazán *et al.*, 1999*a*; Martínez-Bazán *et al.*, 2010).

5.2.1 Bubble lifetime, τ_ℓ

The first fundamental timescale is the expected lifetime of a bubble, τ_ℓ, To define this, we start by defining $p_{\text{frag}}(a;T)$ to be the probability of fragmentation over some measurement interval T, i.e., the probability a bubble of radius a present at time t will fragment before the next measurement at time $t+T$. For convince, let $t = 0$ be the time a bubble is created. The probability distribution function for the bubble lifetime ℓ is $\partial p_{\text{frag}}/\partial T$ and the expected lifetime $\tau_\ell = \mathbb{E}[\ell]$ can be calculated. If we assume, as is done in PBE, that the fragmentation rate of a bubble is independent of the time since its formation, then

$$p_{\text{frag}}(a;T) = 1 - \exp\left[-T\Omega(a)\right]\,, \tag{5.7}$$

and the expected lifetime is $\tau_\ell = 1/\Omega(a)$.

5.2.2 Relaxation time, τ_r

For (5.7) (and by extension (1.5) and the PBE (1.2) in general), we assume that the statistics describing fragmentation are independent of bubble age, which we will refer to as the no-hysteresis assumption. This no-hysteresis assumption means that the (statistical) behavior of a bubble after it is created by fragmentation should be indistinguishable from a bubble that has existed for a much longer time. For PBE modeling, it is desirable to assume the effect of hysteresis is negligible; however, at least over short timescales, this seems unlikely physically, as a young bubble must be significantly deformed from equilibrium. Regardless of the mechanistic explanation for fragmentation (either the result of accumulation of surface oscillations (Risso & Fabre, 1998) or a single-sufficiently strong eddy (Martínez-Bazán *et al.*, 1999*a*)), we expect a young bubble to be more likely to fragment, violating no-hysteresis. Although the physical mechanism for hysteresis is unclear, we posit that there exists a timescale τ_r below which it is relevant, and above which it is negligible.

To define τ_r more formally, we start by investigating how $\Omega(a)$ can be related to measurable quantities. As infinitely small temporal resolution is unobtainable, a finite

measurement interval T is inherent in the measurement of fragmentation events, in both experiments and simulations (Vejražka *et al.*, 2018). The validity of the no-hysteresis assumption depends on the timescale T one uses to define fragmentation events (Solsvik *et al.*, 2016), as some T is inherent to any measurement, we will allow for measured fragmentation statistics to depend on T. We rearrange (5.7) to define the T-dependent fragmentation rate

$$\Omega(a;T) \equiv -\ln\left[1 - p_{\text{frag}}(a;T)\right]/T\,. \tag{5.8}$$

For large *We* where daughter bubbles will eventually fragment, it is clear that $\bar{m}$ must also depend on T, and therefore, by (5.4), so must f_V^*. Thus, let $\bar{m}(a';T)$ be the expected number of daughters present at $t+T$ if the bubble fragments and $f_V^*(v^*;a',T)$ be the size distribution of these daughters. In the absence of hysteresis $\Omega(a;T)$ would be independent of T, but the physical dependence of these statistics on T makes them difficult to relate to the statistics describing fragmentation in the PBE (Solsvik *et al.*, 2016). We posit that there exists a timescale τ_r such that $\Omega(a;T \gg \tau_r) = \Omega(a)$ is independent of T. It follows that $\tau_\ell \gg \tau_r$ is required for the no-hysteresis assumption to be valid in PBE.

5.2.3 Convergence time, τ_c

As presented in Chapter 1, $N(a) \propto a^{-10/3}$ is the equilibrium solution ($\partial N/\partial t = 0$) to the PBE with only the fragmentation source term (Garrett *et al.*, 2000), or as more recently shown fragmentation with cut-off power-law entrainment, where the size distribution of the bubbles injected by entrainment follows a power law $I(a) \propto a^\gamma$ where $\gamma > -4$ (Gaylo *et al.*, 2021). The time, τ_c, it takes to reach these equilibrium solutions is of interest: if τ_c is much less than the timescale over which the flow is transient, we expect an equilibrium fragmentation cascade (generally $N(a) \propto a^{-10/3}$) to be obtained. Gaylo *et al.* (2021) provide an expression for τ_c which allows for arbitrary f_V^* and $\bar{m}$, but its derivation is specific to power-law entrainment. For general fragmentation cascades, τ_c is characterized by the time it takes for the volume of the largest bubble (radius a_{max}) to reach the Hinze scale a_H (Deike *et al.*, 2016; Qi *et al.*, 2020). This characterization is useful because it allows τ_c to be measured independent of the evolution of $N(a)$. Additionally, being directly related to fragmentation, it could provide a constraint on the fragmentation statistics in PBE (Qi *et al.*, 2020). However, current derivations of τ_c from fragmentation statistics assume that bubbles break into identically sized daughters, ignoring the effect of f_V^*. Although Monte Carlo simulation can be used to determine what τ_c is predicted by given fragmentation statistics (Qi *et al.*, 2020), the lack of a general analytic expression relating τ_c to realistic fragmentation statistics precludes the reverse; it is unclear how a given value of τ_c constrains fragmentation statistics.

5.3 Describing τ_c using a Lagrangian description of fragmentation cascades

In this section, we derive a general analytic expression that relates τ_c to realistic fragmentation statistics. Previous derivations of τ_c assume identical fragmentation and that the life of a

bubble is exactly (rather than on the average) equal to τ_ℓ so that the cascade can be treated as a series of discrete deterministic fragmentation events (Deike *et al.*, 2016). While this approach provides the general physical scaling of τ_c, it is unable to directly relate τ_c to realistic fragmentation statistics. In this section we use a Lagrangian air particle-based mathematical description of the speed at which volume moves through fragmentation cascades to derive τ_c. We note that this is a "speed" in the abstract sense as it measures how quickly air moves from large bubbles to small bubbles through the fragmentation cascade rather than through physical space. However, this description is useful as, through this speed, τ_c can be related to realistic fragmentation statistics and this speed is also directly accessible from volume-based bubble-tracking (Gaylo *et al.* 2022; see also Chapter 3). Although T-independence is obvious when τ_c is obtained through the evolution of $N(a)$, it is less clear when τ_c is obtained through fragmentation statistics, which generally depend on T. We show that our approach allows fragmentation statistics-based measurement of τ_c independent of T.

Throughout this section, we consider large $We_B \gg We_H$ so that we can assume that fragmentation statistics are scale-invariant and simplify (5.6) to a Heaviside step function:

$$C_\Omega(We_B) = C_{\Omega,\infty}\,\mathcal{H}\left(1 - We_H/We_B\right) . \tag{5.9}$$

In the following derivation, we also assume no-hysteresis, limiting applicability to timescales much larger than τ_r.

5.3.1 Lagrangian-based mathematical description of fragmentation cascades

Previous work on the movement of volume in fragmentation cascades applies Eulerian descriptions, focusing on volume flux. To find the equilibrium between entrainment and fragmentation, Gaylo *et al.* (2021) balance the flux of volume in and out of the set of bubbles of a given range of sizes. To evaluate locality, Chan *et al.* (2021*b*) study the flux of volume from bubbles larger than a given size to those smaller. Eulerian descriptions are useful to model the volume flow in and out of specified bubble sizes, but to derive τ_c we need to understand how any specific air volume flows through the entire cascade. For this, a Lagrangian description is more direct.

Consider how a single Lagrangian particle of air p moves through a fragmentation cascade, illustrated in figure 5-1. We define the function $a_p(t)$ to be the radius of the bubble that contains the particle of air p at time t. We note that, because one bubble breaks up into two instantaneously, $a_p(t)$ is a step function that does not have a well-defined derivative. Still, we can use $a_p(t)$ to write a simple expression for expression for τ_c: Defining time for a particle such that $a_p(t = 0) = a_{max}$, our interest is the expected time for the particle to reach the Hinze scale,

$$\tau_c \equiv \mathbb{E}\left\{\min\left\{t \;:\; a_p(t) \leq a_H\right\}\right\} . \tag{5.10}$$

The expected value here refers to an ensemble average over an arbitrarily large set of particles p.

Before addressing (5.10) directly, we need to develop a relationship between the ensemble average behavior of these particles and the previous bubble-based statistical description of

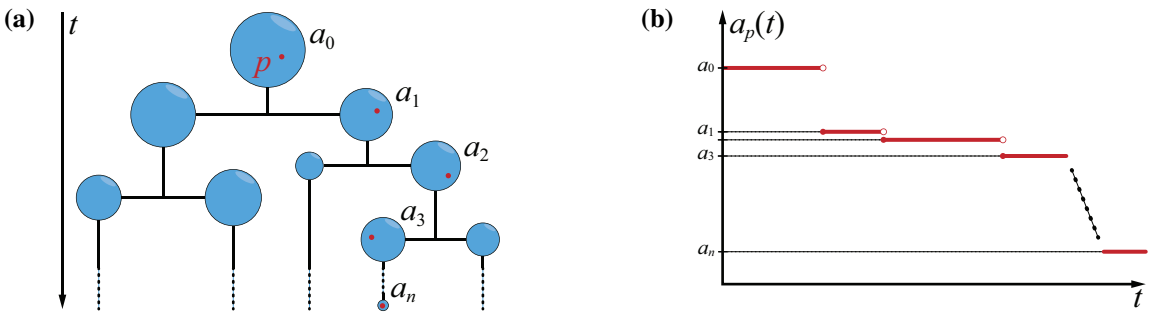


Figure 5-1: (a) Schematic of the Lagrangian description showing the path of a Lagrangian air particle p through a sequence of fragmentations from large to small radii, $a_0, a_1, \ldots a_n$, of the bubble containing p; and (b) the corresponding function $a_p(t)$ describing the evolution of this bubble radius. Describing the radius of the bubble containing p as a function of time allows a propagation speed of p through the cascade to be defined.

fragmentation. This is easier in terms of bubble volume rather than radius. Incorporating the measurement interval T, for a given particle p we define the volume ratio between the bubble containing p at time t and the bubble containing p at time $t+T$:

$$v_R(t;T) \equiv \left[a_p(t+T)/a_p(t)\right]^3 . \tag{5.11}$$

If the bubble containing p at time t does not fragment over the measurement interval T, then $v_R = 1$. Trivially, any moment n of the distribution of v_R given no fragmentation is

$$\mathbb{E}\left\{[v_R(T^*)]^n \;\middle|\; \text{no frag}\right\} = 1 . \tag{5.12}$$

We now consider the case where the bubble containing p at time t fragments, in which case v_R depends on the size of the daughter bubble that p ends up in. It is important to note that the probability p ends up in a given daughter is equivalent to v^*, the ratio of the volume of the daughter to that of the parent. For example, if the bubble containing p broke up into two dissimilar daughter bubbles, the larger bubble three times the volume of the smaller bubble ($v^* = 3/4$ and $v^* = 1/4$ respectively), p would have a $3/4$ probability of ending up in the larger bubble and a $1/4$ probability of ending up in the smaller bubble. With this subtlety in mind, the probability distribution function for v_R given that fragmentation occurs, $f_{V_R \mid \text{frag}}$, is related to the previous fragmentation statistics through

$$f_{V_R \mid \text{frag}}(v_R;t,T) = \bar{m}(a_p(t);T)\; v_R\; f_V^*(v_R;a_p(t),T) . \tag{5.13}$$

We assume these statistics are scale invariant and introduce the non-dimensional parameter $T^* = T\varepsilon^{1/3}a_p(t)^{-2/3}$. This gives

$$f_{V_R \mid \text{frag}}(v_R;T^*) = \bar{m}(T^*)\; v_R\; f_V^*(v_R;T^*) , \tag{5.14}$$

and any moment n of the distribution is given by

$$\mathbb{E}\left\{[v_R(T^*)]^n \;\middle|\; \text{frag}\right\} = \bar{m}(T^*) \int_0^1 v^{*n+1} f_V^*(v^*; T^*)\, \mathrm{d}v^* \,. \tag{5.15}$$

To find the total distribution of v_R, we combine (5.12) and (5.15) using the law of total probability,

$$\mathbb{E}\left\{[v_R(T^*)]^n\right\} = p_{\text{frag}}\, \mathbb{E}\left\{[v_R(T^*)]^n \;\middle|\; \text{frag}\right\} + \left(1 - p_{\text{frag}}\right) \,. \tag{5.16}$$

Rearranging a little,

$$1 - \mathbb{E}\left\{[v_R(T^*)]^n\right\} = p_{\text{frag}} \left(1 - \mathbb{E}\left\{[v_R(T^*)]^n \;\middle|\; \text{frag}\right\}\right) \,. \tag{5.17}$$

Assuming no hysteresis, we use (5.7) to describe p_{frag}. Applying (5.9), the simplified model of $C_\Omega(We_B)$, and writing in terms of $T^* = T\varepsilon^{1/3} a_p(t)^{-2/3}$, we have

$$p_{\text{frag}}(T^*) = 1 - \exp[-C_{\Omega,\infty}\, T^*] \tag{5.18}$$

for within the fragmentation cascade ($a_p > a_H$). Finally, we can relate the distribution of v_R to the classical fragmentation statistics,

$$1 - \mathbb{E}\left\{[v_R(T^*)]^n\right\} = \left(1 - \exp[-C_{\Omega,\infty}\, T^*]\right) \left[1 - \bar{m}(T^*) \int_0^1 v^{*n+1} f_V^*(v^*; T^*)\, \mathrm{d}v^*\right] \,. \tag{5.19}$$

5.3.2 Defining the volume-propagation speed in a fragmentation cascade

To obtain τ_c, we derive a metric that measures the speed at which Lagrangian air particles move through fragmentation cascades. To derive a speed, we must first define the "location," $x(t)$, of a Lagrangian air particle p within the cascade. In this case location refers to some abstract scalar bubble-size metric within the cascade rather than a physical spatial coordinate. This means we have great freedom in how we map $a_p(t)$ to location $x(t)$. We choose a mapping such that the average of the associated speed $s(t) \equiv \dot{x}(t)$ is constant for $a_p(t) > a_H$. A constant average speed is necessary for many of the properties that will follow and, as a result of the scaling in (5.5), is achieved only by $x(t) \propto a_p(t)^{2/3}$. We choose

$$x(t) \equiv -\varepsilon^{-1/3} a_p(t)^{2/3} \,, \tag{5.20}$$

which has dimensions of time, so that, in addition to being constant, the time-derivative of $x(t)$,

$$s(t) = -\frac{2}{3}\varepsilon^{-1/3} a_p(t)^{-1/3} \frac{\mathrm{d}}{\mathrm{d}t} a_p(t) \,, \tag{5.21}$$

is also positive and non-dimensional.

Because $a_p(t)$ is a step function, the derivative in (5.21) is ill defined. However, we can

take the time-average of $s(t)$ over a measurement interval T,

$$\langle s(t)\rangle_T \equiv \frac{1}{T}\int_t^{t+T} s(t')\,\mathrm{d}t' \,. \tag{5.22}$$

This gives

$$\langle s(t)\rangle_T = \frac{x(t+T)-x(t)}{T} = \frac{\varepsilon^{-1/3}a_p(t)^{2/3}}{T}\left(1-[v_R(t;T)]^{2/9}\right), \tag{5.23}$$

where (5.11) defines the volume ratio $v_R(t;T)$. As we have assumed scale-invariance, we can then introduce non-dimensional time $T^* = T\varepsilon^{1/3}a_p(t)^{-2/3}$ to obtain

$$\langle s(t)\rangle_{T^*} = \frac{1-[v_R(T^*)]^{2/9}}{T^*} \,. \tag{5.24}$$

This equation gives the time-averaged propagation speed for a single particle.

Next, we perform an ensemble average to get

$$\mathbb{E}\left\{\langle s(t)\rangle_{T^*}\right\} = \frac{1-\mathbb{E}\left\{[v_R(T^*)]^{2/9}\right\}}{T^*}, \tag{5.25}$$

the expected time-averaged speed for an ensemble of (independent) Lagrangian air particles. Using (5.19) with $n = 2/9$, we obtain

$$\mathbb{E}\left\{\langle s(t)\rangle_{T^*}\right\} = C_{\Omega,\infty}\frac{1-\exp[-C_{\Omega,\infty}T^*]}{C_{\Omega,\infty}T^*}\left[1-\bar{m}(T^*)\int_0^1 v^{*11/9}f_V^*(v^*;T^*)\,\mathrm{d}v^*\right]. \tag{5.26}$$

The limit $T^* \to 0$ gives the expected instantaneous speed,

$$\bar{s} \equiv \lim_{T^*\to 0}\mathbb{E}\left\{\langle s(t)\rangle_{T^*}\right\} = C_{\Omega,\infty}\left[1-\bar{m}\int_0^1 v^{*11/9}f_V^*(v^*)\,\mathrm{d}v^*\right], \tag{5.27}$$

where $\bar{m}$ and $f_V^*(v^*)$ describe the fragmentation statistics for $T^* \to 0$ and are equivalent to those in (5.2).

Hereafter, we refer to $\bar{s}$ as the volume-propagation speed of a fragmentation cascade. Although the size locations of individual Lagrangian air particles in the cascade follow step functions, by commuting time averaging and ensemble averaging, we are able to obtain an average instantaneous speed for particles in the cascade. This speed $\bar{s}$ can be related to fragmentation statistics measured over finite intervals T with (5.26), or the instantaneous statistics used by PBE with (5.27). The relationship between the two is explored in §5.3.4. In §5.3.3 we use $\bar{s}$ to provide τ_c.

5.3.3 Describing convergence time, τ_c

As intended, our choice of the definition of location within the cascade, $x(t)$, makes $\bar{s}$ constant for $a_p(t) > a_H$. This constant speed means that, despite $x(t)$ being a step function,

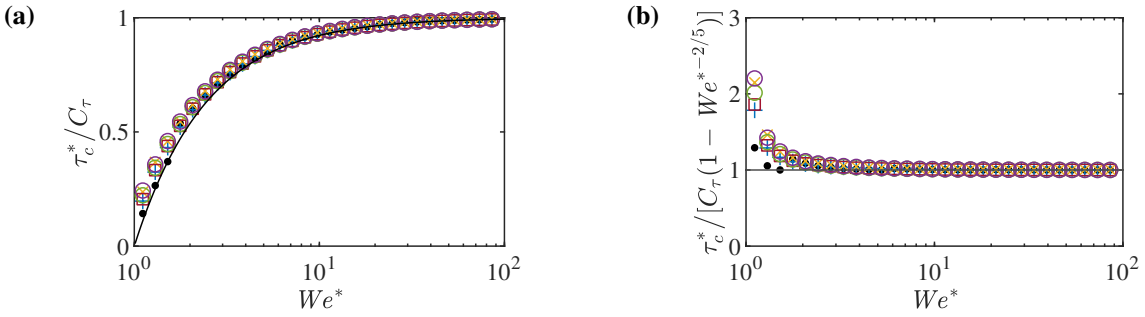


Figure 5-2: The effect of We^* on τ_c^* as modeled by (5.29) (——) compared to Monte Carlo simulations of daughter distributions, •, A; +, B; ×, C; □, D; ○, E; ○, F (see table 5-1). (5.9) is used to model the Hinze scale. The 95% C.I. on all τ_c^* is < 1%.

after a sufficient number of steps, we can treat fragmentation as a continuous process and apply the approximation $x(t) \approx t\bar{s}$ with reasonable (statistical) accuracy. Thus, we can approximate τ_c as the distance in x between a_{max} and a_H divided by this speed,

$$\tau_c = \frac{\left(\varepsilon^{-1/3} {a_{max}}^{2/3}\right) - \left(\varepsilon^{-1/3} {a_H}^{2/3}\right)}{\bar{s}} . \tag{5.28}$$

Nondimensionalizing $\tau_c^* = \tau_c \, \varepsilon^{1/3} {a_{max}}^{-2/3}$ and defining We_{max} to be the We_B associated with a_{max},

$$\tau_c^* = C_\tau \left[1 - (We_{max}/We_H)^{-2/5}\right] \; ; \quad C_\tau \equiv 1/\bar{s} \, . \tag{5.29a, b}$$

Despite the approximation used to derive (5.28) from $\bar{s}$ in (5.27), (5.29) is expected to be valid for $We^* \equiv We_{max}/We_H$ not small (where multiple fragmentation events are generally necessary to reach a_H). This is confirmed by Monte Carlo simulations of prescribed fragmentation statistics (figure 5-2).

For $We \sim \infty$ we recover the same $\tau_c \propto \varepsilon^{-1/3} {a_{max}}^{2/3}$ scaling as previous work which assumes identical fragmentation (Deike *et al.*, 2016). This scaling of τ_c is like τ_ℓ, demonstrating that the fragmentation rate is the dominant factor in determining τ_c. Our propagation speed-based analysis provides the scaling constant C_τ which quantifies the contribution of fragmentation rate, as well as fragmentation statistics $\bar{m}$ and $f_V^*(v^*)$. For large-but-finite We, (5.29) captures the effect of the We-driven separation between a_{max} and a_H on the value of τ_c; however, we note that the scaling or τ_c with We will be more complex for small We ($We \sim We_H$) as we have not incorporated the effect of finite-We on fragmentation rate, such as modeled by (5.6), into our propagation speed-based analysis. In section 5.5.3, DNS shows for what sufficiently large We this effect is negligible.

Although primarily driven by fragmentation rate, τ_c is also related to the fragmentation statistics $\bar{m}$ and $f_V^*(v^*)$ (Qi *et al.*, 2020), which is now quantified by the scaling constant C_τ. To describe these relationships, we follow Gaylo *et al.* (2021) and isolate the effect of f_V^* from $\bar{m}$ through a daughter-distribution constant C_f, defined as the ratio between C_τ and a C_τ found using the same $\bar{m}$ but identical fragmentation, $f_V^*(v^*) = \delta(v^* - 1/\bar{m})$, where δ is

	Daughter Distribution	m	$f_V^*(v^*)$	C_f	$C_f{}^\star$
A	Valentas *et al.* (1966)	2	$\delta(v^* - 1/2)$	1	1
B	Martínez-Bazán *et al.* (1999*b*)	2	$(v^*)^{2/9}\,(1-v^*)^{2/9}$	1.348	1.314
C	Tsouris & Tavlarides (1994)	2	$2^{1/3} - (v^*)^{2/3} - (1-v^*)^{2/3}$	2.432	2.255
D	Martínez-Bazán *et al.* (2010)	2	$(v^*)^{-4/9}\,(1-v^*)^{-4/9}$	1.782	1.712
E	Diemer & Olson (2002)	3	$(v^*)^{1/4}\,(1-v^*)^{3/2}$	1.269	1.253
F	Diemer & Olson (2002)	4	$(v^*)^{1/2}\,(1-v^*)^{7/2}$	1.190	1.185

Table 5-1: Daughter distributions used in Monte Carlo simulations and corresponding daughter-distribution constants C_f defined by equation (5.30) compared to $C_f{}^\star$ defined by Gaylo *et al.* (2021, Eq. (4.3)). Note, a constant to ensure $\int f_V^*(v^*)\mathrm{d}v^* = 1$ is omitted for brevity.

the Dirac delta function. This gives

$$C_\tau = \frac{C_f/C_{\Omega,\infty}}{1-\bar{m}^{-2/9}}\,; \quad C_f = \frac{1-\bar{m}^{-2/9}}{1-\bar{m}\int_0^1 v^{*11/9} f_V^*(v^*)\,\mathrm{d}v^*}\,. \tag{5.30a,b}$$

In table 5-1 we compare this C_f for general fragmentation cascades to the similar constant (hereafter denoted as $C_f{}^\star$) derived by Gaylo *et al.* (2021) for the special case of power-law entrainment. The values are nearly equivalent, and, noting that $(9/2)(\ln\bar{m})^{-1} \approx (1-\bar{m}^{-2/9})^{-1}$, (5.30) predicts similar τ_c as Gaylo *et al.* (2021) for their special case.

5.3.4 Measurement-interval independence of volume-propagation speed

A consequence of $\bar{s}$ being constant for $a_p(t) > a_H$ is that the time-averaged value and the instantaneous speed are equal, $\mathbb{E}\left\{\langle s(t)\rangle_T\right\} = \bar{s}$, so long as $a_p(t+T) > a_H$. Thus, to obtain $\bar{s}$ we must choose a T such that $\Pr\{a(t+T) > a_H\} \approx 1$. For measurements of an initial parent-bubble radius $a = a_p(t)$, we define an upper bound T_U as the interval we expect $a_p(t+T_U) \sim a_H$ and require $T \ll T_U$. Through the same arguments used to derive τ_c, this upper bound is

$$T \ll \varepsilon^{-1/3} a^{2/3} C_\tau \left[1 - (We_B/We_H)^{-2/5}\right]\,. \tag{5.31}$$

For $a = a_{max}$ this is simply $T \ll T_U = \tau_c$. From Monte Carlo simulations of prescribed fragmentation statistics measuring initial bubbles $a = a_{max}$, figure 5-3 confirms that $\mathbb{E}\left\{\langle s\rangle_T\right\}$ gives an exact, T-independent measurement of $\bar{s}$ for $T \ll \tau_c$. T_U provides an upper bound on T for experiments or simulations, although we point out that it is an *a posteriori* measure because C_τ is derived from $\bar{s}$.

Finally, T-independence means $\mathrm{d}\,\mathbb{E}\left\{\langle s(t)\rangle_{T^*}\right\}/\mathrm{d}T^* = 0$. Taking the derivative of (5.26), we obtain

$$T^* \frac{\mathrm{d}\,\mathbb{E}\left\{\left[v_R(T^*)\right]^{2/9}\right\}}{\mathrm{d}T^*} - \mathbb{E}\left\{\left[v_R(T^*)\right]^{2/9}\right\} + 1 = 0\,, \tag{5.32}$$

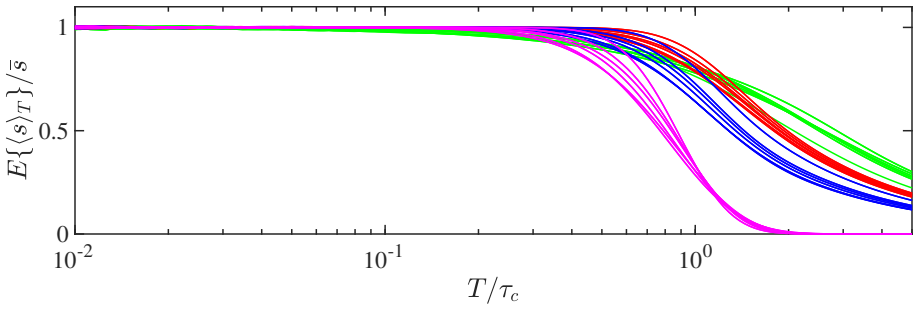


Figure 5-3: Measurements of $\mathbb{E}\{\langle s\rangle_T\}$ from Monte Carlo simulations of daughter distributions A–F (see table 5-1) at a range of T/τ_c, normalized by $\bar{s}$ calculated using (5.27). Colors based on We^*: green, 2; red, 50; blue, 100; magenta, 200, where (5.9) is used to model the Hinze scale. The 95% C.I. on $\mathbb{E}\{\langle s\rangle_T\}$ for $T/\tau_c < 1$ is < 3%.

which is solved by

$$\frac{1-\mathbb{E}\left\{[v_R(T^*)]^{2/9}\right\}}{T^*} = \text{Constant}\,, \tag{5.33}$$

where c is some constant. With (5.19) this gives,

$$\frac{1-\exp[-C_{\Omega,\infty}\,T^*]}{T^*}\left[1-\bar{m}(T^*)\int_0^1 v^{*n+1} f_V^*(v^*;T^*)\,\mathrm{d}v^*\right] = \text{Constant}\,. \tag{5.34}$$

This bounds how scale-invariant fragmentation statistics $\bar{m}(T^*)$ and $f_V^*(v^*;T^*)$ can depend on T^* and provides insight into the relationship between $\bar{m}(T^*)$ and $f_V^*(v^*;T^*)$ measured at large T^* versus the theoretical $T^* \to 0$ limiting case used in PBE. This is useful because a finite relaxation time τ_r implies a lower bound ($T > \tau_r$) for measuring fragmentation statistics that are compatible with the PBE no-hysteresis assumption.

5.4 Direct numerical simulation of bubble fragmentation in homogeneous isotropic turbulence

Using the solver described in Chapter 2, we perform DNS of populations of bubbles fragmenting in HIT. During these simulations, ELA allows direct measurement of individual fragmentation events (Gaylo *et al.* 2022; see also Chapter 3). A summary of the DNS performed is provided in table 5-2. In section 5.5 we will use these measurements of fragmentation to quantify the three fundamental timescales of fragmentation.

5.4.1 Simulation setup

The simulation can be broken into two phases. For the first phase, we develop steady single-phase HIT. To initialize the second step, we take the velocity field from the single-phase HIT and create bubbles by prescribing the VOF field. We measure the fragmentation of these bubbles.

We_T	Re_T	We_B	Δ/η_T	We_Δ	Δ/a_H	N_{sims}	N_{frag}	C_Ω	C_τ
400	200	101–142	1.1	0.66	0.71	7	213	1.64 ± 0.42	8.9 ± 1.9
200	200	50–71	2.2	0.66	0.93	7	106	0.60 ± 0.13	16.1 ± 2.9
			1.5	0.44	0.62	7	189	1.21 ± 0.34	10.2 ± 2.5
			1.1	0.33	0.47	7	208	1.64 ± 0.44	9.8 ± 2.8
			0.7	0.22	0.31	5	187	1.77 ± 0.26	10.3 ± 2.1
100	200	25–36	1.1	0.16	0.31	7	218	1.50 ± 0.27	10.0 ± 2.3
50	200	13–18	1.1	0.08	0.20	7	174	0.93 ± 0.13	15.2 ± 2.9
25	200	6.3–8.9	1.1	0.04	0.13	7	113	0.44 ± 0.12	27.1 ± 5.5

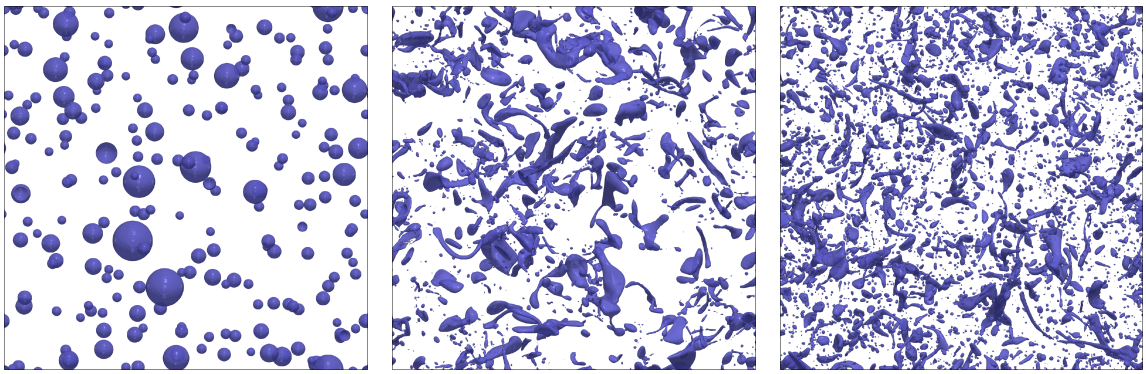
Table 5-2: Summary of HIT simulations performed and values measured using $\Delta t_s/t_\ell = 0.4$, including 95% C.I.. N_{sims} is the number of simulations (each with different initial bubble populations) and N_{frag} is the total number of fragmentation events. a_H is calculated using $We_H \approx 7$ from §5.5.2.

To develop the initial turbulent velocity field for the simulation, we use the linear forcing method described in §2.1.5 on a triply periodic cubic domain. Following Rosales & Meneveau (2005), we choose the forcing parameter $A = 1/3$ and domain length $L = 5.2751$ to obtain a characteristic turbulent dissipation rate $\varepsilon = 1$ and velocity fluctuation $u_{\text{rms}} = 1$. An equivalent interpretation is that we use the characteristic turbulent scales to nondimensionalize all values in the simulation (i.e., to go from (2.1b) to (2.5)). For all simulations we set $Re = 200$ in (2.5). Because we have nondimensionalized by the characteristic turbulent scales, the turbulent Reynolds number is the same, $Re_T = u_{\text{rms}}^4 \rho_w / \varepsilon \mu_w = 200$. The velocity field is initialized with random noise, and then the single-phase simulation runs until u_{rms} and ε reach a statistically steady state.

Using the single-phase HIT as the initial velocity field, we perform simulations with an ensemble of different initial air-water[1] bubble populations, all with void fraction $\sim 1\%$. Populations are created by randomly distributing (without overlap) spherical bubbles with radii between $3L/256$ and $15L/256$ following $N(a) \propto a^{-10/3}$. By repeating the random generation and distribution of bubble populations in the (same) initial HIT velocity field, unique but statistically similar initial bubble populations are generated to provide statistical variation between our ensemble simulations.

After applying the VOF field describing the initial population of (spherical) bubbles, we continue the simulation. Although this abrupt introduction of bubbles to previously single-phase HIT is non-physical, numerical simulations rapidly adjust (Yu *et al.*, 2019; Rivière *et al.*, 2021). As the bubble population evolves under the effects of fragmentation, we continue to apply linear forcing, but (as discussed in §2.1.5) we only apply it to regions of water. This maintains $\varepsilon \approx 1$ and $u_{\text{rms}} \approx 1$. For the two-phase simulations, we apply a range of We to obtain different turbulent Weber numbers, $We_T = u_{\text{rms}}^5 / \varepsilon(\sigma/\rho_w)$, shown in table 5-2.

[1]For this chapter, we use $\lambda = 0.001$ and $\eta = 0.01$, only slightly different than $\lambda = 0.00123$ and $\eta = 0.0159$ used elsewhere.



(a) $t/t_\ell = 0$ **(b)** $t/t_\ell = 1$ **(c)** $t/t_\ell = 3$

Figure 5-4: Evolution of the $f = 0.5$ iso-surface in the three-dimensional HIT simulation with $We_T = 100$.

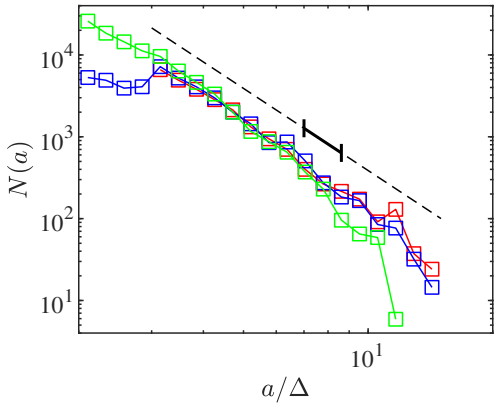


Figure 5-5: Average bubble size distribution $N(a)$ for $We_T = 100$ simulations at times: □, $t/t_\ell = 0$; □, $t/t_\ell = 1$; and □, $t/t_\ell = 3$. $N(a) \propto a^{-10/3}$ is provided for reference over the range of initialized spherical bubbles (– – –) and the range of measured parent bubbles, $a_0 < a < 1.2a_0$ (——).

For one of the $We_T = 100$ simulations, figure 5-4 shows the evolution of the bubble population. Based on averaging over the ensemble of 7 $We_T = 100$ simulations, figure 5-5 shows the evolution of the bubble size distribution $N(a)$. We note that, with our focus on bubbles $a > a_H$, the transition to a distinct power-law regime for $N(a < a_H)$ is not captured (Deane & Stokes, 2002). During the evolution, we use ELA to measure fragmentation statistics for parent bubbles of radii $a_0 < a < 1.2a_0$, where $a_0 = 7L/256$ provides a balance between the number of observed fragmentation events per simulation and resolution of the daughter bubbles. By initializing the bubbles to follow an equilibrium fragmentation cascade $N(a) \propto a^{-10/3}$ (Garrett *et al.*, 2000), the fragmentation of bubbles $a > a_0$ maintains the population of bubbles $a \sim a_0$ for $t/t_\ell < 3$, where

$$t_\ell = (0.42)^{-1}\varepsilon^{-1/3}{a_0}^{2/3} \tag{5.35}$$

is an *a priori* estimate of τ_ℓ (Martínez-Bazán *et al.*, 1999*a*). To exclude the fragmentation of the initial set of spherical bubbles (see figure 5-4), we study fragmentation over $1 < t/t_\ell$. Thus, by measuring fragmentation statistics over $1 < t/t_\ell < 3$, we measure a quasi-steady population of parent bubbles that are realistically formed by a fragmentation cascade.

Chapter 3 explains how ELA allows us to identify fragmentation events and measure the daughter bubbles (see §3.2.2 for details). An important note is that ELA measures the evolution of bubbles between snapshots t^n and $t^{n+1} = t^n + \Delta t_s$. Thus, the measurement interval T discussed in section 5.2.2 for general (experimental or numerical) measurements of fragmentation is equivalent to Δt_s for ELA measurements.

5.4.2 Grid independence

As discussed in Chapter 2 (see §2.1.4), the choice of grid size, Δ, is driven by resolving the Kolmogorov microscale ($\Delta/\eta_T \lesssim 1$) and surface tension ($We_\Delta \lesssim 1$ and/or $\Delta/a_H \lesssim 1$). Based on these metrics we find $L/\Delta = 256$ ($\Delta/\eta_T = 1.1$ in table 5-2) resolves turbulence and surface tension for our entire range of We_T.

In addition to resolving the general physics, another grid resolution concern is resolving the daughter bubbles produced by a fragmentation event. With no clear lower limit to the ratio between the daughter-bubble and parent-bubble volume (v^*), grid resolution limitations require us to filter out daughter bubbles of radius $a < 2\Delta$. Figure 5-6 shows that the bubble size distribution of filtered bubbles, $N(a > 2\Delta)$, is grid independent. For $L/\Delta = 256$ and parent bubbles $a_0 = 7L/256$, daughter bubbles of radius $a < 2\Delta$ correspond to $v^* < 0.02$. While this filter prevents us from measuring the full range of possible daughter bubbles, especially sub-Hinze daughters, we expect this to have little effect on the statistics of interest for two reasons. First, sub-Hinze bubble production by fragmentation happens concurrently with the production of large daughter bubbles (Rivière *et al.*, 2022), so excluding small daughters should not affect the measured rate of fragmentation used to obtain τ_r and τ_ℓ. Second, for τ_c, the integral of the daughter-size distribution in (5.30) weights local daughter production ($v^* \sim 1/\bar{m}$) over non-local daughter production ($v^* \ll 1$), making the contribution of the excluded small daughters small. This is related conceptually to locality, which suggests $v^* \ll 1$ can be neglected when modeling the cascade (Chan *et al.*, 2021*b*,*c*).

To confirm that we resolve turbulence and surface tension, that the filter has a negligible

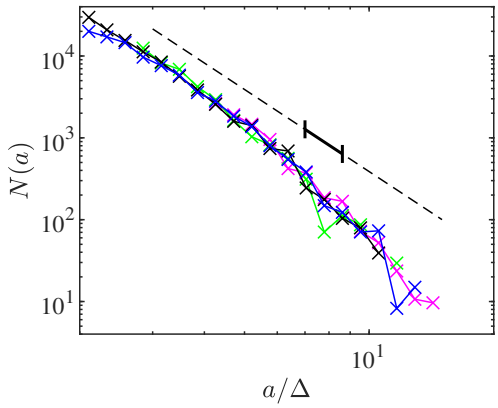


Figure 5-6: Average bubble size distribution $N(a)$ for $We_T = 200$ at time $t/t_\ell = 3$ from simulations with girds: ×, $L/\Delta = 128$; ×, $L/\Delta = 192$; ×, $L/\Delta = 256$; ×, $L/\Delta = 384$; Horizontal axis is normalized by $\Delta = L/256$ and $N(a) \propto a^{-10/3}$ is provided for reference over the range of initialized spherical bubbles (– – –) and the range of measured parent bubbles, $a_0 < a < 1.2a_0$ (——).

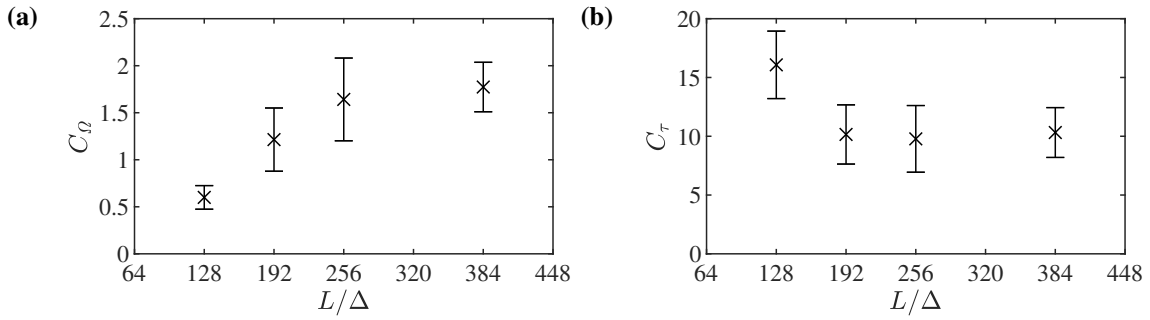


Figure 5-7: Grid-convergence study for (a) fragmentation rate constant C_Ω and (b) convergence constant C_τ based on simulations of $We_T = 200$ (parent bubbles $We_B = 50 - 71$) with different grids, measured using $T/t_\ell = 0.4$. Error bars indicate 95% C.I..

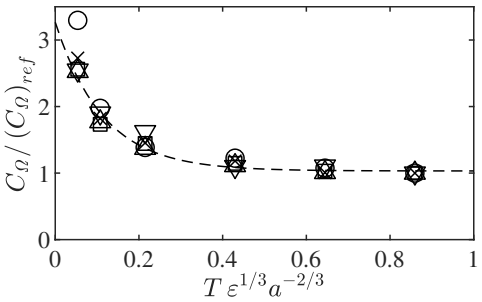


Figure 5-8: Measured fragmentation-rate constant C_Ω normalized by $(C_\Omega)_{ref}$, the value measured using $T/t_\ell = 0.4$, for We_B of (○) 101–142; (×) 50–71; (□) 25–36; (△) 13–18; (▽) 6.3–8.9. Variance-weighted least-squares fit of all data to (5.37) (– – –) gives $C_r = 0.11$ and $A = 2.2$ ($R^2 = 0.954$).

effect, and (more broadly) that the statistics we measure are independent of the grid, we perform a convergence study for $We_T = 200$ using three additional grids, $L/\Delta = 128$, 192, and 384. The results of this convergence study (see figure 5-7) show that our measurements of fragmentation statistics $\mathbb{E}\{\langle s\rangle_T\}$ and $p_{\text{frag}}(a;T)$ (from which the timescales will be calculated) are grid independent for $L/\Delta \geq 256$.

5.5 Estimating the fundamental timescales using Eulerian label advection

5.5.1 Relaxation time, τ_r

We first seek the relaxation time, τ_r. For each simulation, we use 6 instances of ELA with different measurement intervals $T = \Delta t_s$. Using (5.5) and (5.8), we calculate the value of $C_\Omega(We_B;T)$ from each $p_{\text{frag}}(a;T)$ measured using ELA:

$$C_\Omega(We_B;T) = \frac{-\ln\left[1 - p_{\text{frag}}(a;T)\right]}{T\varepsilon^{1/3}a^{-2/3}}\,. \tag{5.36}$$

Figure 5-8 shows how T affects the value of $C_\Omega(We_B;T)$ for a range of We_B. If the no-hysteresis assumption were valid for all T, C_Ω would be a constant for each We_B; however, figure 5-8 shows a strong dependence on small $T^* = T\varepsilon^{1/3}a^{-2/3}$.

We observe that the dependence of $C_\Omega(We_B;T)$ on T is approximately exponential, which provides an empirical definition of the relaxation time τ_r as well as the hysteresis strength A:

$$C_\Omega(We_B;T)/C_\Omega(We_B;T=\infty) = 1 + A\exp[-T/\tau_r]\,. \tag{5.37}$$

We observe that τ_r scales like τ_ℓ. Thus, we define the scaling constant C_r and write

$$\tau_r = C_r\,\varepsilon^{-1/3}a^{2/3}\,, \tag{5.38}$$

where least-squares regression of the combined data for all We_B gives $C_r \approx 0.11$ and $A \approx 2.2$, with a coefficient of determination $R^2 = 0.954$. As shown in figure 5-8, (5.37) does a good

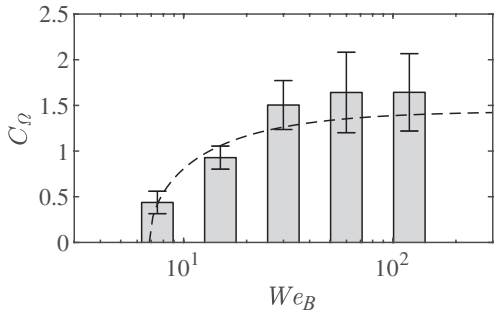


Figure 5-9: Fragmentation rate constant C_Ω as functions of We_B, measured using $T/t_\ell = 0.4$. Error bars indicate 95% C.I.. Variance-weighted least-squares fit to (5.6) (– – –) gives $We_H = 6.9$ and $C_{\Omega,\infty} = 1.4$ ($R^2 = 0.890$).

job explaining the dependence of $C_\Omega(We_B; T)$ on T for all We_B considered.

The scaling of τ_r being like τ_ℓ rather than, say, bubble natural period, $We_B^{-1/2}\varepsilon^{-1/3}a^{2/3}$, is notable because it suggests that, at least for $We_B > We_H$, the physical mechanisms for the decay of hysteresis are not related to surface tension. Future, more detailed, studies of the dynamics of individual bubbles are necessary to understand hysteresis and identify the mechanisms for its decay. Rather than focus on the mechanism of hysteresis, for our statistical study of fragmentation our primary concern is more pragmatic: answering when $T \gg \tau_r$ and hysteresis can be neglected. Hereafter, we measure all results with $T/t_\ell = 0.4$ (corresponding to $T/\tau_r \approx 8$), which guarantees that effect of hysteresis on our estimation of τ_ℓ and τ_c is negligible.

5.5.2 Bubble lifetime, τ_ℓ

We now seek the expected bubble lifetime, τ_ℓ. Figure 5-9 shows our measurements of $C_\Omega(We_B)$ and the fit to the model by Martínez-Bazán *et al.* (1999*a*), repeated here for clarity:

$$C_\Omega(We_B) = C_{\Omega,\infty}\sqrt{1 - We_H/We_B}\,. \tag{5.6}$$

Non-linear least-squares regression gives $We_H = 6.9$ and $C_{\Omega,\infty} = 1.4$, with a coefficient of determination $R^2 = 0.890$. For $We_B \gg We_H$, this gives a bubble lifetime

$$\tau_\ell = \left(C_{\Omega,\infty}\right)^{-1} \varepsilon^{-1/3}a^{2/3}\,. \tag{5.39}$$

Compared to previous work, our value for the Hinze scale, $We_H = 6.9$, is similar to $We_H = 4.7$ measured by Martínez-Bazán *et al.* (1999*a*) and $We_H = 2.7 - 7.8$ by Risso & Fabre (1998). However, we obtain $C_{\Omega,\infty} = 1.4$, greater than $C_{\Omega,\infty} = 0.42$ measured by Martínez-Bazán *et al.* (1999*a*) and $C_{\Omega,\infty} = 0.95$ from HIT simulations by Rivière *et al.* (2021). This means we predict a shorter τ_ℓ than previous work. An important distinction between our fragmentation measurements and previous experimental and numerical measurements is that we measure bubbles that have been formed as the daughters of previous fragmentation, so the bubbles are already distorted by fragmentation.

To confirm that our larger value of $C_{\Omega,\infty}$ is due to the way bubbles are created, we

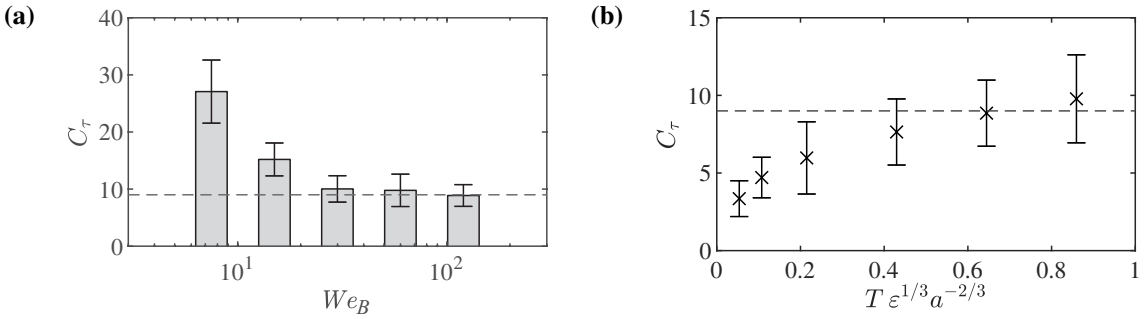


Figure 5-10: Convergence constant C_τ (a) as functions of We_B, measured using $T/t_\ell = 0.4$ and (b) as a function of T for $We_B = 50$–71. Error bars indicate 95% C.I.. The estimated large-We_B value of $C_\tau = 9$ (– – –) is included for reference.

repeat the measurement of fragmentation statistics, but this time over an earlier time in our simulation, $0 < t/t_\ell < 1$, when (as opposed to the later time $1 < t/t_\ell < 3$) many parent bubbles which started spherical have not yet fragmented. When we measure this earlier time range (denoted by $(\cdot)_{t<t_\ell}$), we obtain a similar $(We_H)_{t<t_\ell} = 7.0$ but an appreciably smaller $(C_{\Omega,\infty})_{t<t_\ell} = 0.88$ (with $R^2 = 0.974$). As our interest is bubbles *within* fragmentation cascades, our value of $C_{\Omega,\infty} \approx 1.4$ is more relevant for bubbles formed by fragmentation.

Finally, we note that $1/C_{\Omega,\infty}$ is an order of magnitude larger than C_r, meaning that $\tau_\ell \gg \tau_r$. As discussed in section 5.2.2, this confirms that no-hysteresis assumption, a key assumption for PBE, is reasonable when modeling fragmentation cascades.

5.5.3 Convergence time, τ_c

We now seek the convergence time, τ_c. ELA gives us direct access to measure the (T-dependent) size distribution of daughter bubbles (shown in Appendix E), from which $\mathbb{E}\{[v_R(T^*)]^n\}$ can easily be calculated. With (5.25), this gives us the time-averaged speed $\mathbb{E}\{\langle s\rangle_{T^*}\}$. If (5.31) is satisfied and $T \gg \tau_r$ so we can neglect hysteresis, we expect $\mathbb{E}\{\langle s\rangle_T\}$ to give a T-independent measurement of $C_\tau = 1/\bar{s}$. Figure 5-10a shows the value of C_τ we obtain over a range of We_B using $T/t_\ell = 0.4$ ($T/\tau_r \approx 8$). Recall that the model we developed in section 5.3, as a result of large-We_B assumptions, predicts a constant C_τ. We find that this is accurate for $We \gg We_H$, or more specifically $We > 30$. In this large-We_B regime, we measure $C_\tau \approx 9$.

To validate that our measurement is T-independent, for $We_B = 50$–71 we also measure C_τ using a range of T (figure 5-10b). As expected, for $T \lesssim \tau_r$ we see a dependence on T due to hysteresis, but for $T \gg \tau_r$ C_τ is independent of T. In addition to the τ_r lower bound, (5.31) implies an upper bound on the choice of T. However, for all the T we consider, (5.31) (with $C_\tau = 9$) gives $T \ll T_U$ for $We_B > 30$. This means the Hinze scale driven upper bound on T-independence described in §5.3.4 is not relevant to these measurements.

5.6 Discussion

We now examine how our quantification of the three fundamental fragmentation timescales informs the general study of fragmentation.

5.6.1 τ_r informs choice of measurement interval

For τ_r, our results suggest that the physical mechanism for the decay of hysteresis with bubble age is independent of surface tension for $We_B > We_H$ and that τ_r scales like τ_ℓ. The respective scaling constants we estimate from DNS of HIT differ by an order of magnitude ($C_r \ll 1/C_{\Omega,\infty}$), suggesting that $\tau_r \ll \tau_\ell$ is always true for $We_B > We_H$. Although the physical mechanism for the decay of hysteresis is still unclear, this shows that hysteresis can be assumed negligible when modeling fragmentation, validating an essential assumption of PBE. More practically, knowledge of τ_r also informs the choice of measurement interval in experiments and simulations. $T \gg \tau_r$ makes the effect of hysteresis on measurements negligible, ensuring that the measured fragmentation statistics are compatible with PBE. Applied to ELA, this provides a lower bound on the choice of snapshot interval, $\Delta t_s \gg \tau_r$.

5.6.2 τ_c provides a new constraint on fragmentation models

The insight that the convergence time τ_c provides into the evolution of the bubble size distribution in fragmentation-dominated bubbly flows has been discussed by Qi *et al.* (2020) and Deike *et al.* (2016), and we have now quantified τ_c directly. now quantified τ_c directly. For large We_B where the effect of surface tension on fragmentation rates is negligible, we find

$$\tau_c = C_\tau \varepsilon^{-1/3} {a_{max}}^{2/3} \left[1 - (We_{max}/We_H)^{-2/5}\right] , \tag{5.40}$$

where We_{max} is the bubble Weber number We_B of the largest bubble in the cascade (radius a_{max}) and we estimate $C_\tau \approx 9$ and $We_H \approx 6.9$ from DNS. In addition, as we can now express τ_c in terms of realistic fragmentation statistics for $We > 30$, τ_c also informs large-We fragmentation models. Inspired by (5.4), we rearrange (5.30) to provide a new bound on a moment of the daughter-size distribution f_V^*:

$$\bar{m}(a') \int_0^1 {v^*}^{11/9} f_V^*(v^*; a')\, \mathrm{d}v^* = 1 - \left(C_\tau C_{\Omega,\infty}\right)^{-1} , \tag{5.41}$$

where our estimations of $C_\tau \approx 9$ and $C_{\Omega,\infty} = 1.4$ from DNS give 0.92 for the right side of (5.41), independent of parent bubble radius a'. For a physical interpretation, (5.4) bounds the relationship between daughter-size distributions and $\bar{m}$ to guarantee volume conservation, while (for $We_B > 30$) (5.41) bounds the relationship to match the empirical value of τ_c.

Many existing fragmentation models assume binary breakup ($\bar{m} = 2$). To evaluate how well these meet (5.41), we focus on the proposed daughter-size distributions through C_f, which includes the integral in (5.41). With $\bar{m} = 2$, $C_\tau \approx 9$, and $C_{\Omega,\infty} = 1.4$, (5.30) gives $C_f \approx 1.8$. Because C_f indicates how much longer τ_c is compared to the case of identical fragmentation, this shows that τ_c is 1.8 times longer for fragmentation in HIT than what would be predicted if one assumes identical binary-fragmentation. Comparing to more

realistic binary daughter-distributions (see B–D in table 5-1), we see good agreement with the distribution proposed by Martínez-Bazán *et al.* (2010). We also compare our C_f to the binary daughter-distribution model by Qi *et al.* (2020),

$$f_V(v^*) = \omega \left[\frac{1-\cos(2\pi v^*)}{\int_0^1 [1-\cos(2\pi v^*)]\, \mathrm{d}v^*} \right] + (1-\omega) \left[\frac{(v^*)^{-4/3}+(1-v^*)^{-4/3}}{\int_{0.02}^{0.98} \left[(v^*)^{-4/3}+(1-v^*)^{-4/3}\right] \mathrm{d}v^*} \right], \tag{5.42}$$

who set the tuning parameter $\omega = 0.3$ to match experimental measurements of τ_c. For this daughter-distribution model, (5.30) gives $C_f = 1.741$, in very good agreement with our value of $C_f \approx 1.8$. Although we assume $\bar{m} = 2$ here for illustration, this analysis is applicable to any $\bar{m}$. Rather than attempting to compare the details of disparate fragmentation models, relating τ_c to the fragmentation statistics specified by these models allows us to directly compare the physical predictions each model makes regarding the evolution of the bubble size distribution through a simple scalar quantity.

5.7 Conclusion

As discussed in Chapter 1, a well-known equilibrium for the bubble size distribution is $N(a) \propto a^{-10/3}$, which is the result of fragmentation cascades at moderate and large We_B (Garrett *et al.*, 2000; Gaylo *et al.*, 2021). More generally, there is an interest in a statistical model of fragmentation for use in the PBE (1.2). In this chapter we describe three fundamental timescales characterizing the statistics of fragmentation and the resulting fragmentation cascade. These timescales directly support statistical modeling of fragmentation, and, although our focus here is on statistical descriptions of fragmentation, the results here also help inform future mechanistic study of fragmentation.

One fundamental timescale is the relaxation time τ_r which characterizes the time after fragmentation over which hysteresis cannot be neglected. From DNS measurements, we provide an empirical definition of τ_r based on when measured fragmentation rates become independent of the measurement interval T. We find that $\tau_r = C_r \varepsilon^{-1/3} a^{2/3}$, where $C_r \approx 0.11$ independent of moderate/large We. This We_B-independence suggests the physical mechanism causing τ_r at these We_B is unrelated to surface tension. Although understanding hysteresis and its decay is an area of future work, by providing τ_r we identify the timescales over which hysteresis can be neglected.

A second fundamental timescale is the expected lifetime τ_ℓ of a bubble from formation by fragmentation to further fragmentation. For $\tau_\ell \gg \tau_r$, $\tau_\ell = [C_\Omega(We_B)]^{-1} \varepsilon^{-1/3} a^{2/3}$ is the inverse of the fragmentation rate. Fitting our DNS results for bubbles within the fragmentation cascade to the square-root model of We_B-dependence by Martínez-Bazán *et al.* (1999*a*) (eq. (5.6)), we find the Hinze-scale $We_H \approx 6.9$, in agreement with previous experiments, but measure a smaller τ_ℓ corresponding to a higher scaling constant (at large We) $C_{\Omega,\infty} \approx 1.4$ (compared to $C_{\Omega,\infty} \approx 0.42$ reported by Martínez-Bazán *et al.* (1999*a*)). We show that this higher value of $C_{\Omega,\infty}$ is related to formation of the bubbles by a fragmentation cascade. For modeling fragmentation cascades, this higher $C_{\Omega,\infty}$ is likely more relevant. In either case, we find $\tau_r \ll \tau_\ell$ for all We, validating the use of the no-hysteresis assumption in modeling fragmentation.

Finally, we consider the fundamental timescale $\tau_c = C_\tau[1-(We_{max}/We_H)^{-2/5}]\varepsilon^{-1/3}a_{max}^{2/3}$, which measures the time for a Lagrangian air particle to go from the largest bubble to the Hinze scale. This also characterizes the time for fragmentation cascades to reach the $N(a) \propto a^{-10/3}$ equilibrium. For large We_B, we derive τ_c based on the (constant) expected speed $\bar{s}$ at which a Lagrangian air particle moves through the cascade. We show that, $C_\tau = 1/\bar{s}$ and can thus be measured independent of T. This result is valid for $\tau_r \ll T \ll \tau_c$, which provides a bound on the choice of T in experiments and simulations. The T-independence of C_τ is confirmed by DNS measurements, which give $C_\tau \approx 9$ for $We_B > 30$, which agrees well with the values obtained from the fragmentation model of Martínez-Bazán *et al.* (2010) and an experimentally-constrained fragmentation model of Qi *et al.* (2020). The relationship between C_τ and fragmentation statistics in PBE provides new constraints on these statistics at large We_B, limiting the possible forms of fragmentation models.

In the context of modeling the evolution of bubble populations in air entraining FST, τ_c is particularly relevant. If the bubble population is dominated by the effects of the fragmentation cascade, then τ_c gives the characteristic time for the bubble population to converge to $N(a) \propto a^{-10/3}$. By quantifying C_τ, we can now obtain this convergence time τ_c based on the strength of turbulence (ε) and the largest bubble entrained by the turbulence (a_{max}). What remains then is to determine if bubble population is actually dominated by fragmentation. In Chapter 7 we will show it is not beneath air entraining FST. Evidence of this will include that we do not obtain $N(a) \propto a^{-10/3}$, despite considering bubble population evolution on timescales much larger than τ_c.

Chapter 6

Bubble Entrainment in Free-Surface Turbulence

In this chapter we focus on bubble entrainment, the $I(a)$ term in the population balance equation (1.2). The presence of air entrainment is of course the defining characteristic of air entraining free-surface turbulence, and, by a simple argument, entrainment is the most important mechanism because it is the original source of all the air beneath the free surface. In this chapter we consider entrainment by free-surface turbulence, specifically entrainment of large bubbles where surface tension effects are small. While Yu *et al.* (2020) made a prediction about large bubble entrainment size distribution, they could only measure the total bubble size distribution $N(a)$ and infer the entrainment size distribution, and it turns out that their prediction is incorrect. Previously the only measurements of entrainment were for ~ 100 events (Wei *et al.*, 2019). Using ELA (Gaylo *et al.* 2022; see also Chapter 3) with DNS of multiple free-surface flows across a range of scales, we obtain direct measurements of $\sim 60,000$ entrainment events, elucidating the entrainment size distribution $I(a)$.

Key results from this chapter are summarized in "Size distribution of large air bubbles entrained by strong free-surface turbulence" by Gaylo & Yue (2025).

6.1 Introduction

For the development of general statistical models of the evolution of bubble populations beneath air entraining free surfaces (i.e., through the PBE), an immediate challenge is that entrainment itself is one of the least understood classes of bubble evolution mechanisms. Previous work on entrainment tends to focus on entrainment by specific large-scale flow structures. One example is the entrapment of a cavity by a plunging breaking wave (Deike *et al.*, 2016; Chan *et al.*, 2021*a*; Gao *et al.*, 2021). Another example is entrainment by a plunging jet (Kiger & Duncan, 2012; Bertola *et al.*, 2018). While these features may be relevant to specific air entraining flows, Brocchini & Peregrine (2001*a*) note that a common feature in a large class of air entraining flows is strong turbulence beneath the free surface. Examples include the wake behind a vessel, spilling breakers, hydrologic jumps, upwelling of existing bubble populations, and flows in steep rivers or spillways. Motivated by the

potential broad applicability, in this chapter we seek to characterize air entrainment that is a direct result of this strong free-surface turbulence (FST).

As discussed in §1.1, one flow where FST entrainment is relevant is open channel flow, often studied in the context of hydraulic modeling of rivers and spillways (Falvey & Ervine, 1988; Chanson, 1996). The quantity and size of bubbles is necessary to predict how the dissolved oxygen content changes due to rapids or spillways (Gulliver & Rindels, 1993), so a model of bubble entrainment is needed. While some previous work considered bubble formation by the impact of droplets at the free surface, this mechanism is not believed to be energetic enough to explain the observed bubbles (Rein, 1998). Wei *et al.* (2019) use high speed photography to obtain time series of individual entrainment events. Their direct measurement of individual entrainment events is unlike the more typical approach of measuring the entire resulting bubble population (e.g. Chanson & Toombes, 2003). This gave Wei *et al.* (2019) new insight into the exact mechanism for entrainment, and they show that entrainment in open channel is the direct result of the interaction of turbulence with the free surface. That work shows FST entrainment is relevant to channel flow and provides what we believe to be the first direct experimental measurement of entrainment by FST.

An inherent challenge in direct experimental measurement of entrainment by FST is that the free surface is strongly distorted, making visual access difficult. In total, Wei *et al.* (2019) were only able to observe 108 entrainment events, which is insufficient to obtain detailed statistics on how entrainment scales with bubble size and turbulence strength. For this purpose, DNS is useful, as all features of the flow are immediately accessible. As discussed in Chapter 3, identifying individual mechanisms from this wealth of data is not trivial; however, ELA enables us to, for the first time, identify and measure individual entrainment events within DNS of complex bubbly flow. In this chapter we perform DNS of a canonical flow which isolates entrainment by FST from other entrainment mechanisms and use ELA to obtain a large $O(10^4)$ data set of FST entrainment events.

Based on the direct measurement of entrainment events in DNS, we can elucidate the scaling of the size distribution of FST entrainment with the strength of turbulence and gravity. This observed scaling is explained and supported by a simple mechanistic model. Although our results are based on a flow where FST is the only entrainment mechanism, we show that the same scaling closely describes entrainment size distributions observed in other, more complex free-surface flows where additional entrainment mechanisms are also present.

6.2 Characterizing air entrainment by FST

From the PBE (1.2), our interest is the entrainment size distribution $I(a)$, where $I(a)\delta a\delta t$ is defined to be the number of bubbles of radius $[a, a+\delta a]$ created at the free surface over time $[t, t+\delta t]$ in the region of interest. As discussed in Chapter 4, an entraining free surface can be thought of as a locally flat surface ($\bar{\eta}$) plus perturbations by turbulence, in which case it is natural to normalize $I(a)$ by A_{FS}, the area of the mean free surface in the region of interest. In §6.2.1 we perform dimensional analysis to obtain, given some basic properties of strong FST, the set of possible scalings of $I(a)/A_{FS}$. In §6.2.2 we describe a mechanistic model of FST entrainment which elucidates the scaling. In §6.3.2 this scaling is confirmed by direct measurement of $I(a)/A_{FS}$ in DNS.

6.2.1 Dimensional analysis

As introduced in Chapter 4, the critical parameter for FST is the strength of the near-surface turbulence relative to the restoring force of gravity, described by a turbulent Froude number (squared)

$$Fr_T^2 = \frac{\varepsilon}{u_{\text{rms}}\text{g}}\,, \tag{1.1}$$

where g is gravitational acceleration and the strength of near surface turbulence is described by the dissipation rate ε and the root-mean velocity fluctuations u_{rms}. In Chapter 4 we show that for strong FST ($Fr_T^2 > 0.1$) the turbulence will be nearly isotropic and follow the Kolmogorov $-5/3$ scaling. This means that within the inertial sub range the strength of turbulence is fully characterized by $\varepsilon = [L^2T^{-3}]$. After normalizing by the density of water ρ_w, the energy to entrain a bubble of radius $a = [L]$ is characterized by gravitational acceleration $\text{g} = [LT^{-2}]$ and surface tension $(\sigma/\rho_w) = [L^3T^{-2}]$. We will assume that constitutive property ratios, such as the density ratio ρ_a/ρ_w and the viscosity ratio μ_a/μ_w, are fixed for air and water. This gives that for strong FST $I(a)/A_{FS}$ (dimensions $[L^{-3}T^{-1}]$) only depends on

$$I(a)/A_{FS} = \mathcal{F}(\varepsilon, a, \text{g}, \sigma/\rho_w)\,. \tag{6.1}$$

By dimensional analysis, we find that the system (6.1) is described by three dimensionless parameters. While the choice of these parameters is not unique, it is useful to choose one and only one parameter which includes surface tension σ/ρ_w. We choose this parameter to be a Bond number $Bo \equiv \text{g}(2a)^2/(\sigma/\rho_w)$, which describes the ratio of gravitational to surface energy needed to form a bubble. The critical value $Bo = 1$ corresponds to the capillary scale,

$$a_c \equiv 0.5\sqrt{\sigma/\text{g}\rho_w}\,. \tag{6.2}$$

For air water ($\sigma/\rho_w \approx 7.03 \times 10^{-5}$ m^3/s^2) on Earth ($\text{g} \approx 9.81$ m/s^2), the capillary scale is $a_c \approx 1.3$mm. For large bubbles with radii $a \gg a_c$ (i.e., $Bo \gg 1$), we expect the effects of surface tension on entrainment to be negligible, so we drop Bo as a parameter and (6.1) becomes,

$$I(a)/A_{FS} = \mathcal{F}(\varepsilon, a, \text{g}) \quad \text{for } a \gg a_c\,. \tag{6.3}$$

By dimensional analysis, the system (6.3) now has only two dimensionless parameters, say Π_1 and Π_2. We choose ε and $(2a)$ as the repeating variables and obtain

$$\Pi_1 = \frac{I(a)/A_{FS}}{\varepsilon^{1/3}(2a)^{-11/3}}\,; \qquad \Pi_2 = \left(\frac{\varepsilon^{2/3}(2a)^{-1/3}}{\text{g}}\right)^{1/2}\,. \tag{6.4a, b}$$

Far from critical values of Bo or Fr_T^2, we assume the underlying mechanisms are scale invariant. This assumption implies a power law relationship, $\Pi_1 = C_I{\Pi_2}^{\alpha}$, where α describes the scaling and C_I is a scaling constant. Thus, dimensional consistency and scale invariance gives that

$$I(a)/A_{FS} = C_I\, g^{[-\alpha/2]}\, \varepsilon^{[\alpha/3+1/3]}\, (2a)^{[-\alpha/6-11/3]} \quad \text{for } a \gg a_c\,, \tag{6.5}$$

where α and C_I are to be determined.

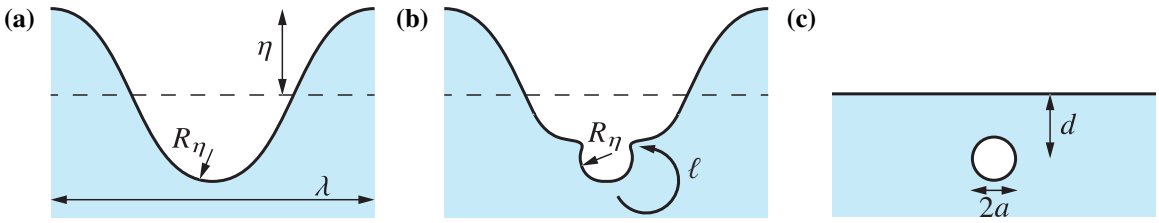


Figure 6-1: Illustration of the three stages of air entrainment: (a) A free-surface deformation exists with amplitude η, wavelength λ, and minimum radius of curvature R_η. (b) It interacts with an eddy of size ℓ. (c) a bubble of radius $a \sim R_\eta$ is formed at depth d.

For FST it is useful to normalize by near-surface turbulence scales $L_T = u_{\text{rms}}^3/\varepsilon$ and $T_T = u_{\text{rms}}^2/\varepsilon$. Thus, (6.5) can be nondimensionalized

$$[I(a)/A_{FS}]^* = C_I \ (Fr_T)^{[\alpha]} \ (2a^*)^{[-\alpha/6-11/3]} \quad \text{for } Bo \gg 1\,, \tag{6.6}$$

where the nondimensionalized bubble radius is $a^* \equiv a/L_T$ and the nondimensionalized entrainment size distribution is $[I(a)/A_{FS}]^* \equiv [I(a)/A_{FS}]{L_T}^3 T_T$. In this nondimensional form, it is clear that our choice of Πs was such that the value α describes how large-bubble air entrainment by FST scales with turbulent Froude number.[1]

6.2.2 Mechanistic model for large-bubble air entrainment by FST

We now propose a mechanistic model of entrainment which obtains the value of α and predicts the scaling we observe in DNS (see §6.3.2). As a simplified mechanistic model we consider three stages for entrainment, illustrated in figure 6-1:

1. The interaction of the free surface with the turbulence beneath creates a deformation of wavelength λ height η with a minimum radius of curvature R_η.
2. Interaction of this surface deformation with a turbulent eddy of size ℓ causes the deformation to collapse.
3. The collapse entrains a single bubble of radius a at depth d.

Central to our mechanistic model is that during the collapse of the surface the minimum radius does not change significantly, such that the radius a of the bubble formed scales with the minimum radius of curvature R_η of the initial surface deformation,

$$a \sim R_\eta\,. \tag{6.7}$$

This insight is consistent with observation by Wei *et al.* (2019) that as air entraining surface deformations collapse the minimum radius of curvature approaches an asymptotic value r_c which increases with the size of bubble eventually obtained, and is consistent with our qualitative observations from DNS (figure 6-2).

[1]While the choice of Πs is not unique, any other valid Πs would simply replace α with some $\alpha' = C\alpha$.

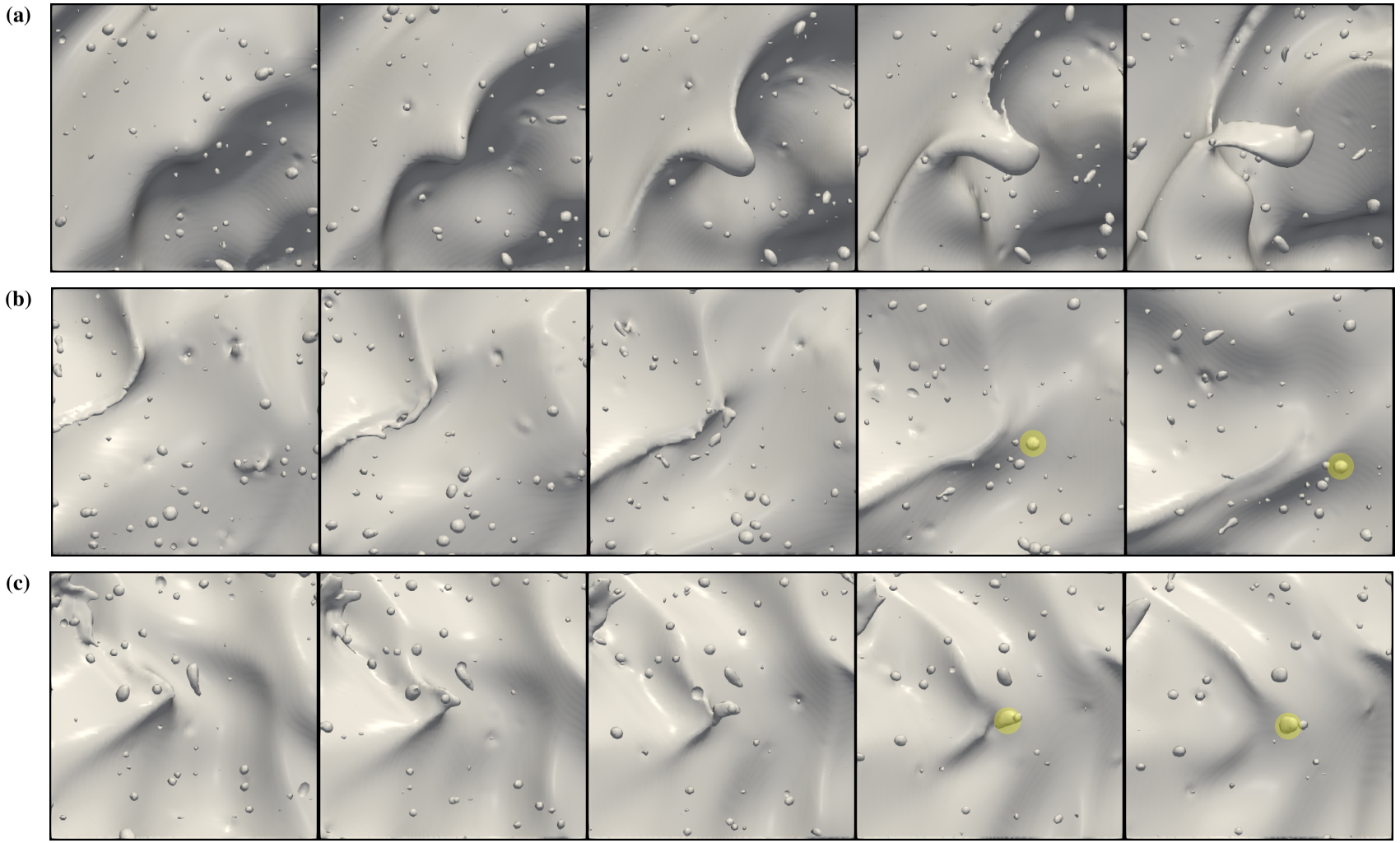


Figure 6-2: Three examples of air entrainment events, viewed from beneath the free surface, at $u_{\text{rms}} = 0.220$ m/s and $\varepsilon = 0.291$ W/kg (see table 6-2). Each frame is 62mm × 62mm and separated in time by 17ms. A large surface deformation collapses into a single bubble, highlighted in yellow for (b) and (c). The minimum radius of curvature does not change significantly, leading to $a \sim R_\eta$.

Turbulence interacting with the free surface causes inter-scale energy transfer

We first seek to determine which turbulent scales are responsible for entraining a bubble based on the insight that $a \sim R_\eta$, starting with an expression for R_η. As confirmed in Chapter 4, near-surface turbulence is isotropic for $Fr_T^2 > 0.1$, allowing application of the Kolmogorov energy cascade. Within the inertial range, the power spectrum of pressure fluctuations is given by $E_{pp}(\kappa) = 2.97\,\rho_w^2 \varepsilon^{4/3} \kappa^{-7/3}$ (George *et al.*, 1984). For $Bo \gg 1$ such that surface tension is negligible, the linearized relationship between pressure fluctuations and surface elevation gives $p \sim \rho_w \mathrm{g} \eta$. This linear relationship means the surface elevation spectrum $E_{\eta\eta}$ can be related to the pressure spectrum by $E_{pp} \sim (\rho_w \mathrm{g})^2 E_{\eta\eta}$, which gives

$$E_{\eta\eta}(\kappa) \sim 2.97\,\mathrm{g}^{-2} \varepsilon^{4/3} \kappa^{-7/3}\,. \tag{6.8}$$

For a surface disturbance of wave number $\kappa = 2\pi/\lambda$, the characteristic amplitude (squared) is $\eta^2 \propto E_{\eta\eta}(\kappa)\delta\kappa$ where $\delta\kappa \sim \kappa$. Taking the square root,

$$\eta \sim 1.72\, g^{-1} \varepsilon^{2/3} \kappa^{-2/3}\,. \tag{6.9}$$

For a sinusoidal disturbance, the minimum radius of curvature is $R_\eta = \kappa^{-2}\eta^{-1}$. Other disturbance geometries, such as Gaussian (Wei *et al.*, 2019), follow the same scaling $R_\eta \propto \kappa^{-2}\eta^{-1}$. Thus, for general geometries we have

$$R_\eta \propto \mathrm{g}\, \varepsilon^{-2/3} \kappa^{-4/3}\,. \tag{6.10}$$

Applying $a \sim R_\eta$ and rearranging (6.10), we determine the turbulence scales κ that are responsible (through interaction with the free surface) for generating bubbles of radius a,

$$\kappa \propto \mathrm{g}^{3/4} \varepsilon^{-1/2} a^{-3/4}\,. \tag{6.11}$$

This is distinct from previous energy-based mechanistic models of air entrainment (Rein, 1998; Yu *et al.*, 2020), which assume bubbles of radius a are entrained by turbulence eddies of similar scales, i.e., $\kappa \sim a^{-1}$. While those models are dimensionally consistent, they do not capture the inter-scale energy transfer we observe with (6.11).

Before moving on, we take a moment to analyze how the inter-scale energy transfer predicted by (6.11) introduced a significant Froude number effect that was missing from previous models. On the local scale κ of the turbulent eddy, we can define an eddy Froude number (squared) $Fr_\kappa^2 \equiv \kappa^{-1/3} \varepsilon^{2/3} \mathrm{g}^{-1}$, and (6.11) can be written as

$$a\kappa \propto Fr_\kappa^{-2}\,. \tag{6.12}$$

This is opposed to $a\kappa \sim$ constant from previous models (Rein, 1998; Yu *et al.*, 2020). To interpret the Froude-dependence of this inter-scale energy transfer, consider a fixed g and fixed eddies of wave number κ. Increasing ε increases η, but the steeper free surface leads to a smaller volume of air trapped when the deformation collapses, and thus a smaller radius a for the entrained bubble. Smaller bubbles take less energy to overcome gravity, and the energy of the eddies increases with ε. The combination of these effects leads to a very Froude number sensitive entrainment mechanism.

Energy balance

To obtain the scaling of the entrainment distribution, we now describe the balance between the energy available in the free surface to the work necessary to entrain bubbles. Starting with the energy available in the free surface, for a small range of wave numbers $[\kappa, \kappa + \delta\kappa]$, the potential energy associated with the surface deformations (per free surface area A_{FS}) is

$$PE_{\text{waves}}/A_{FS} = \rho_w \mathrm{g}\, E_{\eta\eta}(\kappa)\, \delta\kappa\,. \tag{6.13}$$

Applying the wave spectrum (6.8),

$$PE_{\text{waves}}/A_{FS} \propto \rho_w \mathrm{g}^{-1} \varepsilon^{4/3} \kappa^{-7/3}\, \delta\kappa\,. \tag{6.14}$$

We now consider the energy needed to entrain bubbles. The rate at which bubbles of radius $[a, a + \delta a]$ are entrained is $I(a)\delta a$. For $Bo \gg 1$ where surface energy is negligible, the potential energy associated with each bubble of radius a entrained to depth is d (see figure 6-1c) is $\rho_w g v d$ where the volume of the bubble $v = (4\pi/3)\, a^3$. Following Yu *et al.* (2020) we assume bubbles are entrained to a depth $d \sim 2a$. Multiplying the rate at which bubbles are entrained by the energy needed for each bubble, we obtain the necessary work done, W_{ent}, to entrain bubbles of radius $[a, a + \delta a]$,

$$W_{\text{ent}} = I(a)\ (8\pi/3)\, \rho_w g a^4\, \delta a\,. \tag{6.15}$$

Figure 6-1b illustrates that for the surface deformation to collapse into a bubble, some perturbation is necessary. We argue that this perturbation will come from eddies of size $\ell \sim R_\eta$ (or equivalently $\ell \sim a$) and T_{ent} is the associated eddy turnover time. From the Kolmogorov energy cascade,

$$T_{\text{ent}} \sim \varepsilon^{-1/3} a^{2/3}\,. \tag{6.16}$$

To balance the energy available in waves to the energy to entrain bubbles, we set $W_{\text{ent}} T_{\text{ent}} = PE_{\text{waves}}$ and obtain

$$I(a)/A_{FS} \propto g^{-2} \varepsilon^{5/3} \kappa^{-7/3} a^{-14/3}\, (\delta\kappa/\delta a)\ . \tag{6.17}$$

We now apply the inter-scale energy transport which related κ to a. (6.11) gives $\kappa^{-7/3} \propto \mathrm{g}^{-7/4} \varepsilon^{7/6} a^{7/4}$. The derivative of (6.11) gives $\delta\kappa/\delta a \propto \mathrm{g}^{3/4} \varepsilon^{-1/2} a^{-7/4}$. Finally,

$$I(a)/A_{FS} \propto g^{-3} \varepsilon^{7/3} a^{-14/3}\,. \tag{6.18}$$

This corresponds to $\alpha = 6$ in (6.5), and FST entrainment scaling with Fr_T^6 in (6.6).

Limits on largest entrained bubble, a_{max}

The mechanistic model we develop is based on $Bo \gg 1$, leading to the lower limit $a \gg a_c$ for applicability of (6.18). Here we briefly address the upper limit on applicability of (6.18). Our interest is some maximum entrained bubble $a_{\max}$ such that $I(a > a_{\max}) \approx 0$. The previous mechanistic argument gives two possible limits. First, in our mechanism the collapse of the deformation is caused by an eddy of scale $\ell \sim a$. The largest turbulent eddies

Fr^2	We	T_{sim}	u_{rms}	ε	L_T	Fr_T^2	Re_T	We_T	N_I	$a_{\max}/L_T$
0.3	∞	256	0.27	0.027	0.76	0.03	41	–	820	0.20
0.6	∞	128	0.29	0.031	0.81	0.06	47	–	1908	0.30
0.9	∞	128	0.29	0.033	0.78	0.10	46	–	7920	0.38
1.2	∞	128	0.26	0.025	0.74	0.11	39	–	9308	0.34
1.8	∞	128	0.28	0.021	0.99	0.14	54	–	24170	0.27
1.2	400	127	0.31	0.032	0.93	0.12	58	36	526	0.29
1.2	200	128	0.34	0.039	1.06	0.13	73	25	433	0.23
1.2	100	128	0.30	0.044	0.60	0.18	35	5.2	321	0.62

Table 6-1: List of forced FST simulations used for entrainment measurements. Turbulence properties u_{rms} and ε are measured using (6.21). The characteristic length scale $L_T = u_{\text{rms}}^3/\varepsilon$ is used to calculate the near-surface turbulent Froude number $Fr_T^2 = (u_{\text{rms}}^2/L_T)Fr^2$, turbulent Reynolds number $Re_T = (u_{\text{rms}}L_T)Re$, and turbulent Weber number $We_T = (u_{\text{rms}}^2 L_T)We$. N_I is the number of (resolved) entrainment events recorded and $a_{\max}$ is the radius of the largest entrained bubble observed.

are on the order of the characteristic length scale $L_T = u_{\text{rms}}^3/\varepsilon$, i.e., $\ell \lesssim L_T$. This implies a maximum entrained bubble size

$$a_{\max}/L_T \propto \text{constant}\,. \tag{6.19}$$

Second, our mechanism shows an inter-scale energy transfer, where the energy to form a bubble comes from a different scale of turbulence. In terms of wave number, the largest scales of turbulence are $\kappa_{\min} \sim 2\pi/L_T$. By (6.11), this gives $a_{\max} \propto g\varepsilon^{-2/3}L_T^{4/3}$. Rearranging,

$$a_{\max}/L_T \propto Fr_T^{-2}\,. \tag{6.20}$$

Which of these two conditions is more restrictive will depend on Fr_T^2.

6.3 Quantifying entrainment by free-surface turbulence

To study entrainment by FST, we first seek a flow which isolates FST entrainment from any other entrainment mechanism. This is obtained with the forced FST simulations used in Chapter 4, where isotropic turbulence is continuously forced deep beneath the surface to obtain statistically steady turbulence and bubble population at the free surface. In section 6.4 and section 6.5 we will investigate how the entrainment distribution we elucidate here applies to more general air entraining flows.

6.3.1 Simulation setup

We use the same forced FST simulation setup described §4.3, with many results coming from the same exact simulations. Table 6-1 provides a summary of the forced FST simulations we use for this chapter.

So that our results can be more easily compared to previous work (Yu *et al.*, 2019; Gaylo *et al.*, 2024), in this chapter the method we use to obtain near-surface turbulence values u_{rms} and ε differs from Chapter 4. We perform averaging only in the water phase (where the color function $c = 1$),

$$\langle \cdot \rangle_\delta \equiv \frac{\iiiint \cdot \, c \, \mathrm{d}\mathbf{x} \, \mathrm{d}t}{\iiiint c \, \mathrm{d}\mathbf{x} \, \mathrm{d}t} \quad \text{for } z > -\delta \text{ and } t \in [t_0, t_0 + T_{\text{sim}}] \, . \tag{6.21}$$

Previous work used $\delta / L_T \approx 0.5$ to define "near-surface." Seeking to match this definition, we choose $\delta = 0.3$ based on *a priori* estimates of L_T. This averaging defines a near-surface $u_{\text{rms}} = \sqrt{\langle \mathbf{u} \cdot \mathbf{u} \rangle_\delta / 3}$ and a dissipation rate $\varepsilon = \langle \boldsymbol{\tau} : \boldsymbol{\nabla} \mathbf{u} \rangle_\delta$. Chapter 4 shows that measuring these values at $(z - \bar{\eta})/\delta_s = -0.5$ is based on a more robust definition of "near-surface"; however comparing table 6-1 and table 4-1 we see using (6.21) instead does not significantly affect the calculated values of Fr_T^2.

During the steady state portion of the simulations ($t \in [t_0, t_0 + T_{\text{sim}}]$), we use ELA (Gaylo *et al.* 2022; see also Chapter 3) to identify and measure entrainment events. As described in §3.2.2, by identifying the label which corresponds to the bulk air above the free surface (the 'sky') at t^n and tracking this air up to $t^{n+1} = t^n + \Delta t_s$, we obtain direct measurement of entrainment events over the snapshot interval Δt_s (see figure 3-1c). In general, $\Delta t_s \, \Omega(a_{\text{res}}) \geq 0.1$ avoids spurious events (Chan *et al.* 2021*a*; Gaylo *et al.* 2022; see also Chapter 3). Here we use $\Delta t_s = 0.16$.

For each individual entrainment event, ELA gives us the entrained volume v, from which we calculate the effective radius $a \equiv (3v/4\pi)^{1/3}$. To ensure the relevant physics are resolved by the numerical grid, we only report bubbles $a > a_{\text{res}}$, where $a_{\text{res}} = 1.5\Delta$ for simulations where surface tension is not modeled (Yu *et al.* 2019; Gaylo *et al.* 2024, see also Appendix G), and $a_{\text{res}} = 3.5\Delta$ for simulations where surface tension is modeled (Yu *et al.*, 2020). The total number of resolved entrainment events N_I as well as the largest single entrained bubble a_{max} are shown in table 6-1. For the $Fr^2 = 0.3$ run, to obtain a sufficient number of resolve entrainment events, we increase T_{sim} from 128 (used in Chapter 4) to 256.

6.3.2 Entrainment size distribution, $I(a)$

From each simulation, we have a list of all entrainment events and the associated bubble radius. To obtain the entrainment size distribution, we first bin the entrainment events by radius. Starting with $\sqrt{N_I}$ bins evenly spaced over $\log(a)$, we identify the bin with the smallest number of events and split it between the two neighboring bins, combining three bins into two. This is done iteratively until there are at least 15 events per bin. Normalizing by the width of each bin, T_{sim}, and $A_{FS} = 4\pi^2$ gives $I(a)/A_{FS}$.

Figure 6-3 shows the entrainment size distribution for simulations without surface tension (see table 6-1). We first perform regression to obtain α and C_I. Because (6.6) is based on the assumption of isotropic near-surface turbulence, we initially consider the results from only the three strong FST ($Fr_T^2 \geq 0.1$) simulations. Least-squares regression in log-log space (after binning, $n = 226$ data points) gives 95% confidence intervals $\alpha \in [5.6, 6.1]$ and $C_I \in [2.65, 3.87]$. The mechanistic model we derive in §6.2.2 gives $\alpha = 6$ consistent with the regression, and if we fix $\alpha = 6$, the regression gives the 95% confidence interval

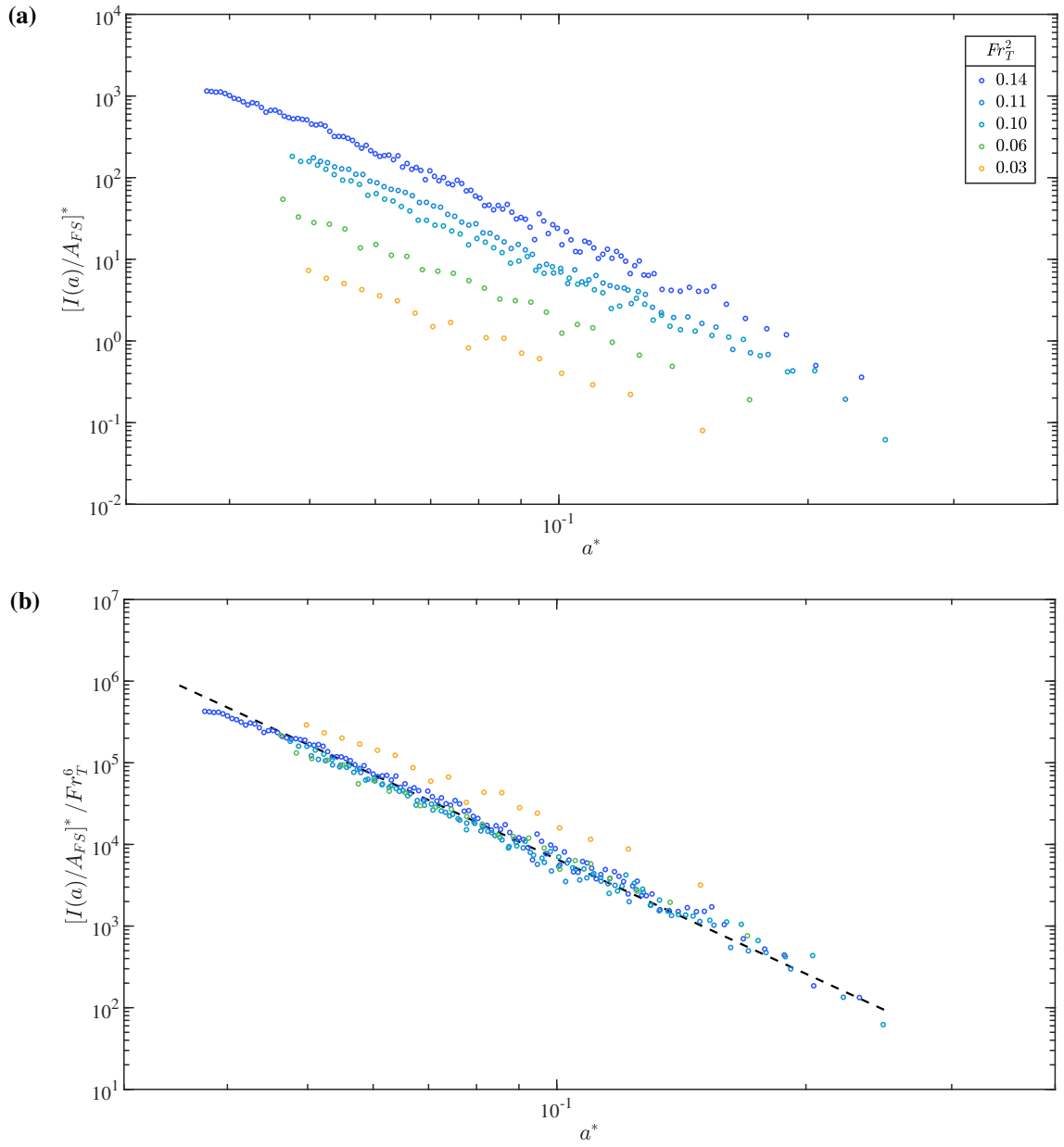


Figure 6-3: Entrainment size distribution in forced FST (a) as measured; and (b) normalized by Fr_T^6 for different turbulent Froude number $Fr_T^2 = \varepsilon/u_{\text{rms}}\text{g}$. In (b), (– – –) shows (6.6) with $\alpha = 6$ and $C_I = 3.62$ ($R^2 = 0.990$ excluding $Fr_T^2 = 0.03$). Recall $a^* = a\,\varepsilon\,u_{\text{rms}}^{-3}$ and $[I(a)/A_{FS}]^* = [I(a)/A_{FS}]\,u_{\text{rms}}^{11}\,\varepsilon^{-4}$.

Fr_T^2	We_T	u_{rms} m/s	ε W/kg	L_T m	a_{res} mm	a_H mm	$a_{\max}$ mm
0.12	36	0.234	0.282	0.046	4.2	4.5	13.1
0.13	25	0.220	0.291	0.037	3.0	4.4	8.4
0.18	5.2	0.159	0.276	0.015	2.1	4.5	9.0

Table 6-2: Dimensional values for forced FST simulations including surface tension. a_{res} is the smallest bubble resolved by the grid. The Hinze scale a_H is calculated using (5.1) with $We_H = 4.7$ (Martínez-Bazán *et al.*, 1999*a*).

$C_I = 3.62 \pm 0.10$. Figure 6-3b shows the collapse of the entrainment size distribution when normalized by Fr_T^6 for all but the lowest $Fr_T^2 = 0.03$, where the turbulence deviates furthest from isotropy. Even including $Fr_T^2 = 0.06 < 0.1$ where turbulence is weakly anisotropic, our model shows a very strong agreement ($R^2 = 0.990$) with the numerical results. The strong observed correlation with DNS confirms that (6.5) describes the scaling of FST entrainment for large bubbles ($a \gg a_c$) where surface tension is negligible. These results also show that $\alpha = 6$, so, consistent with (6.18) derived from our mechanistic model, we have

$$I(a)/A_{FS} = C_I \, \mathrm{g}^{-3} \, \varepsilon^{7/3} \, (2a)^{-14/3} \quad \text{for } a \gg a_c \,, \tag{6.22}$$

where $C_I \approx 3.62$ for this limiting case of negligible surface tension.

We see that (6.22) applies up to bubbles as large as $a_{\max}/L_T \approx 0.3$, above which there is no entrainment (see Table 6-1). This constant $a_{\max}/L_T$ is what we expect from (6.19), suggesting that interaction with a similar sized eddy is the cause of the limit on the largest entrained bubble size in the Fr_T^2 regime of these simulations. It is not a surprise that (6.19) is the limit rather than (6.20), as $1 \ll Fr_T^{-2}$ for all simulations.

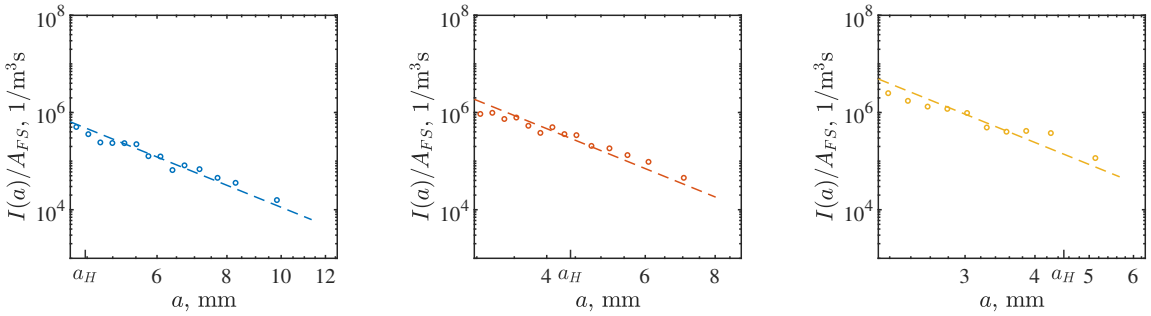
6.3.3 Effect of weak surface tension

So far, we have considered the limiting case of negligible surface tension. We now consider the effects of finite surface tension on FST entrainment. The turbulent Weber number,

$$We_T \equiv \frac{u_{\text{rms}}^5}{\varepsilon(\sigma/\rho_w)} \,. \tag{6.23}$$

characterizes the strength of near-surface turbulence relative to surface tension. Holding $Fr_T^2 \approx 0.11$ constant (strong FST), we perform DNS of 3 different $We_T \gg 1$. While the computational limits of DNS prevent us from directly matching viscosity, for each We_T we can match $\sigma = 72$ mN/m, $\rho_w = 1024$ kg/m^3, and g = 9.81 m/s^2 for air-water on Earth, which we can use to dimensionalize our results (see Table 6-2). As it is independent of turbulence levels, the capillary scale is $a_c = 1.3$ mm for all simulations.

As discussed in Chapter 5, it is well known that surface tension prevents fragmentation of bubbles smaller than the Hinze scale (Hinze, 1955), defined by (5.1). For bubble populations dominated by fragmentation, this causes a change in the power-law slope of $N(a)$ at a_H



(a) $We_T = 36$ **(b)** $We_T = 25$ **(c)** $We_T = 5.2$

Figure 6-4: Entrainment size distribution (per unit-free surface area) for different We_T. In each, (– – –) shows the fit to (6.22) with the corresponding value of C_I and the Hinze scale a_H is indicated.

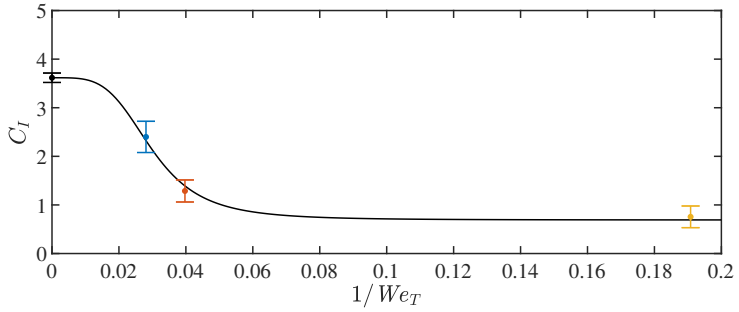


Figure 6-5: Value of C_I obtained through regression as a function of We_T, with error bars indicating the 95% confidence intervals. (——) shows the empirical fit, (6.24).

(Deane & Stokes, 2002). For these simulations $a_H \approx 4.4$mm. Figure 6-4 shows the obtained entrainment size distributions, and we highlight that for the large bubbles we study ($a > a_c = 1.3$mm), there is no departure from the $I(a) \propto a^{-14/3}$ power law given by (6.22), including around a_H. It is not surprising that the Hinze scale is not relevant to large bubble entrainment; the Hinze scale comes from the relationship between surface tension and turbulence, where large bubble entrainment is driven by the relationship between gravity and turbulence.

Except for the case with the strongest surface tension (b.iii) where surface tension suppresses L_T, we also see that $a_{\max} \approx 0.3L_T$ consistent with the negligible surface tension results in §6.3.2. Rather than changes in the shape of $I(a)$, we find that the effect of surface tension on large-bubble entrainment is to decrease the magnitude of $I(a)$ through C_I. Consistent with $C_I(We_T \to \infty) = 3.62$ from §6.3.2, we use the C_I from these simulations to obtain the empirical fit

$$C_I(We_T) = \frac{3.62 - A}{1 + \left(We_{T,\mathrm{cr}}/We_T\right)^4} + A\,, \tag{6.24}$$

where $We_{T,\mathrm{cr}} = 34$ and $A = 0.69$, shown in figure 6-5. Our results show that any effects of surface tension on the entrainment of large bubbles become negligible for $We_T > We_{T,\mathrm{cr}}$, which is reflected in the denominator of (6.24).

In theory the mechanistic model in §6.2.2 could be extended to include surface tension effects. While this could provide a mechanistic explanation for (6.24), it is not trivial. The first challenge that arises is that the linearized relationship between pressure fluctuations and surface elevation is now $p \sim (\rho_w g + \sigma\kappa^2)\,\eta$. This makes the relationship between η and κ in (6.9) non-monotonic. The result is that the inter-scale energy transfer becomes more complicated, as one scale of turbulence can lead to the creation of bubbles at two different scales.

6.4 Comparison to air entraining free-surface shear flow

In section 6.3 we considered entrainment in a flow where the only entrainment mechanism present is FST. For general air entraining flows, FST is often present (Brocchini & Peregrine, 2001*a*) so we expect the FST entrainment mechanism to play a role in the total entrainment; however, other entrainment mechanisms may also be relevant. The question then is how significant FST entrainment is to the total entrainment. If FST entrainment is dominant, we expect $I(a)$ to follow the scaling (6.22) we develop. In this section, as an illustration of a more general air entraining flow, we consider a canonical free-surface shear flow, which also serves as a model for ship wakes (Shen *et al.*, 1999; Yu *et al.*, 2019).

The canonical free-surface shear flow we consider is that generated by the initial shear profile,

$$u(z, t = 0)/U = 1 - 0.9988\ \mathrm{sech}(0.88137\, z/L)\,, \tag{6.25}$$

characterized by the shear velocity U and shear length L. At sufficient Reynolds number $Re = UL\rho_w/\mu_w$ a small, random initial perturbation causes the shear profile to generate turbulence (Shen *et al.*, 1999). Yu *et al.* (2019, 2020) show that for Froude number squared, $Fr^2 = U^2/Lg$, greater than a critical $Fr^2_{cr} \approx 5$ the turbulence leads to bubble entrainment. There has also been extensive DNS characterizing this FST at low Fr^2 (Shen *et al.*, 1999, 2000), where it has been shown that the shear instability leads to surface waves propagating in the direction of U (Dimas & Triantafyllou, 1994; Longuet-Higgins, 1998). In Appendix F we extend the linear stability analysis by Longuet-Higgins (1998) to include finite depth effects. For our simulations this linear analysis shows that wavenumbers $kL \in [0.67, 1.23]$ are unstable, and analysis of the wave spectrum confirms these long waves are present even at the large Fr^2 we consider. We highlight these long waves as they represent an alternative mechanism for air entrainment, and the presence of these non-isotropic waves also must have some effect on the near-surface turbulence.

6.4.1 Direct numerical simulation of free-surface shear flow

For DNS of free-surface shear flow we follow the same setup as Yu *et al.* (2019, 2020). We set U and L to unity, which is equivalent to using U and L to nondimensionalize all values in the simulation (i.e., to go from (2.1b) to (2.5)). We set the domain size to $(2\pi/0.6)^2 \times 6$, where initially $z \in [-4, 0]$ is water and $z \in [0, 2]$ is air. The horizontal domain size $2\pi/0.6 \approx 10.472$ ensures that the longest unstable wave number from the shear instability is captured (Shen *et al.*, 1999). This gives an initial free surface area $A_{FS} = 10.472^2$. We set $Re = 1000$ and $We = \infty$ (surface tension not modeled) and study a

Fr^2	N_{sim}	u_{rms}	$\varepsilon \times 10^4$	L_T	Fr_T^2	Re_T	N_I	$a_{\max}/L_T$
5	10	0.065	5.1	0.54	0.04	36	358	0.28
8	8	0.066	5.7	0.51	0.07	34	1344	0.40
10	6	0.067	5.8	0.52	0.09	35	1848	0.41
15	6	0.069	5.5	0.61	0.12	42	4214	0.46
20	6	0.074	5.4	0.74	0.15	55	5877	0.33

Table 6-3: List of free-surface shear flow simulations used for entrainment measurements. All simulations are performed with $We = \infty$ (surface tension not modeled). N_{sim} is the number of ensemble simulations. Turbulence properties u_{rms} and ε are measured using (6.21) during $t \in [40, 70]$. The characteristic length scale $L_T = u_{\text{rms}}^3/\varepsilon$ is used to calculate the near-surface turbulent Froude number $Fr_T^2 = (u_{\text{rms}}^2/L_T)Fr^2$, turbulent Reynolds number $Re_T = (u_{\text{rms}}L_T)Re$. N_I is the number of (resolved) entrainment events recorded and $a_{\max}$ is the radius of the largest entrained bubble observed.

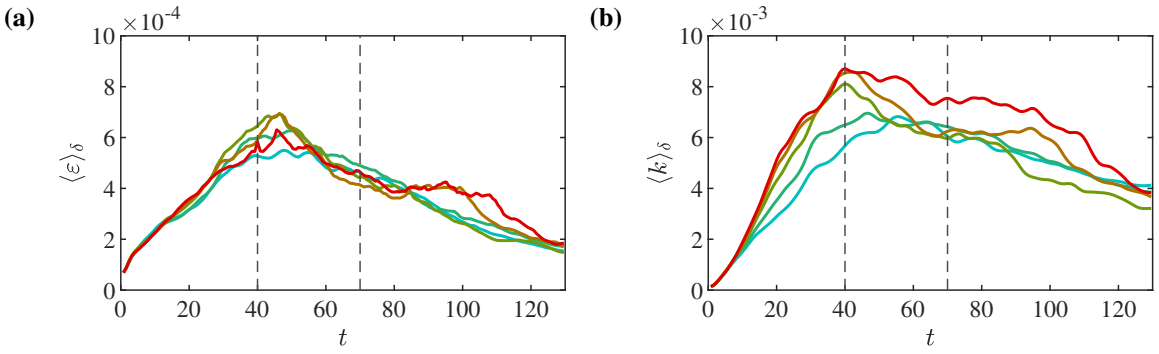


Figure 6-6: Ensemble average (a) turbulent dissipation rate and (b) turbulent kinetic energy in the near-surface region for: ——, $Fr_T^2 = 0.04$; ——, $Fr_T^2 = 0.07$; ——, $Fr_T^2 = 0.09$; ——, $Fr_T^2 = 0.12$; ——, $Fr_T^2 = 0.15$. (- - - -) indicate $t \in [40, 70]$ over which we perform a temporal average to obtain the values in table 6-3.

range of $Fr^2 \geq Fr_{cr}^2$ (see table 6-3). Yu *et al.* (2019) perform a grid convergence study to show a grid $384^2 \times 256$ is sufficient for DNS of this flow, giving a grid size $\Delta \approx 0.027$. We include an additional grid convergence study in Appendix G to confirm that this grid size sufficiently resolves entrainment (and degassing). As in section 6.3, only resolved bubbles (radius larger than $a_{\text{res}} = 1.5\Delta$) are reported. To obtain sufficient entrainment statistics, for each Fr^2 studied (see table 6-3), we repeat the simulation with different realizations of the random initial perturbation to obtain ensemble statistics. The rendering in figure 1-2 comes from a simulation of $Fr^2 = 15$ at time $t = 60$, with the shear velocity going from left to right.

We highlight that, unlike the forced FST in section 6.3, this is an unsteady flow. The initial shear profile generates turbulence which then reaches the surface; however, because there is no further injection of energy, the turbulence slowly decays when dissipation becomes stronger than the shear turbulence production. Figure 6-6 shows the evolution of ensemble-averaged turbulent dissipation rate $\langle \varepsilon \rangle_\delta$ and turbulent kinetic energy $\langle k \rangle_\delta = \frac{1}{2}\langle \mathbf{u} \cdot \mathbf{u} \rangle_\delta$, using the water phase average (6.21), where $\delta = 0.2$ captures the near-surface region for this flow (Yu *et al.*, 2019). We observe a quasi-steady period $t \in [40, 70]$ over which we perform a

temporal average to obtain the turbulence values ε and $u_{\text{rms}} = \sqrt{2\langle k \rangle_\delta / 3}$ used to characterize near-surface turbulence (see values reported in table 6-3). For each Fr^2 studied we perform at least 6 simulations to obtain at least 10^3 resolved entrainment events during this quasi-steady period for all but the smallest Fr^2 (close to Fr^2_{cr}).

For ELA, we use a snapshot interval $\Delta t_s = 0.4$. This corresponds to $\Delta t_s \approx 0.1\Omega(a_{\text{res}})^{-1}$, which avoids spurious events (Chan *et al.* 2021*a*; Gaylo *et al.* 2022; see also Chapter 3). We also use $\Delta t_s = 0.8$ and 1.6 and find the shapes of the entrainment size distribution $I(a)$ (and the degassing size distribution $D(a)$ addressed in Chapter 7) are independent to this range of Δt_s.

6.4.2 Results

Using the same method described in §6.3.2, we bin the entrainment events identified by ELA and calculate the entrainment size distribution (per unit free-surface area) $I(a)/A_{FS}$. Figure 6-7 shows the results, before and after scaling by (6.22). We see that our model does a very good job of predicting the measured $I(a)$ (in log-log space, $R^2 = 0.891$), without even the need to adjust $C_I = 3.62$. Similar to §6.3.2, we find $a_{\max}/L_T \approx 0.3$–0.4 (see table 6-3).

While the fit here ($R^2 = 0.891$) may not be quite as perfect as for the forced FST simulations ($R^2 = 0.990$), this is still a very strong agreement with (6.22). For comparison, Gaylo *et al.* (2024, Figure 4b) tried to scale these same results with Fr^2_T predicted by Yu *et al.* (2020) rather than Fr^6_T predicted by (6.22), and the collapse here in figure 6-7b is clearly much better. As noted, this free-surface shear flow is much more complex than the forced FST we first considered: turbulence is only quasi-steady, and the shear instability generates long waves. While both could have some effect on the entrainment size distribution, the still strong agreement with (6.22) suggests that the FST entrainment mechanism is nonetheless the dominant entrainment mechanism for this free-surface shear flow.

6.5 Comparison to open-channel flow experiments

We now compare our FST entrainment model (6.22) to the experimental measurements by Wei *et al.* (2019) in open-channel flow. The first challenge is that Wei *et al.* (2019) only report the size of bubble produced by 108 entrainment events. Due to the small number of events, we modify the binning method described in §6.3.2 to only require at least 4 events per bin (rather than 15). While this introduced more statistical noise, we find the shape of $I(a)$ is still relatively clear. The second challenge is that, because some number of entrainment events would have been obscured from view and A_{FS} was not reported, the magnitude of $I(a)/A_{FS}$ is unclear; while we can still report the shape of $\propto I(a)$ across radii, there is a chance of bias in what size of entrainment events were obscured.

Figure 6-8 shows the entrainment size distribution we calculate from the entrained bubble sizes measured by Wei *et al.* (2019). Many of the entrainment events they measure produce bubbles smaller than the capillary scale ($a \lesssim a_c$). Focusing on bubbles larger than the capillary scale, it appears that the results are converging to $I(a) \propto a^{-14/3}$ in the limit $a \gg a_c$, consistent with our FST entrainment model (6.22). This implies that in this flow too, FST entrainment is the dominant entrainment mechanism.

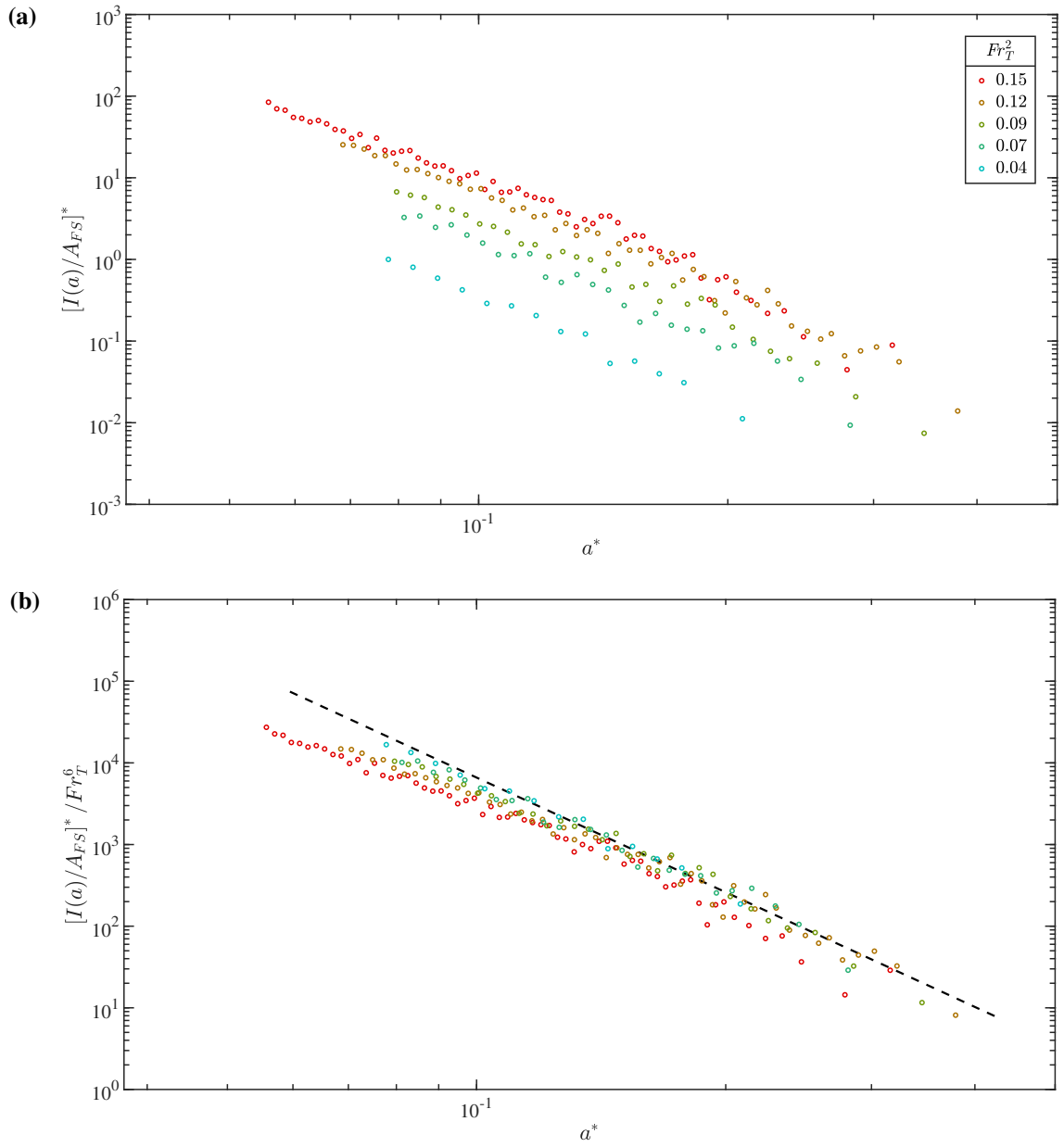


Figure 6-7: Entrainment size distribution in free-surface shear flow (a) as measured; and (b) normalized by Fr_T^6 for different turbulent Froude number $Fr_T^2 = \varepsilon/u_{\rm rms}g$. In (b), (– – –) shows our model (6.22) with $C_I = 3.62$ ($R^2 = 0.891$). Recall $a^* = a\,\varepsilon\,u_{\rm rms}^{-3}$ and $[I(a)/A_{FS}]^* = [I(a)/A_{FS}]\,u_{\rm rms}^{11}\,\varepsilon^{-4}$.

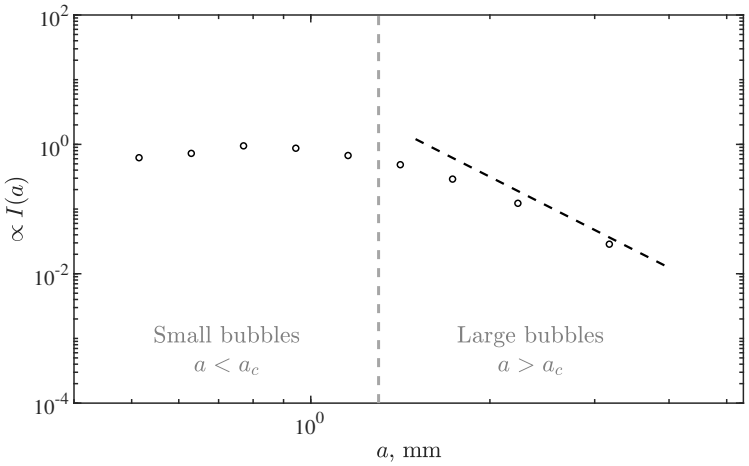


Figure 6-8: Entrainment size distribution calculated from the results of Wei *et al.* (2019) for open-channel flow experiments. The scaling of the y-axis is arbitrary. (– – –) shows $I(a) \propto a^{-14/3}$ from (6.22).

6.6 Conclusion

In air entraining flows a variety of different mechanisms could be involved in creating the entrainment size distribution $I(a)$. However, a common feature of many air entraining flows is the presence of strong turbulence beneath the free surface. In this chapter we have specifically addressed entrainment by this strong FST and characterized the size distribution of bubbles larger than the capillary scale $a_c \approx 1.3\,\text{mm}$.

Applicable to strong FST ($Fr_T^2 > 0.1$) where the near-surface turbulence is nearly isotropic (see Chapter 4), dimensional analysis in §6.2.1 shows that large-bubble FST entrainment is governed by a single scaling parameter α and a single scaling constant C_I. Using DNS of forced turbulence under a free surface, which isolates the FST entrainment mechanism, we obtain that $\alpha = 6$, and $C_I \approx 3.62$. This gives

$$I(a)/A_{FS} = C_I\, \mathrm{g}^{-3}\, \varepsilon^{7/3}\, (2a)^{-14/3} \quad \text{for } a \gg a_c\,, \tag{6.22}$$

to describe how large-bubble entrainment scales with gravity g, the strength of turbulence ε, and bubble radius a. This distribution is in near perfect agreement with the measured distributions ($R^2 = 0.990$), even when including $Fr_T^2 \lesssim 0.1$. Introducing weak surface tension effects, we obtain (6.24) to explain how surface tension decreases C_I, but we highlight that $I(a) \propto a^{-14/3}$ from (6.22) is still true for all $a > a_c$. In §6.2.2 we describe the mechanism for FST entrainment, which is consistent with observations of individual entrainment events and obtains the same g^{-3} (i.e., Fr_T^6) and $a^{-14/3}$ scaling of $I(a)$. Missing from previous models (Yu *et al.*, 2020), the very strong Froude number dependence is explained by the Froude-dependent movement of energy from larger scales of turbulence to smaller scales of bubble given by (6.12).

We start with studying a flow chosen to isolate entrainment by FST, but find our results apply to more general, complex flows where other entrainment mechanisms could be present. We perform DNS of a free-surface shear flow which models aspects of the flow behind a

ship. Without changing any parameters in (6.22), we obtain a very strong agreement with the measured $I(a)$ ($R^2 = 0.891$). We also consider open-channel flow experiments by Wei *et al.* (2019). While many more observations will be necessary to obtain definitive entrainment statistics, our predictions are consistent with their results for large-bubble entrainment.

In addition to validation of our model, the agreement between $I(a)$ observed in free-surface shear flow and open-channel flow and our model suggests FST entrainment is the dominant entrainment mechanism in these flows. We note that there are many air entraining flows and some may involve different (dominant) entrainment mechanisms, notably cavity entrapment in breaking waves (Deike *et al.*, 2016; Chan *et al.*, 2021*c*; Gao *et al.*, 2021) or plunging jets (Kiger & Duncan, 2012; Bertola *et al.*, 2018). While we do not expect one model to be able explain the features of all conceivable air entraining flows, if entrainment by FST mechanism is a significant contributor to the total entrainment one expects to see its features reflected in $I(a)$. Noting the common presence of near-surface turbulence (Brocchini & Peregrine, 2001*a*), this suggests that our model for FST entrainment could be applicable broad classes of air entraining free-surface flows. In particular, the confirmed applicability to free-surface shear flow suggests our model is relevant for modeling flows like ship wakes.

Chapter 7

Bubble Degassing in Free-Surface Turbulence

In this chapter we focus on bubble degassing, the $D(a)$ term in the population balance equation (1.2). Using ELA (Gaylo *et al.* 2022; see also Chapter 3), we are able to, for the first time, directly measure the degassing size distribution $D(a)$ and elucidate the scaling. After quantifying degassing, we examine the bubble size distribution $N(a)$ predicted by (1.2). By comparing the strength of degassing to fragmentation, we identify a class of free-surface flows that are degassing, rather than fragmentation, dominated. We show degassing-dominance leads to an equilibrium bubble population that is clearly distinct from the $N(a) \propto a^{-10/3}$ equilibrium solution for fragmentation dominated bubble populations, such as in plunging breaking waves.

Many of the key results of this chapter are summarized in "Effect of degassing on bubble populations in air entraining free-surface turbulent flows" by Gaylo, Hendrickson & Yue (2024). A major difference is that here we use the new model for $I(a)$ (Gaylo & Yue 2025, see also Chapter 6) in our model of degassing-dominated bubble populations ($N(a) = I(a)/\Lambda(a)$). This leads to significantly better agreement between the bubble population predicted by our model and observed in DNS.

7.1 Introduction

A fundamental property of the bubble size distribution is the power law slope β describing how the distribution depends on bubble size, $N(a) \propto a^{\beta}$. As discussed in section 1.3, $\beta = -10/3$ is the equilibrium solution for air entraining flows which are dominated by fragmentation (Garrett *et al.*, 2000; Gaylo *et al.*, 2021). Studying plunging breaking waves, Deane & Stokes (2002) observe $\beta = -10/3$ for super-Hinze bubbles ($a > a_H$, where a_H is given by (5.1)) during the air entraining period (referred to by them as the acoustically active period) and conclude that fragmentation and entrainment are the dominant mechanisms during this period. Many have since reported $\beta(a > a_H) \approx -10/3$ during the air entraining period of plunging breaking waves (Deike, 2022).

After the air entraining period, when the total volume of air begins to decrease, Deane & Stokes (2002) observe steepening of the bubble size distribution ($\beta < -10/3$), which

they attribute to degassing and dissolution becoming relevant, in the absence of entrainment and fragmentation. They refer to this period as the quiescent period. Previous work on degassing in the context of breaking waves focuses on this quiescent period, seeking to explain how degassing evolves a bubble population after entrainment and fragmentation have already established it (e.g., Callaghan *et al.*, 2013; Deike *et al.*, 2016). What has not been addressed is the effect of degassing during the air entraining period, when degassing acts to balance entrainment. The wide agreement on $\beta(a > a_H) \approx -10/3$ suggests that this effect is negligible compared to fragmentation for plunging breaking waves; however, in this work we will show that for air entraining FST, degassing is in fact the dominant balance to entrainment during the air entraining period.

Here we evaluate the effect of degassing in FST. In contrast to breaking waves where the energy to create bubbles comes from a (downward) mean flow, in this flow the energy comes directly from the underlying turbulence. In the absence of a mean downward flow, we theorize that degassing will be stronger relative to other entrainment-balancing mechanisms, such as fragmentation, and our results confirm degassing is dominant. In section 7.2 we start by reviewing the population balance equation (PBE) and deriving quantities to measure degassing dominance. We then derive the equilibrium solution to the PBE for degassing dominated bubble populations in air entraining flow. Using a simple model of degassing and the entrainment size distribution from Chapter 6, we obtain the power-law slope β for this equilibrium solution.

In section 7.3 we use the same DNS of canonical free-surface shear flow (Shen *et al.*, 1999; Yu *et al.*, 2019) as in Chapter 6, and ELA (Gaylo *et al.* 2022; see also Chapter 3) directly obtains degassing statistics. Studying large Weber numbers, where fragmentation is strongest, we find that degassing is dominant over fragmentation, independent of Fr_T^2. ELA measurements of the degassing rate confirm the accuracy of our simple model of degassing. Finally, measurements of $N(a)$ (independent of ELA) agree with our predicted equilibrium solution for degassing dominated bubble populations. In section 7.4 we discuss how this result elucidates how degassing-dominated bubble populations (expected in air entraining FST) scale with Froude number.

7.2 Modeling degassing-dominated bubble populations

Using the PBE, we seek a model for β in degassing-dominated air entraining flows, as well as formal definitions of "air entraining" and "degassing-dominated." We start with a brief review of the PBE introduced in Chapter 1. Recall that, because fragmentation and degassing of an individual bubble is independent of other bubbles, the distributions are linearly dependent on $N(a)$ and we can write

$$S_f^-(a) = \Omega(a)N(a)\,, \tag{1.5}$$

and

$$D(a) = \Lambda(a)N(a)\,. \tag{1.6}$$

This defines the fragmentation rate $\Omega(a)$, the focus of Chapter 5, and the degassing rate $\Lambda(a)$, a focus of this chapter. With these two relationships, the PBE (1.3) (which neglects

dissolution and coalescence) can be split into positive and negative terms:

$$\partial N/\partial t(a) = \left[I(a) + S_f^+(a)\right] - \left[\Lambda(a) + \Omega(a)\right]N(a)\,. \tag{1.7}$$

Weighting the PBE by a^3 and integrating over all bubble sizes, we obtain the evolution of the total entrained volume

$$\mathrm{d}V/\mathrm{d}t = Q_I - Q_D\,. \tag{1.8}$$

Because it only moves volume between bubble sizes, fragmentation does not contribute $\mathrm{d}V/\mathrm{d}t$.

To classify flows as air entraining, we consider the ratio between degassing flux and entrainment flux,

$$\mathcal{D} \equiv Q_D/Q_I\,. \tag{7.1}$$

We define an air entraining flow to be one where $\mathcal{D} \in [0, 1]$, meaning the amount of entrained air is increasing ($\mathrm{d}V/\mathrm{d}t > 0$) or at equilibrium ($\mathrm{d}V/\mathrm{d}t = 0$). This excludes flows with negligible entrainment ($\mathcal{D} \gg 1$), such as the quiescent period of breaking waves. For breaking waves during the air entraining period, Deane & Stokes (2002) show that entrainment and fragmentation are the primary mechanisms, implying that degassing is negligible ($\mathcal{D} \ll 1$). Our interest is air entraining flows where degassing is important ($\mathcal{D} \lesssim 1$).

7.2.1 Defining degassing-dominated bubble populations

We define a degassing-dominated flow to be one where degassing, rather than fragmentation, dominates the evolution of the bubble population. As a simple measure, we could look at the negative terms in (1.7) and see that if $\Lambda(a) > \Omega(a)$ a bubble is more likely to degases than fragment. Formally we need to also consider the positive term $S_f^+(a)$, but it turns out that $\Lambda(a)/\Omega(a)$ is still an appropriate measure of degassing versus fragmentation dominance.

To understand the effect of $S_f^+(a)$, we take the same approach as in §1.3.1 and assume identical fragmentation, where all bubbles fragment into exactly m identically sized daughter bubbles.[1] This gives

$$S_f^+(a) \sim m^{4/3}\Omega(m^{1/3}a)N(m^{1/3}a)\,. \tag{7.2}$$

While this may not be the most realistic model fragmentation, in Chapter 5 we showed that unrealistic fragmentation models ($C_f = 1$ here) can capture the behavior of fragmentation cascades if m is chosen such that a realistic value of C_τ is obtained in (5.30). Using $C_\tau \approx 9$ and $C_\Omega \approx 1.4$ (Gaylo *et al.* 2023, see also Chapter 5), (5.30) gives $m \approx 1.45$ is needed for (7.2) to realistically capture the fragmentation cascade.

Assuming $a \gg a_H$ such that we can use the power-law relationship (1.11), we have $\Omega(m^{1/3}a) = m^{-2/9}\Omega(a)$. If we assume a power-law for the bubble size distribution, $N(a) \propto a^\beta$, we obtain

$$S_f^+(a) \sim S_f^-(a)\; m^{10/9+\beta/3}\,. \tag{7.3}$$

[1] In (5.2) this means setting $\beta(a; a') = \delta(a' - m^{-1/3}a)$, where δ is the Dirac delta function.

This allows us to rewrite (1.7) as

$$\partial N/\partial t(a) \sim I(a) - \left[\Lambda(a) + \Omega(a)(1 - m^{10/9+\beta/3})\right]N(a)\,, \tag{7.4}$$

Which makes it clear that the ratio

$$\frac{\text{Effect of Degassing}}{\text{Effect of Fragmentation}} \sim \frac{\Lambda(a)}{\Omega(a)}\left(\frac{1}{\left|1 - m^{10/9+\beta/3}\right|}\right)\,. \tag{7.5}$$

Using $m \approx 1.45$ as an example, we have $|1 - m^{10/9+\beta/3}| < 1$ for any $\beta < 2.26$. As β is typically negative (more smaller bubbles than larger bubbles), we can be confident that $\Lambda(a)/\Omega(a) > 1$ means that the effects of degassing are larger than fragmentation. If β is known, a tighter bound could be determined. In summary, if $\Lambda(a)/\Omega(a) \gg 1$ we can be sure that the effect of fragmentation on $N(a)$ is small compared to the effect degassing, defining a degassing dominated population.

7.2.2 Degassing-dominated population balance equation

For degassing-dominated bubble populations ($\Lambda(a)/\Omega(a) \gg 1$) we can remove the negligible fragmentation term from the PBE, giving

$$\partial N/\partial t(a) = I(a) - \Lambda(a)N(a)\,. \tag{7.6}$$

This is significantly more simple than (1.7) because, by removing the integral coming from $S_f^+(a)$, we have removed any dependence of $N(a)$ on $N(a')$ for any $a' \neq a$. With no dependence between bubble sizes, we could write (7.6) as a set of *independent* first-order linear ordinary differential equations,

$$\left[\mathrm{d}N_i/\mathrm{d}t = I_i - \Lambda_i N \quad \forall \quad \text{bubble radii } a_i\right]\,. \tag{7.7}$$

which are easy to solve for a given initial condition, entrainment rate I_i (which could depend on time) and degassing rate Λ_i (which could also depend on time).

Equilibrium and non-equilibrium regimes

A useful way to interpret (7.6) comes from rearranging it:

$$I(a) = \Lambda(a)N(a) + \partial N/\partial t(a)\,. \tag{7.8}$$

We see that bubble entrainment is balanced by two terms: degassing, $\Lambda(a)N(a)$, and/or an increase in the bubble size distribution, $\partial N/\partial t(a)$. With a little more rearranging, the ratio of the first to the second term defines

$$\gamma(a) \equiv \frac{\Lambda(a)N(a)}{I(a) - \Lambda(a)N(a)}\,. \tag{7.9}$$

For $\gamma(a) \ll 1$, entrainment is balanced by a corresponding increase in the bubble size distribution, $\partial N/\partial t(a) \approx I(a)$. Suppose we start from $N(a) = 0$ at $t = 0$. $\gamma(a) \ll 1$ suggests a non-equilibrium regime of linear growth $N(a) \approx tI(a)$. Plugging this back into (7.9), we get

$$\gamma(a) \sim \frac{\Lambda(a)t}{1 - \Lambda(a)t} \quad \text{for } \gamma(a) \ll 1\,. \tag{7.10}$$

We see that $\gamma(a) \ll 1$ can only be true for timescales $t \ll 1/\Lambda(a)$. Beyond this, we enter a regime $\gamma(a) \gg 1$ where entrainment is primally balanced by degassing and the bubble population does not change significantly ($\partial N/\partial t(a) \sim 0$). For a given radius, the characteristic time to reach this equilibrium regime is $1/\Lambda(a)$.

While we stress that the evolution of each individual radius is independent for a degassing-dominated bubble population, it can still be useful to define measures of equilibrium/non-equilibrium which describe the entire bubble population. If we weigh the numerator and denominator of (7.9) by bubble volume a^3 and integrate separately, we obtain $\Gamma \equiv Q_D/(Q_I - Q_D)$. This can be related to $\mathcal{D}$ from (7.1):

$$\Gamma = \frac{\mathcal{D}}{1 - \mathcal{D}}\,. \tag{7.11}$$

This Γ gives a measure of $\gamma(a)$ for all radii of a bubble population. We expect the non-equilibrium regime for $\Gamma \ll 1$ and the equilibrium regime for $\Gamma \gg 1$.

Power-law slope of the equilibrium solution, β

Here, our primary interest is the equilibrium solution to (7.6). As demonstrated, we expect this solution if $I(a)$ is constant or changes over timescales much longer than $1/\Lambda(a)$. Setting $\partial N/\partial t(a) = 0$, the balance between entrainment and degassing in (7.6) gives

$$N(a) \approx I(a)/\Lambda(a)\,. \tag{7.12}$$

Chapter 6 shows that $I(a) \propto a^{-14/3}$ for $a \gg a_c$, where a_c is the capillary scale.[2] If we assume that degassing follows some power law $\Lambda(a) \propto a^{\alpha}$, the power law slope of $N(a)$ in degassing-dominated bubble populations is

$$\beta = -14/3 - \alpha \quad \text{for } a \gg a_c\,. \tag{7.13}$$

Next, we determine the value of α.

7.2.3 Power-law scaling of degassing in free-surface turbulence

To predict α, we derive a characteristic bubble depth L_Λ and bubble rise velocity U_Λ such that the degassing rate is $\Lambda(a) \propto U_\Lambda/L_\Lambda$. We start with the characteristic depth L_Λ, which we described the depth of a recently entrained bubble. For the energy argument in Chapter 6, we consider an initial depth $d \sim 2a$ (see figure 6-1). This initial depth d was used to relate

[2] Gaylo *et al.* (2024) use $I(a) \propto a^{-10/3}$ predicted by Yu *et al.* (2020), which we now know is wrong (Gaylo & Yue 2025; see also Chapter 6).

the potential energy of the surface deformation to the initial potential energy of the entrained bubble. What we did not consider in §6.2.2 is that turbulence could then advect the bubble further downward.

As confirmed in Chapter 4, near-surface turbulence is isotropic for $Fr_T^2 > 0.1$, allowing application of the Kolmogorov energy cascade. In the inertial sub range, an eddy of size ℓ has a velocity $u_\ell \sim \varepsilon^{1/3}\ell^{1/3}$ and is coherent over $t_\ell \sim \varepsilon^{-1/3}\ell^{2/3}$. A bubble of radius a has an added mass $m \sim \rho_w a^3$ and (at large Reynolds numbers) feels a force from the eddy $F \sim \rho_w a^2 u_\ell^2$. Assuming this force is greater than buoyancy ($F \gg \rho_w \mathrm{g}(4\pi/3)a^3$) the bubble accelerates downward with $\dot{w} \sim u_\ell^2/a$. This moves the bubble to a depth $z \sim t_\ell{}^2\dot{w} = \ell^2/a$. The eddies with the most energy are $\ell \sim L_T = {u_{\mathrm{rms}}}^3\varepsilon^{-1}$, and we assume these are primarily responsible for pushing bubbles downwards, giving a model

$$L_\Lambda = C_L \, {u_{\mathrm{rms}}}^6 \varepsilon^{-2} a^{-1} \,, \tag{7.14}$$

where C_L is an unknown constant of proportionality.

For characteristic velocity U_Λ, either buoyant rise or turbulent advection can be relevant. For buoyant rise, we consider W_T, the terminal rise velocity of a bubble in quiescent flow. This is characterized by a Reynolds number,

$$Re_W = \frac{(2a)W_T}{\nu_w} \,. \tag{7.15}$$

Park *et al.* (2017) summarize the three regimes of bubble terminal rise velocity:

$$W_T(a) = \begin{cases} \frac{1}{12}\,\mathrm{g}\,\nu_w^{-1}\,(2a)^2 & Re_W < 1 \\ 0.144\,\mathrm{g}^{5/6}\,\nu_w^{2/3}\,(2a)^{3/2} & 1 < Re_W < 100 \\ 0.711\,\mathrm{g}^{1/2}\,(2a)^{1/2} & Re_W > 150, \text{ and } Bo > 40 \end{cases} \tag{7.16}$$

(Davies & Taylor, 1950; Mendelson, 1967; Wallis, 1974; Clift *et al.*, 2013). We consider bubbles in the inertial regime $1 < Re_W < 100$, like those in our DNS, and bubbles in the spherical-cap regime $Re_W > 150$, which we expect for larger physical-scale free-surface flows where the Reynolds number is much larger than captured by our DNS (e.g., Hendrickson *et al.*, 2019). While our focus here is bubbles $a \gg a_c$ where surface tension effects are negligible, Park *et al.* (2017) provide the full equation for the spherical-cap regime which includes the effect of surface tension.

In addition to buoyant rise, the movement of bubbles can be affected by turbulent advection, which we model as $U_\Lambda = C_U\, u_{\mathrm{rms}}$, where C_U is an unknown constant of proportionality. For small bubbles, we expect their movement to be dominated by turbulent advection. For large bubbles, we expect their movement to be dominated by buoyant rise. Setting $W_T = C_U u_{\mathrm{rms}}$ defines the transition between these two regimes, and we solve (7.16) for a to define the critical radius

$$a_\Lambda \equiv \begin{cases} 1.82\;\; (C_U\, u_{\mathrm{rms}})^{2/3}\;\; \nu_w^{4/9}\,\mathrm{g}^{-5/9} & \text{DNS-scale } Re_W \\ 0.99\;\; (C_U\, u_{\mathrm{rms}})^2\;\; \mathrm{g}^{-1} & \text{Physical } Re_W \end{cases} . \tag{7.17}$$

Thus, we define a simple model for the two regimes of the characteristic rise velocity,

$$U_\Lambda = \begin{cases} C_U\, u_{\text{rms}} & a < a_\Lambda \\ W_T(a) & a > a_\Lambda \end{cases} . \tag{7.18}$$

Although more advanced models consider the interaction between buoyancy and turbulence (relevant near $a = a_\Lambda$) (e.g., Salibindla *et al.*, 2020; Ruth *et al.*, 2021), we find this simple two-regime model is sufficient to explain the power law scaling of degassing.

Using our models for L_Λ and U_Λ with $\Lambda(a) = U_\Lambda / L_\Lambda$, we obtain

$$\Lambda(a) = \Lambda_0 \begin{cases} (a/a_\Lambda) & a < a_\Lambda \\ (a/a_\Lambda)^{5/2} & a > a_\Lambda, \text{ DNS-scale } Re_W \\ (a/a_\Lambda)^{3/2} & a > a_\Lambda, \text{ Physical } Re_W \end{cases} , \tag{7.19a}$$

where

$$\Lambda_0 = \frac{C_U}{C_L} {u_{\text{rms}}}^{-5} \varepsilon^2 a_\Lambda = \begin{cases} 1.82\; {C_L}^{-1}\, {C_U}^{5/3}\, {u_{\text{rms}}}^{-13/3}\, \varepsilon^2\, {\nu_w}^{4/9}\, \text{g}^{-5/9} & \text{DNS-scale } Re_W \\ 0.99\; {C_L}^{-1}\, {C_U}^{3}\; {u_{\text{rms}}}^{-3}\, \varepsilon^2\, \text{g}^{-1} & \text{Physical } Re_W \end{cases} . \tag{7.19b}$$

This gives the power-law slope α for degassing in the turbulence-driven regime ($a < a_\Lambda$) and buoyancy-driven regime ($a > a_\Lambda$),

$$\alpha = \begin{cases} 1 & a < a_\Lambda \\ 5/2 & a > a_\Lambda, \text{ DNS-scale } Re_W \\ 3/2 & a > a_\Lambda, \text{ Physical } Re_W \end{cases} , \tag{7.20}$$

Using (7.13) for degassing dominated populations, we obtain

$$\beta = \begin{cases} -17/3\ (= -5.6\bar{6}) & a < a_\Lambda \\ -43/6\ (= -7.1\bar{6}) & a > a_\Lambda, \text{ DNS-scale } Re_W \\ -37/6\ (= -6.1\bar{6}) & a > a_\Lambda, \text{ Physical } Re_W \end{cases} . \tag{7.21}$$

In all these regimes, we predict that degassing-dominated bubble populations have a β significantly more negative (more smaller bubbles, fewer larger bubbles) than $\beta = -10/3$ ($= -3.\bar{3}$) for fragmentation-dominated bubble populations.

7.3 Quantifying degassing in air entraining free-surface shear flow

To quantify the effects of degassing on the bubble population, we consider the same canonical free-surface shear flow we used to quantify the effects of entrainment in Chapter 6 (see §6.4 for details). Figure 7-1 lists the simulations performed. Note that these are the same simulations as in Chapter 6 (see table 6-3). In these simulations, we use ELA (Gaylo *et al.*

Fr^2	N_{sim}	u_{rms}	$\varepsilon \times 10^4$	L_T	Fr_T^2	Re_T	N_D	$\mathcal{D}$
5	10	0.065	5.1	0.54	0.04	36	231	0.60 ± 0.09
8	8	0.066	5.7	0.51	0.07	34	976	0.72 ± 0.05
10	6	0.067	5.8	0.52	0.09	35	1344	0.71 ± 0.05
15	6	0.069	5.5	0.61	0.12	42	3150	0.75 ± 0.06
20	6	0.074	5.4	0.74	0.15	55	4167	0.68 ± 0.05

Table 7-1: List of free-surface shear flow simulations used for degassing measurements. All simulations are performed at $We = \infty$ (surface tension not modeled). N_{sim} is the number of ensemble simulations. Turbulence properties u_{rms} and ε are measured using (6.21) during $t \in [40, 70]$. The characteristic length scale $L_T = u_{\text{rms}}^3/\varepsilon$ is used to calculate the near-surface turbulent Froude number $Fr_T^2 = (u_{\text{rms}}^2/L_T)Fr^2$, turbulent Reynolds number $Re_T = (u_{\text{rms}}L_T)Re$. N_D is the number of (resolved) degassing events recorded. The 95% confidence interval is given for the average value of $\mathcal{D}$ from (7.1).

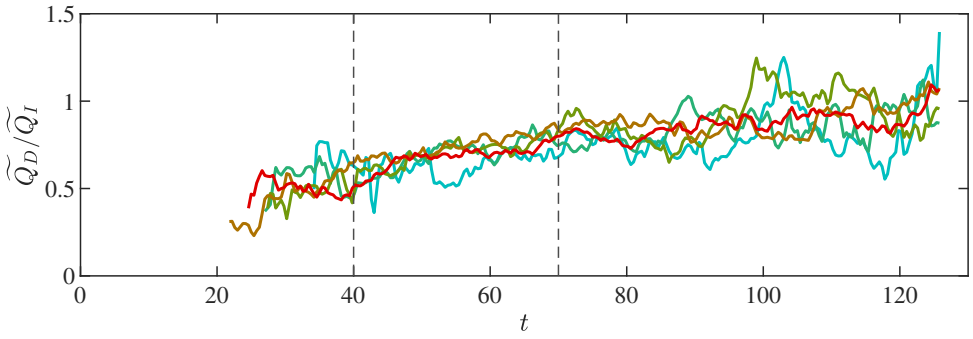


Figure 7-1: Ratio of degassing flux to entrainment flux for: ——, $Fr_T^2 = 0.04$; ——, $Fr_T^2 = 0.07$; ——, $Fr_T^2 = 0.09$; ——, $Fr_T^2 = 0.12$; ——, $Fr_T^2 = 0.15$. (- - - -) indicate $t \in [40, 70]$ over which we perform a temporal average to obtain the values in table 7-1. For clarity, we apply a top-hat filter (width 8 in time) denoted $\widetilde{\cdot}$.

2022; see also Chapter 3) to identify and measure degassing events over the snapshot interval Δt_s (see figure 3-1c). As in §6.4, we use a snapshot interval $\Delta t_s = 0.16$ and only report bubbles $a > a_{\text{res}}$, where $a_{\text{res}} = 1.5\Delta$.

7.3.1 Volume flux ratio, $\mathcal{D}$

We start by considering $\mathcal{D}$ from (7.1), the ration of degassing volume flux Q_D to entrainment volume flux Q_I. Figure 7-1 shows instantaneous $\mathcal{D}$ over the evolution of the free-surface shear flow. First, we note that the evolution of $\mathcal{D}$ is consistent across the range of Fr_T^2. Recall in section 7.2.2 we showed the characteristic timescale for the evolution of $\gamma(a)$ is $\propto 1/\Lambda(a)$. For $a < a_\Lambda$, (7.19) gives $\Lambda(a) \propto {u_{\text{rms}}}^{-5}\varepsilon^2 a$, independent of g. Γ (related to $\mathcal{D}$ through (7.11)) is a volume weighted average of $\gamma(a)$ across bubble radius. This means that if the majority of air volume is in bubbles $a < a_\Lambda$, we expect the evolution of $\mathcal{D}$ to be Fr_T^2 independent.

Consistent with the analysis in §7.2.2, we see $\Gamma < 1$ ($\mathcal{D} < 0.5$) for a short time near the start of entertainment ($t < 30$), before $\Gamma > 1$ is obtained. We focus on the period $t \in [40, 70]$ when the turbulence is quasi-steady (see figure 6-3). Unlike $\mathcal{D} \ll 1$ seen for breaking waves (Deane & Stokes, 2002), our free-surface shear flow shows $\mathcal{D} \approx 0.7$ (or $\Gamma \approx 2.3$) during this air entraining period, which, as expected, is independent of Fr_T^2. This $\mathcal{D} \sim O(1)$ shows that degassing is an important mechanism balancing entrainment. Next, we will prove this flow is degassing dominated, in which case $\Gamma > 1$ shows that although the turbulence and entrainment size distribution $I(a)$ are changing in time, those timescales are large enough relative to $1/\Lambda(a)$ such that the bubble population is in the equilibrium regime.

7.3.2 Degassing rate, $\Lambda(a)$

We use ELA to obtain the degassing events over $t \in [40, 70]$. Using the same binning method described in §6.3.2, we calculate the degassing size distribution $D(a)$, where there are at least 15 events per bin.[3] For each bin, we then divide by the average $N(a)$ (calculated from 75 evenly spaced samples over $t \in [40, 70]$ from each simulation) to obtain the degassing rate $\Lambda(a)$, shown in figure 7-2. Least-squares regression in log-log space (after binning, $n = 160$ data points) gives the best fit $C_U = 3.0$ and $C_L = 0.219$.[4] Looking at figure 7-2b we see that our model (7.19), with these two fitting parameters, shows a good agreement ($R^2 = 0.761$) with the numerical results.

To directly quantify the power-law slope $\Lambda(a) \propto a^\alpha$, we also perform least-square regression to

$$\log\left[\Lambda(a)/\Lambda_0\right] = \hat{\alpha} \log\left[a/a_\Lambda\right] + C\,, \tag{7.22}$$

for $a/a_\Lambda < 0.8$ and $a/a_\Lambda > 1.25$ separately. As discussed in section 7.2.3, our model is designed to capture the power law far from $a/a_\Lambda \approx 1$, where $C_U\, u_{\text{rms}} \approx W_T(a)$. For $a/a_\Lambda < 0.8$ we obtain a 95% confidence intervals $\hat{\alpha} \in [0.82, 1.08]$, consistent with $\alpha(a < a_\Lambda) = 1$ predicted by (7.20). For $a/a_\Lambda > 1.25$ we only have 16 data points, so the range $\hat{\alpha} \in [-2.07, 2.90]$ is quite wide, but not inconsistent with $\alpha(a > a_\Lambda) = 2.5$ predicted by (7.20). Qualitatively, figure 7-2b is compatible with $\alpha(a > a_\Lambda) = 2.5$.

7.3.3 Degassing dominance

We now examine the strength of degassing relative to fragmentation. Figure 7-3 shows that $\Lambda(a)/\Omega(a) > 1$ for most bubble radii. Using (7.5), we can calculate the ratio of degassing versus fragmentation including the cancellation between S_f^+ and S_f^-. Based on $m \approx 1.45$ chosen to match C_τ and $\beta = -17/3$ from (7.21), (7.5) becomes

$$\frac{\text{Effect of Degassing}}{\text{Effect of Fragmentation}} \sim \frac{\Lambda(a)}{\Omega(a)} \left(\frac{1}{0.251}\right). \tag{7.23}$$

In figure 7-3 we see $\Lambda(a)/\Omega(a) > 0.251$ for all bubble radii. These results show that degassing is not only important in free-surface shear flow, but dominant over fragmentation.

[3] Gaylo *et al.* (2024) use a different binning method, leading to only slightly different plots.

[4] 95% confidence intervals: $C_U \in [2.6, 3.3]$ and $C_L \in [0.197, 0.244]$

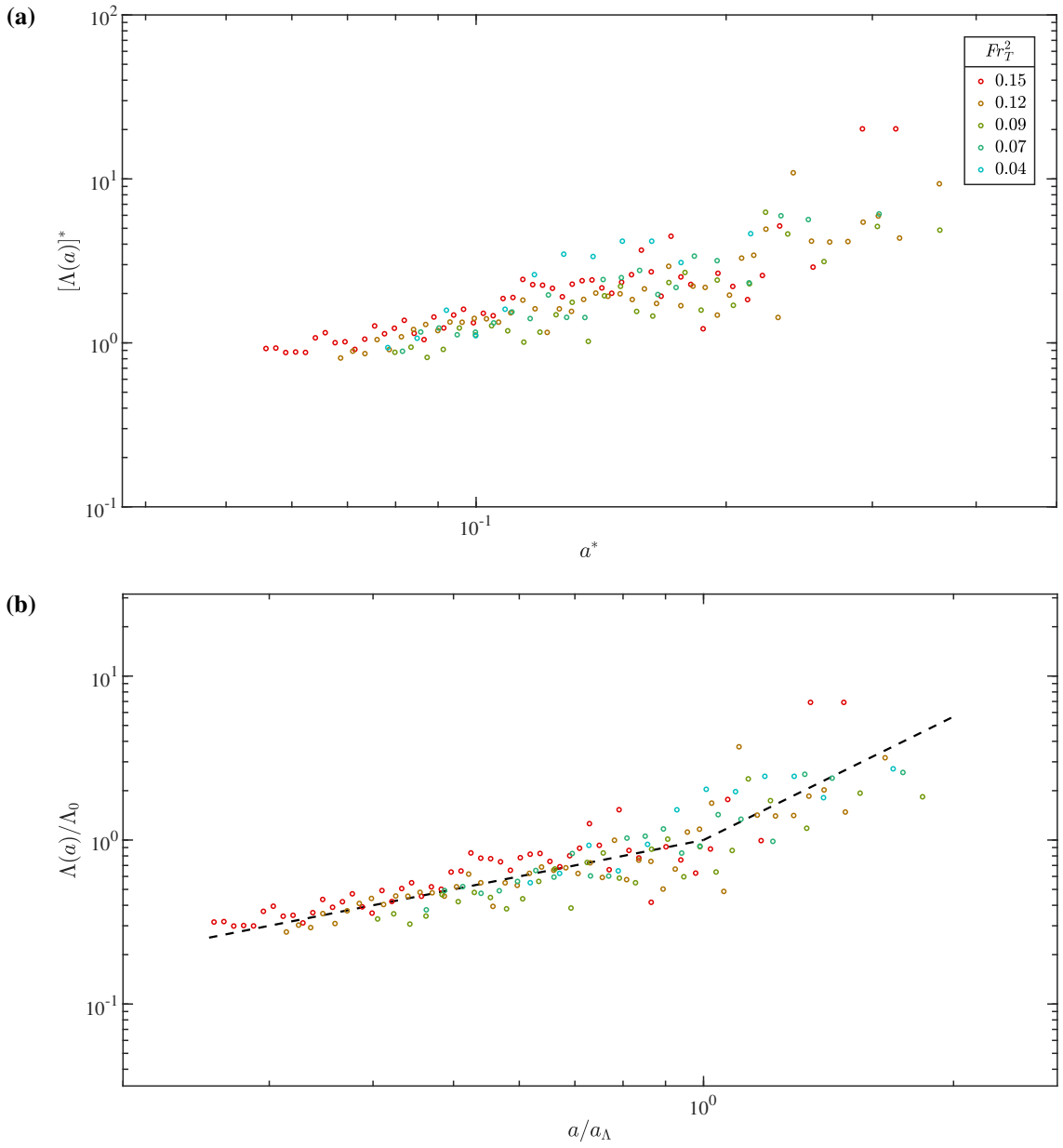


Figure 7-2: Degassing rate in free-surface shear flow (a) as measured and (b) compared to our model (7.19) for different turbulent Froude number $Fr_T^2 = \varepsilon/u_{\rm rms}g$. In (a), $a^* = a\,\varepsilon\,u_{\rm rms}^{-3}$ and $[\Lambda(a)]^* = [\Lambda(a)]\,\varepsilon^{-1}\,u_{\rm rms}^2$. In (b), (– – –) shows our model (7.19) with $C_U = 3.0$ and $C_L = 0.219$ ($R^2 = 0.761$). Recall $a^* = a\,\varepsilon\,u_{\rm rms}^{-3}$.

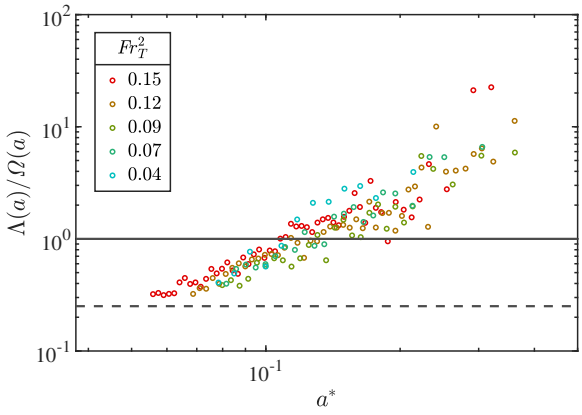


Figure 7-3: Degassing rate in free-surface shear flow compared to the fragmentation rate for different turbulent Froude number $Fr_T^2 = \varepsilon/u_{\mathrm{rms}}\mathrm{g}$. $\Omega(a)$ is given by (5.5) with $C_\Omega = 0.42$ (Martínez-Bazán *et al.*, 1999*a*). (——) shows $\Lambda(a)/\Omega(a) = 1$ and (– – –) shows $\Lambda(a)/\Omega(a) = 0.251$.

7.4 Equilibrium bubble size distributions in degassing-dominated flow

In the previous section, we first confirmed that the degassing size distribution $\Lambda(a)$ follows (7.19) and elucidated the scaling coefficients, and second confirmed that degassing is dominant over fragmentation. As discussed in section 7.2.2, degassing dominance means that we expect the equilibrium bubble size distribution $N(a)$ to follow (7.12). In Chapter 6 we obtain (6.22) to describe $I(a)$ (per mean free surface area A_{FS}) for super-capillary scale bubbles in FST. With (7.19) obtained here to describe $\Lambda(a)$ in FST, we can now solve (7.12) for the equilibrium bubble size distribution in degassing-dominated bubble populations,

$$N(a) = N_0 \begin{cases} (a/a_\Lambda)^{-17/3} & a < a_\Lambda \\ (a/a_\Lambda)^{-43/6} & a > a_\Lambda, \text{ DNS-scale } Re_W \\ (a/a_\Lambda)^{-37/6} & a > a_\Lambda, \text{ Physical } Re_W \end{cases}, \tag{7.24a}$$

where

$$N_0 = I(a_\Lambda)/\Lambda_0 = A_{FS}\left[C_L\, C_I\, {C_U}^{-1}\, 2^{-14/3}\right] \mathrm{g}^{-3} u_{\mathrm{rms}}^5 \varepsilon^{-2} a_\Lambda^{-17/3}, \tag{7.24b}$$

$I(a_\Lambda)$ comes from (6.22), and Λ_0 comes from (7.19). This applies to large bubbles ($a \gg a_c$) where (6.22) is valid.

7.4.1 Measurements in free-surface shear flow

Figure 7-4a shows the bubble size distribution we measure from the free-surface shear flow simulations, nondimensionalized by the near-surface turbulence characteristic scale $L_T = u_{\mathrm{rms}}^3/\varepsilon$. The bubble size distribution is much steeper than $\beta = -10/3$ we would expect for fragmentation-dominated bubble populations. The clear difference between the simulation results and $\beta = -10/3$ also excludes the prediction $N(a)/A_{FS} \propto \mathrm{g}^{-1}\varepsilon^{2/3}a^{-10/3}$ from Yu *et al.* (2020), which misses the inter-scale energy transfer associated with entrainment

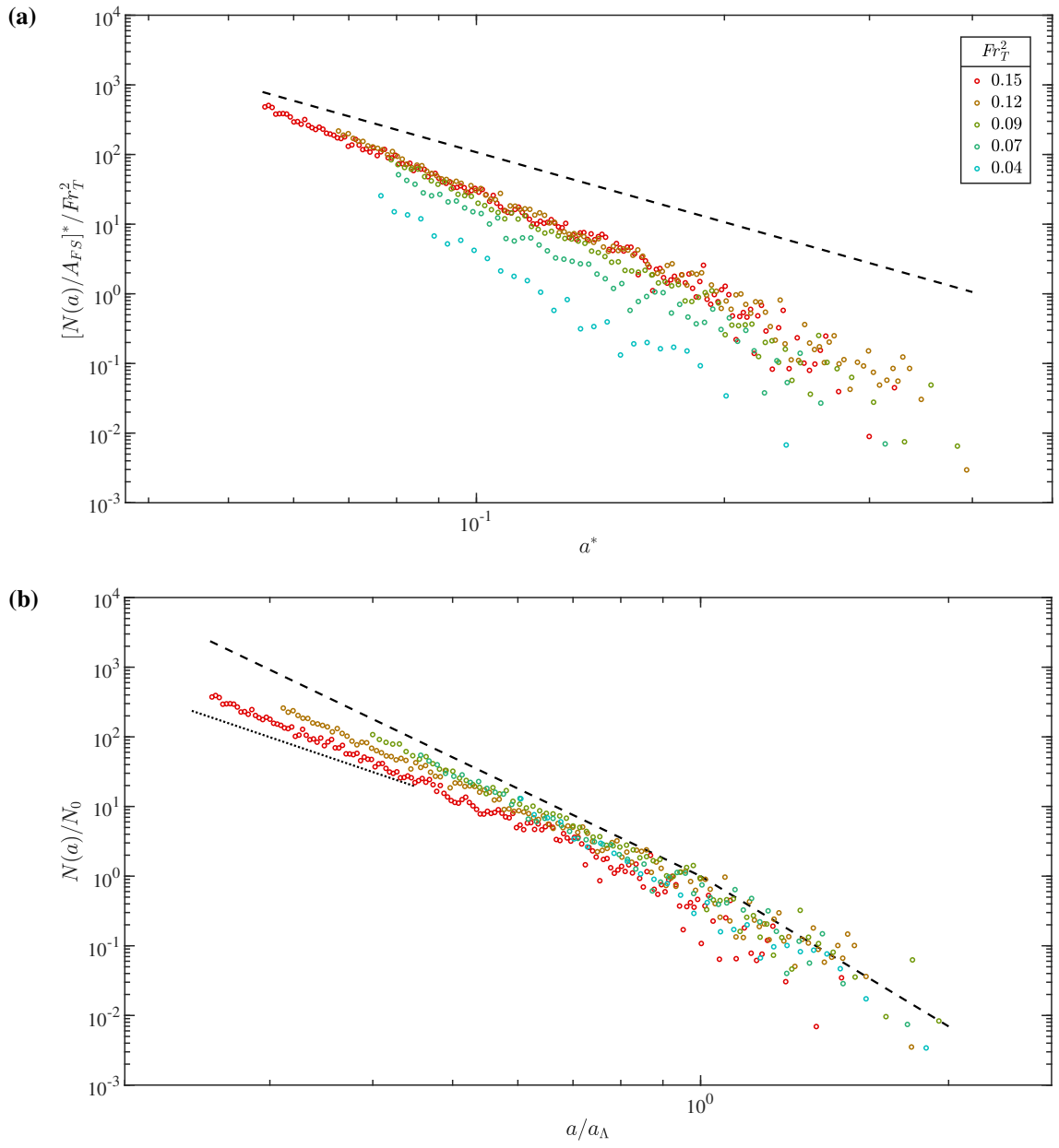


Figure 7-4: Bubble size distribution in free-surface shear flow compared to (a) the model by Yu *et al.* (2020) and (b) our model (7.24) for different turbulent Froude number $Fr_T^2 = \varepsilon/u_{\rm rms}\mathrm{g}$. In (a), $(--)$ shows $N(a)/A_{FS} \propto \mathrm{g}^{-1}\varepsilon^{2/3}a^{-10/3}$. In (b), $(--)$ shows (7.24) with $C_I = 3.62$, $C_U = 3.0$ and $C_L = 0.219$ ($R^2 = 0.849$), and for the smallest bubbles $(\cdots)$ shows $N(a) \propto a^{-4}$ from (1.17) with $\gamma = -14/3$. Recall $a^* = a/L_T$ and $[N(a)/A_{FS}]^* = [N(a)/A_{FS}]\, L_T^3$, where $L_T = u_{\rm rms}^3/\varepsilon$.

(see Chapter 6) and does not consider degassing.

Direct measurement of degassing in the previous section showed this flow is degassing dominated. In figure 7-4b we compare our measured $N(a)$ to (7.24) we predict as the equilibrium solution to degassing-dominated bubble populations. Without introducing any additional fitting parameters or modifying previous ones ($C_I = 3.62$ obtained by direct measurement of $I(a)$ in Chapter 6; $C_U = 3.0$ and $C_L = 0.219$ obtained by direct measurement of $\Lambda(a)$ in this chapter), we see a strong agreement ($R^2 = 0.849$). We note that, for small bubbles where fragmentation is expected to have some effect (see figure 7-3), we do see some evidence for the beginning of a shift toward $N(a) \propto a^{-4}$ predicted by (1.17) for fragmentation-dominated bubble populations with $I(a) \propto a^{-14/3}$ (Gaylo *et al.*, 2021). Generally, comparing figure 7-4a and figure 7-4b, we conclude that, as expected based on the ELA measurements of $\Lambda(a)/\Omega(a)$, $N(a)$ is much better described by the degassing-dominated model.

7.4.2 Scaling with Froude number

We have derived the equilibrium solution (7.24) for bubble populations in air entraining flow which dominated by degassing rather than fragmentation and observed this distribution in DNS of canonical free-surface shear flow, which has been used to model ship wakes. We now explore the scaling of this bubble size distribution with turbulent Froude number Fr_T^2.

We start by expanding (7.24). Using (6.22), (7.17), and (7.19) we obtain

$$N(a)/A_{FS} = 2C_L C_I \begin{cases} C_U^{-1}\, \mathrm{g}^{-3} u_{\mathrm{rms}}^5 \varepsilon^{1/3}\, (2a)^{-17/3} & a < a_\Lambda \\ (3.64)^{3/2}\, \mathrm{g}^{-23/6} \nu_w^{2/3} u_{\mathrm{rms}}^6 \varepsilon^{1/3}\, (2a)^{-43/6} & a > a_\Lambda, \text{ DNS-scale } Re_W \\ (1.98)^{1/2}\, \mathrm{g}^{-7/2} u_{\mathrm{rms}}^6 \varepsilon^{1/3}\, (2a)^{-37/6} & a > a_\Lambda, \text{ Physical } Re_W \end{cases} . \tag{7.25}$$

which shows explicitly how the predicted degassing-dominated bubble size distribution depends on bubble radius a, gravity g, turbulence strength ε and u_{rms}, and (for DNS-scale moderate Re_W) the viscosity of water ν_w. Next, we nondimensionalize (7.25) using the characteristic length scale $L_T = u_{\mathrm{rms}}^3/\varepsilon$, $[N(a)/A_{FS}]^* = [N(a)/A_{FS}]L_T^3$, and $a^* = a/L_T$. Focusing on the large Re_W we expect in large-scale flows like that behind a ship, the non-dimensional value of a_Λ is

$$a_\Lambda^* = a_\Lambda/L_T = 0.99\, C_U^2 Fr_T^2 . \tag{7.26}$$

Thus, nondimensionalizing (7.25) for the case of large Re_W gives

$$[N(a)/A_{FS}]^* = 2C_L C_I \begin{cases} C_U^{-1} Fr_T^6\, (2a^*)^{-17/3} & a^* < 0.99 C_U^2 Fr_T^2 \\ (1.98)^{1/2} Fr_T^7\, (2a^*)^{-37/6} & a^* > 0.99 C_U^2 Fr_T^2 \end{cases} . \tag{7.27}$$

For the constants we can use $C_I = 3.62$ obtained by direct measurement of $I(a)$ in Chapter 6 and $C_U = 3.0$ and $C_L = 0.219$ obtained by direct measurement of $\Lambda(a)$ in section 7.3. This

gives

$$[N(a)/A_{FS}]^* = \begin{cases} 0.537\ Fr_T^6\ (2a^*)^{-17/3} & a^* < 17.2\ Fr_T^2 \\ 2.23\ \ Fr_T^7\ (2a^*)^{-37/6} & a^* > 17.2\ Fr_T^2 \end{cases}. \tag{7.28}$$

We see that the number of bubbles scales with Fr_T^6 for bubbles in the turbulence-driven degassing regime and Fr_T^7 for bubbles in the buoyancy-drive degassing regime. Compared to, for example, Yu *et al.* (2020) who predicted $[N(a)/A_{FS}]^* \propto Fr_T^2\,(a^*)^{-10/3}$ for all $a \gg a_c$, we see that, in addition to the bubble size distribution being much steeper, it is significantly more sensitive to turbulent Froude number.

7.5 Conclusion

We study degassing in air entraining flows, where the total volume of bubbles is increasing. Through incorporating degassing into the PBE, our goal is to characterize the bubble size distribution, $N(a)$, including its power-law slope β. In air entraining flow, entrainment is clearly important, while the other mechanism(s) that balances entrainment determines β. For plunging breaking waves at large *We*, fragmentation is the dominant balancing mechanism, giving $\beta = -10/3$ (Garrett *et al.* 2000; Deane & Stokes 2002, see also §1.3). In contrast, we find that for FST, specifically a canonical free-surface shear flow (Shen *et al.*, 1999; Yu *et al.*, 2019), degassing is the dominant balancing mechanism.

From the PBE, we derive metrics to determine the importance of degassing in air entraining flows. The ratio of degassing flux to entrainment flux, $\mathcal{D} = Q_D/Q_I$, broadly quantifies the relevance of degassing. More directly, the ratio of degassing rate to fragmentation rate, $\Lambda(a)/\Omega(a)$, determines whether degassing or fragmentation is dominant (with (7.5) giving the critical value of $\Lambda(a)/\Omega(a)$ where the two are equal). For FST, DNS gives $\mathcal{D} \approx 0.7$ over a broad range of air entraining Fr^2, consistent with degassing being a relevant mechanism; and $\Lambda(a)/\Omega(a) > O(1)$, indicating that degassing is dominant over fragmentation.

From the PBE, we show that at equilibrium degassing-dominated bubble populations depend on entrainment $I(a)$ and degassing rate $\Lambda(a)$,

$$N(a) \approx I(a)/\Lambda(a)\,. \tag{7.12}$$

Using a simple model of degassing, we derive (7.19) which gives the degassing rate $\Lambda(a)$ in two regimes, separated by a critical radius a_Λ (7.17). We find a turbulence-driven regime where $\Lambda(a) \propto a$ and a buoyancy-driven regime where $\Lambda(a) \propto a^{5/2}$ for moderate, DNS-scale bubble Reynolds number Re_b. DNS measurements of $\Lambda(a)$ confirm this split power-law scaling and obtain the two scaling coefficients. Based on this $\Lambda(a)$ as well as $I(a)$ from Chapter 6, (7.12) gives that that degassing-dominated $N(a)$ follow a split power law $\beta(a < a_\Lambda) = -5.\bar{6}$ and $\beta(a > a_\Lambda) = -7.1\bar{6}$ (for moderate Re_b), which is confirmed independently by DNS. We note that this model can be extended theoretically to large physical-scale Re_b, where the buoyancy-driven regime becomes $\Lambda(a) \propto a^{3/2}$ leading to $\beta = -6.1\bar{6}$.

We have identified that it is possible for an air entraining free-surface flow to be degassing dominated, even at large *We* when fragmentation is strongest. With ELA we have direct

access to measure the physical mechanism and can prove degassing dominance through the metrics $\mathcal{D}$ and $\Lambda(a)/\Omega(a)$. Where such access is not possible, degassing or fragmentation dominance can be inferred from the size distribution power-law slope(s) β. For plunging breaking waves, the wide agreement on $\beta = -10/3$ (Deike, 2022) suggests that such flows are generally fragmentation dominated. The free-surface shear flow we consider is one definitive example of an air entraining flow which is degassing rather than fragmentation dominated. For another possible example, Hendrickson *et al.* (2019) reported $\beta \in [-8, -5]$ in the converging-corner-wave regions of the flow behind a dry transom stern, indicative of the degassing-dominated flows considered here. Because fragmentation-dominated and degassing-dominated bubble populations give such distinct equilibrium bubble populations, determining if a flow is fragmentation- or degassing-dominated is critical to understanding the bubble population.

Chapter 8

Conclusion

Turbulent, air entraining bubbly flows are a feature of many free-surface flows in nature and engineering, and predicting the size distribution of bubbles is critical to modeling these flows. The framework for describing the evolution of the bubble size distribution, $N(a)$, is the population balance equation (PBE), which has terms describing each of the individual bubble evolution mechanisms: fragmentation, entrainment, and degassing. We study these evolution mechanisms in free-surface turbulence (FST), a ubiquitous feature of air entraining free surfaces. Broadly, the two challenges to predicting $N(a)$ are modeling the near surface turbulence, and modeling the bubble evolution mechanisms, which depend on the turbulence. Using DNS of FST, this thesis contributes to the fundamental understanding of both near surface turbulence and bubble evolution mechanisms. This understanding also informs the development of models for air entraining FST. We make progress towards a RANS model for the surface layer and are able to predict a new degassing-dominated regime, which is confirmed by DNS and distinct from the fragmentation-dominated regime often assumed for bubbly flow.

We close with a summary of the contributions of this thesis and discussion of the future research work these contributions point towards.

8.1 Thesis contributions

I Numerical methods for free-surface bubbly flows

Development of ELA, enabling direct measurement of bubble evolution mechanisms in FST (Chapter 3; Gaylo *et al.* 2022)

A previous barrier to understanding the individual bubble evolution mechanisms near air entraining free surfaces was the difficulty in measuring the mechanisms directly. For example, a previous dataset on entrainment had only 108 entrainment events (Wei *et al.*, 2019). Because every mechanism affects $N(a)$, attempts to infer an individual mechanism from only the behavior of $N(a)$ requires assumptions about all other mechanisms. Incorrect assumptions have led to incorrect conclusions about evolution mechanisms (Yu *et al.*, 2020).

Bubble tracking enables direct measurement of the evolution mechanisms, but previous Lagrangian methods (Chan *et al.*, 2021*a*; Gao *et al.*, 2021; Basak *et al.*, 2026) struggle near air entraining free surfaces, where bubble evolution is complex. Instead, we develop Eulerian label advection (ELA) which uses the flow field available from DNS to track the movement of air. This gives robust bubble tracking, independent of the complexity of the bubble evolution. Numerically, ELA is built upon cVOF (Weymouth & Yue, 2010) in a way that maintains volume conservation and minimizes computational cost.

ELA enables us to, for the first time, directly measure individual bubble evolution mechanisms near air entraining free surfaces in DNS. This allows us to obtain huge datasets for each evolution mechanism. For example, analysis of entrainment in Chapter 6 is based on $60,000$ entrainment events. These datasets allow detailed characterization of the statistics of each mechanism, particularly the scaling with bubble size and Froude number.

II Describing turbulence near air entraining free surfaces

Identification of the critical Froude number to obtain strong FST (Chapter 4; Gaylo & Yue 2025)

Yu *et al.* (2019) identify that for sufficient Froude number near-surface turbulence becomes isotropic, the defining feature of strong FST. In this work we determine a robust definition of a near-surface turbulent Froude number (squared), $Fr_T^2 = \varepsilon / u_{\text{rms}}\text{g}$, based on turbulence statistics ε and u_{rms} measured at the bottom of the surface layer. DNS measurements of turbulence isotropy across a range of Fr_T^2 show the transition to strong FST occurs at $Fr_T^2 \approx 0.1$. For $Fr_T^2 > 0.1$, near-surface turbulence is nearly perfectly isotropic (isotropy parameter $J \approx 0.95$), independent of any further change in Fr_T^2.

Characterizing the surface layer in strong FST (Chapter 4)

Brocchini & Peregrine (2001*b*) define the surface layer as the region where air and water are highly mixed and note the difficulty in modeling it. We develop a definition of the surface layer thickness δ_s based on the vertical derivative of the average fraction of water, γ, at the mean free-surface height $\bar{\eta}$,

$$\delta_s \equiv \frac{6}{\sqrt{2\pi}} \left(\left. \frac{\mathrm{d}\gamma}{\mathrm{d}z} \right|_{z=\bar{\eta}} \right)^{-1} . \tag{4.6}$$

Unlike previous definitions (Brocchini & Peregrine, 2001*b*; Hendrickson & Yue, 2019), this does not depend on the tail behavior of the γ distribution. The surface layer thickness defines a nondimensional depth, $z^* \equiv (z - \bar{\eta})/\delta_s$. Scaling by z^* collapses the distribution of γ across a wide range of Fr_T^2. The distribution follows logistic tail behavior, rather than Gaussian. Relevant turbulence statistics within the surface layer ($z^* \in [-0.5, 0.5]$) also collapse with z^* when appropriately nondimensionalized by u_{rms} and ε measured at the bottom of the surface layer ($z^* = -0.5$). Particularly relevant to turbulence modeling, we show that there is no direct effect of the free surface on turbulence beneath the surface layer ($z^* < -0.5$), especially for strong FST.

III Quantifying individual bubble evolution mechanisms in the PBE

Fragmentation – Determining when the memoryless assumption is valid (Chapter 5; Gaylo *et al.* 2023)

A central assumption for modeling fragmentation in the PBE is that the probability a bubble fragments over a given time interval is independent of previous intervals. Physically, we do not expect this memoryless assumption to be valid for short intervals immediately after the bubble was formed by previous fragmentation, as the bubble would still be highly distorted. We elucidate the relaxation timescale $\tau_r = C_r\,\varepsilon^{-1/3}a^{2/3}$, which gives the minimum time interval where the observed statistics of bubble fragmentation are reasonably consistent with the memoryless assumption. This informs selection of measurement intervals to ensure the observed fragmentation statistics are applicable to the PBE. We also find that, for all super-Hinze scale bubbles, $\tau_r \ll \tau_\ell$, the typical lifetime of a bubble, showing that the memoryless assumption in the PBE is valid.

Fragmentation – A new bound on daughter-size distributions (Chapter 5; Gaylo *et al.* 2023)

Fragmentation models need to describe the statistics of the daughter bubbles produced by fragmentation. These are $\bar{m}$, the number of daughters; and $f_V^*(v^*)$, the probability distribution function describing the volume ratio of daughter to parent, v^*. There is a great variety of models for these two statistics. Previously, Martínez-Bazán *et al.* (2010) showed that volume conservation constrains the relationship between the $n = 1$ moment of $f_V^*(v^*)$ and $\bar{m}$, as expressed in (5.4).

In fragmentation cascades, the evolution of the bubble population is characterized by the timescale τ_c, which gives the average time needed for air starting in the largest bubble to go through the cascade and reach the Hinze scale (Deike *et al.*, 2016; Qi *et al.*, 2020; Gaylo *et al.*, 2021). We determine the empirical value of τ_c from DNS, and also derive an analytic expression for τ_c in terms of $f_V^*(v^*)$ and $\bar{m}$. The relationship between empirical value of τ_c and the fragmentations statistics provides a new constraint on the relationship between the $n = 11/9$ moment of $f_V^*(v^*)$ and $\bar{m}$,

$$\bar{m}(a')\int_0^1 v^{*11/9} f_V^*(v^*;a')\,\mathrm{d}v^* = 1 - \left(C_\tau C_{\Omega,\infty}\right)^{-1}. \tag{5.41}$$

Our estimations of C_τ and $C_{\Omega,\infty}$ from DNS give 0.92 for the right side at large *We*.

Entrainment – Scaling of the large-bubble entrainment size distribution in FST (Chapter 6; Gaylo & Yue 2025)

We investigate the scaling of entrainment by FST. For strong FST where turbulence is isotropic ($Fr_T^2 > 0.1$), dimensional analysis shows that, for bubbles larger than the capillary scale ($a_c \approx 1.3$mm for air-water on Earth), the entrainment size distribution is governed by a single scaling parameter α and a single scaling constant C_I. We develop a mechanistic model of FST entrainment based on the observation that the minimum radius of curvature

of the proceeding surface disturbance is similar to the radius of the resulting bubble. This implies an inter-scale energy transfer that was missed by previous entrainment models (Yu *et al.*, 2020) and gives $\alpha = -3$. This corresponds to an entrainment size distribution (per unit free surface area)

$$I(a)/A_{FS} = C_I \, \mathrm{g}^{-3} \, \varepsilon^{7/3} \, (2a)^{-14/3} \quad \text{for } a \gg a_c \, . \tag{6.22}$$

We perform DNS across a wide range of Fr_T^2 of a flow selected to isolate FST entrainment and a more realistic free-surface shear flow. ELA directly measures $I(a)$ and gives $C_I \approx 3.62$. Excluding $Fr_T^2 \ll 0.1$, both flows agree well with our model ($R^2 = 0.990$ and 0.891). We also find evidence of $I(a) \propto a^{-14/3}$ in previous experimental observations of open-channel flow (Wei *et al.*, 2019). The ubiquity of FST suggests that our model for FST entrainment can be important to broad classes of air entraining flows.

Degassing – Scaling of the degassing rate in FST (Chapter 7; Gaylo *et al.* 2024)

To determine the scaling of degassing, we develop a simple model with two regimes separated by a critical radius a_Λ. For $a < a_\Lambda$ degassing is driven by turbulence, giving a characteristic rise velocity $U_\Lambda \propto u_{\mathrm{rms}}$. For $a > a_\Lambda$ degassing is driven by buoyancy, and we set U_Λ equal to the terminal rise velocity in quiescent flow, which depends on bubble Reynolds number Re_W. By a mechanistic argument we determine L_Λ, the characteristic depth a bubble is entrained to. Through $\Lambda(a) = U_\Lambda / L_\Lambda$, this model gives the degassing rate

$$\Lambda(a) = \Lambda_0 \begin{cases} (a/a_\Lambda) & a < a_\Lambda \\ (a/a_\Lambda)^{5/2} & a > a_\Lambda, \text{ DNS-scale } Re_W \\ (a/a_\Lambda)^{3/2} & a > a_\Lambda, \text{ Physical } Re_W \end{cases} . \tag{7.19a}$$

From DNS of free-surface shear flow at a wide range of Fr_T^2, ELA measures $\Lambda(a)$ and we determine the two scaling coefficients in our model. Despite the simplicity of our degassing model, we see a good agreement with the numerical results ($R^2 = 0.761$).

IV Advancements in modeling free-surface bubbly flow

Progress towards a model of near-surface turbulence for use in RANS (Chapter 4)

We observe that the direct effects of the free surface are restricted to the surface layer and that the surface layer is well characterized by (only) the turbulence properties at its bottom. This shows that, rather than attempting to model k and ε within the surface layer, a RANS simulation could simply apply an appropriate boundary condition at the bottom edge of the surface layer. Towards implementation of this approach, we characterize the two quantities necessary for such a boundary condition in k-ε RANS: the surface layer thickness,

$$\delta_s = C_\delta \, u_{\mathrm{rms}}^2 \mathrm{g}^{-1} \, ; \tag{4.35}$$

and the energy flux into the surface layer,

$$W/\rho_w = (C_W/C_\delta)\,\varepsilon\,\delta_s\,. \tag{4.36}$$

DNS measurements give $C_\delta \approx 11.1$ and $C_W \approx 4.6$. $C_W/C_\delta \sim 1/2$ is consistent with the observation that the turbulent dissipation rate (per unit mass) is roughly constant within the surface layer.

A distinct equilibrium solution for bubble populations dominated by degassing, like in FST (Chapter 7; Gaylo *et al.* 2024)

For bubble populations dominated by fragmentation, $N(a) \propto a^{-10/3}$ is the equilibrium solution (Garrett *et al.*, 2000). This power law is often observed in plunging breaking waves for super Hinze-scale bubbles (Deane & Stokes, 2002; Deike *et al.*, 2016). From DNS of free-surface shear flow at a wide range or Fr_T^2, we compare the degassing rate $\Lambda(a)$ to fragmentation rate $\Omega(a)$. We identify that $\Lambda(a) \gg \Omega(a)$, meaning degassing is dominant, rather than fragmentation. For bubble populations dominated by degassing, we show the equilibrium solution is

$$N(a) \approx I(a)/\Lambda(a)\,. \tag{7.12}$$

With the degassing rate $\Lambda(a)$ and the entrainment size distribution $I(a)$ we obtained based on ELA, this allows us to predict the bubble size distribution. Without any additional fitting parameters, our prediction agrees very well with the $N(a)$ we measure (independent of ELA) in DNS ($R^2 = 0.849$).

For large-scale flows (very large Reynolds numbers), degassing-dominance predicts a bubble size distribution (nondimensionalized by the turbulence length scale $L_T = u_{\mathrm{rms}}^3/\varepsilon$)

$$[N(a)/A_{FS}]^* = \begin{cases} 0.537\ Fr_T^6\ \ (2a^*)^{-17/3} & a^* < 17.2\ Fr_T^2 \\ 2.23\ \ \ Fr_T^7\ \ (2a^*)^{-36/6} & a^* > 17.2\ Fr_T^2 \end{cases}, \tag{7.28}$$

which applies to bubbles larger than the capillary scale ($a \gg a_c$). For the scaling with bubble radius a, we see a split distribution, and in both regimes the power-law slope is more negative (i.e., there are fewer large bubbles) than $N(a) \propto a^{-10/3}$. We also see that the bubble population is incredibly sensitive to Froude number; in the turbulence-driven degassing regime $N(a)$ scales with Fr_T^6, and is even more sensitive in the buoyancy-driven degassing regime.

As an illustration of the insight provided by this new equilibrium solution, we consider the flow behind the transom stern of a vessel, which the canonical free-surface shear flow models. Previous work (Yu *et al.*, 2020) predicted that $N(a)$ scales with Fr_T^2, meaning that when the speed of the vessel doubles, the number of bubbles increases by a factor of 4. We show the bubble population actually scales with Fr_T^6, meaning that when the speed of the vessel doubles, the number of bubbles increases by a factor of 64.

8.2 Future work

Implementation of the surface-layer model in RANS

Our DNS measurements of FST suggest that the effects of an air entraining free surface on turbulence could be modeled in RANS using a boundary condition applied at the bottom edge of the surface layer. For the forced FST setup, where turbulence levels deep beneath the surface are prescribed and mean values are only a function of z, it would be straight forward to solve the k-ε RANS equations with our proposed boundary condition and compare the predicted k and ε to the DNS results in Chapter 4. For a flow which includes realistic turbulence production, our DNS of free-surface shear flow provides a useful reference to test a RANS model against.

Elucidating the scaling of small-bubble entrainment by FST

Our entrainment model predicts $I(a) \propto a^{-14/3}$ for bubbles larger than the capillary scale. This scaling cannot apply to arbitrarily small bubbles, as that would imply an infinite entrainment flux, $Q_I = (4\pi/3) \int I(a) a^3 \, \mathrm{d}a$. To predict the entrainment flux, the scaling of the entrainment size distribution for bubbles smaller than the capillary scale is necessary. In theory the mechanistic model developed in Chapter 6 could be extended to include surface tension; however, the relationship between turbulence and bubble scales becomes multi-valued, suggesting a more complicated inter-scale energy exchange.

Describing the effects of coalescence

For small and moderate void fractions with significant space between bubbles, the effect of coalescence on the bubble population is negligible. However, our results suggest the quantity of bubbles scales with Fr_T^6, meaning the void fraction will increase significantly with only slightly larger Froude numbers. ELA can quantify coalescence; however, coalescence involves thin films on scales not resolved by our DNS. Our DNS solver would require explicit models for coalescence, otherwise the VOF scheme implicitly coalesces any bubbles with interfaces within a grid cell.

Investigating non-equilibrium regimes

For degassing-dominated bubble populations, we have elucidated the equilibrium ($\partial N/\partial t = 0$) solution $N(a) = I(a)/\Lambda(a)$. The natural question then is how the bubble population evolves in non-equilibrium. As noted in Chapter 7 the degassing-dominated PBE is a set of independent ordinary differential equations. For the simple case where $N(a) = 0$ at $t < 0$ and $I(a)$ and $\Lambda(a)$ are steady for $t > 0$,

$$N(a;t) = [I(a)/\Lambda(a)] \left(1 - \mathrm{e}^{-\Lambda(a)t}\right). \tag{8.1}$$

For non-steady $I(a)$ or $\Lambda(a)$, or non-negligible fragmentation the evolution is more complex. The evolution of the total entrained volume, $\mathrm{d}V/\mathrm{d}t$ given by (1.8), is also of interest, but requires the entrainment flux Q_I, which requires the scaling of small-bubble entrainment.

Appendix A

Stability of the Viscous Diffusion Term with Arithmetic-Mean Viscosity

For single-phase flows the stability of the viscous diffusion term leads to a timestep restriction

$$\Delta t < \frac{1}{6} Re \, \min \left[\Delta x_d\right]^2 \, (\rho/\mu) \tag{2.11}$$

Tryggvason *et al.* (2011, §3.1). For multi-phase flows, this appendix addresses the appropriate value to use for ρ/μ.

The appropriate value for ρ/μ is closely tied to the methods used to interpolate density ρ and viscosity μ from the pressure and VOF grid where they are explicitly defined by (2.15), to calculations on the velocity grid. In this section we will use the u-grid as an example, and by swamping indices the equations can be trivially modified to apply to the v- and w-grids. The DNS solver used in this thesis uses

$$\rho_{i+1/2\,j\,k} = \frac{\rho_{i\,j\,k} + \rho_{i+1\,j\,k}}{2} \tag{A.1}$$

for density and the same for viscosity,

$$\mu_{i+1/2\,j\,k} = \frac{\mu_{i\,j\,k} + \mu_{i+1\,j\,k}}{2} \,. \tag{A.2}$$

However, the viscous diffusion term requires additional interpolation,

$$\mu_{i+1/2\,j+1/2\,k} = \frac{\mu_{i+1/2\,j\,k} + \mu_{i+1/2\,j+1\,k}}{2} \tag{A.3}$$

(and similar for $\mu_{i+1/2\,j\,k+1/2}$). This expression is equivalent to the arithmetic mean (2.17) discussed in section 2.1.3. The extra interpolation required for viscosity compared to density is the reason that simply choosing the minimum of ρ_w/μ_w and ρ_a/μ_a for ρ/μ in (2.11) is insufficient to guarantee stability.

A.1 Spatial discretization of the viscous diffusion term

We first review the calculation of the viscous term, as described by Tryggvason *et al.* (2011, §3.1). For the u velocity, the viscous diffusion term in (2.5) is

$$\frac{\partial}{\partial t} u_{i+1/2\,j\,k} = \cdots + \frac{1}{\rho_{i+1/2\,j\,k}} \left(D_x\right)_{i+1/2,j,k} + \ldots \tag{A.4}$$

For a constant grid $\Delta = \Delta x = \Delta y = \Delta z$, the viscous term is given in terms of the stress matrix

$$\begin{aligned}\left(D_x\right)_{i+1/2\,j\,k} = \frac{T^{xx}_{i+1\,j\,k} - T^{xx}_{i\,j\,k}}{\Delta} \\ + \frac{T^{xz}_{i+1/2\,j+1/2\,k} - T^{xz}_{i+1/2\,j-1/2\,k}}{\Delta} \\ + \frac{T^{xy}_{i+1/2\,j\,k+1/2} - T^{xy}_{i+1/2\,j\,k-1/2}}{\Delta}.\end{aligned} \tag{A.5}$$

The stress matrix is calculated

$$T^{xx}_{i\,j\,k} = \frac{1}{Re} 2\mu_{i\,j\,k} \frac{u_{i+1/2\,j\,k} - u_{i-1/2\,j\,k}}{\Delta}, \tag{A.6a}$$

$$T^{xz}_{i+1/2\,j+1/2\,k} = \frac{1}{Re} \mu_{i+1/2\,j+1/2\,k} \left(\frac{u_{i+1/2\,j+1\,k} - u_{i+1/2\,j\,k}}{\Delta} + \frac{w_{i+1\,j+1/2\,k} - w_{i\,j+1/2\,k}}{\Delta} \right), \tag{A.6b}$$

$$T^{xy}_{i+1/2\,j\,k+1/2} = \frac{1}{Re} \mu_{i+1/2\,j\,k+1/2} \left(\frac{u_{i+1/2\,j\,k+1} - u_{i+1/2\,j\,k}}{\Delta} + \frac{v_{i+1\,j\,k+1/2} - v_{i\,j\,k+1/2}}{\Delta} \right), \tag{A.6c}$$

and similar for $T^{xy}_{i+1/2\,j\,k+1/2}$.

A.2 Linear stability analysis

Before addressing the viscous term specifically, we start with a general analysis of the effects of the temporal discretization. We will consider the evolution of the spatially discrete u-velocity, $u_{i+1/2\,j\,k}$ (see figure 2-1). In continuous time but discrete space, for linear stability analysis our interest is the value of C in the linear term that multiplies $u_{i+1/2\,j\,k}$,

$$\frac{\partial}{\partial t} u_{i+1/2\,j\,k} = C\, u_{i+1/2\,j\,k} + \cdots. \tag{A.7}$$

Applying the two-stage Runge-Kutta time discretization described in section 2.1.2,

$$u^{(n+1/2)}_{i+1/2\,j\,k} = (1 + \Delta t C)\; u^{(n)}_{i+1/2\,j\,k} + \ldots, \tag{A.8a}$$

$$u^{(n+1)}_{i+1/2\,j\,k} = \frac{1}{2} u^{(n)}_{i+1/2\,j\,k} + \left(\frac{1}{2} + \frac{\Delta t}{2} C \right)\; u^{(n+1/2)}_{i+1/2\,j\,k} + \ldots. \tag{A.8b}$$

Combining these two steps gives

$$u^{(n+1)}_{i+1/2\,j\,k} = \left(1 + (\Delta tC) + \frac{1}{2}(\Delta tC)^2\right) u^{(n)}_{i+1/2\,j\,k} + \dots . \tag{A.9}$$

The linear stability requirement is

$$-2 < (\Delta tC) + \frac{1}{2}(\Delta tC)^2 < 0 \tag{A.10}$$

The lower bound is always satisfied (i.e., there are no oscillatory linear instabilities). Expanding the upper bound,

$$-2 < \Delta tC < 0\,. \tag{A.11}$$

We now seek the equivalent value of C in (A.4). Because the viscous term is linear, we assume a solution of the form

$$u_{i+1/2,j,k} = U\mathrm{e}^{\mathrm{i}\left(\kappa_x i + \kappa_z j + \kappa_y k\right)}\,. \tag{A.12}$$

This gives, for example, $u_{i+1/2,j+1,k} = \mathrm{e}^{\mathrm{i}\kappa_z} u_{i+1/2,j,k}$. We put (A.12) into (A.5) and, after a fair bit of algebra, we obtain (A.18) shown on the next page.

The stability criteria for the viscous term can be made looser by applying continuity (2.8a), which gives us that the velocity field is divergence free. Discretized on the staggered grid, this gives

$$\frac{w_{i\,j+1/2\,k} - w_{i\,j-1/2\,k}}{\Delta} + \frac{v_{i\,j\,k+1/2} - v_{i\,j\,k-1/2}}{\Delta} = -\frac{u_{i+1/2\,j\,k} - u_{i-1/2\,j\,k}}{\Delta}\,. \tag{A.13}$$

As shown in (A.19), this simplifies the third and fourth lines of (A.18) into a term which partially cancels a term in the first line of (A.18). Finally, we obtain (A.20), shown on the next page.

A.2.1 Single-phase stability criteria

To check (A.20), we start by considering the case where μ and ρ are constant throughout the domain. In this case, we expect to obtain (2.11). For constant μ and ρ, (A.20) simplifies to

$$(D_x)_{i+1/2\,j\,k} = \frac{\mu}{Re\Delta^2}\left\{2\left[\cos\kappa_x - 1\right] + 2\left[\cos\kappa_z - 1\right] + 2\left[\cos\kappa_y - 1\right]\right\} u_{i+1/2,j,k}\,. \tag{A.14}$$

The most unstable mode is $\kappa_x = \kappa_z = \kappa_y = \pi$, which gives

$$(D_x)_{i+1/2\,j\,k} = -12\frac{\mu}{Re\Delta^2} u_{i+1/2\,j\,k}\,. \tag{A.15}$$

Recalling that we also divide by ρ in (A.4), we obtain

$$C = -\frac{12}{Re\Delta^2}\,\nu\,, \tag{A.16}$$

$$
\begin{aligned}
Re\,(D_x)_{i+1/2\,j\,k} = &\left\{\frac{4\mu_{i+1/2\,j\,k}}{\Delta^2}\left[\cos\kappa_x - 1\right] + \frac{\mu_{i+1/2\,j+1\,k}+2\mu_{i+1/2\,j\,k}+\mu_{i+1/2\,j-1\,k}}{2\Delta^2}\left[\cos\kappa_z - 1\right] + \frac{\mu_{i+1/2\,j\,k+1}+2\mu_{i+1/2\,j\,k}+\mu_{i+1/2\,j\,k-1}}{2\Delta^2}\left[\cos\kappa_y - 1\right]\right\} u_{i+1/2\,j\,k} \\
&+ \mathrm{i}\left\{\frac{\mu_{i+1\,j\,k}-\mu_{i\,j\,k}}{\Delta^2}\sin\kappa_x + \frac{\mu_{i+1/2\,j+1\,k}-\mu_{i+1/2\,j-1\,k}}{2\Delta^2}\sin\kappa_z + \frac{\mu_{i+1/2\,j\,k+1}-\mu_{i+1/2\,j\,k-1}}{2\Delta^2}\sin\kappa_y\right\} u_{i+1/2\,j\,k} \\
&+ \mu_{i+1/2\,j\,k}\left[\frac{w_{i+1\,j+1/2\,k}-w_{i+1\,j-1/2\,k}}{\Delta^2} - \frac{w_{i\,j+1/2\,k}-w_{i\,j-1/2\,k}}{\Delta^2}\right] \\
&+ \mu_{i+1/2\,j\,k}\left[\frac{v_{i+1\,j\,k+1/2}-v_{i+1\,j\,k-1/2}}{\Delta^2} - \frac{v_{i\,j\,k+1/2}-v_{i\,j\,k-1/2}}{\Delta^2}\right] \\
&+ \frac{\mu_{i+1/2\,j+1\,k}-\mu_{i+1/2\,j\,k}}{\Delta}\left(\frac{w_{i+1\,j+1/2\,k}-w_{i\,j+1/2\,k}}{2\Delta}\right) + \frac{\mu_{i+1/2\,j\,k}-\mu_{i+1/2\,j-1\,k}}{\Delta}\left(\frac{w_{i+1\,j-1/2\,k}-w_{i\,j-1/2\,k}}{2\Delta}\right) \\
&+ \frac{\mu_{i+1/2\,j\,k+1}-\mu_{i+1/2\,j\,k}}{\Delta}\left(\frac{v_{i+1\,j\,k+1/2}-v_{i\,j\,k+1/2}}{2\Delta}\right) + \frac{\mu_{i+1/2\,j\,k}-\mu_{i+1/2\,j\,k-1}}{\Delta}\left(\frac{v_{i+1\,j\,k-1/2}-v_{i\,j\,k-1/2}}{2\Delta}\right)
\end{aligned}
\tag{A.18}
$$

$$
\begin{aligned}
&\frac{\mu_{i+1/2\,j\,k}}{\Delta^2}\left[\left(w_{i+1\,j+1/2\,k} - w_{i+1\,j-1/2\,k}\right) - \left(w_{i\,j+1/2\,k} - w_{i\,j-1/2\,k}\right) + \left(v_{i+1\,j\,k+1/2} - v_{i+1\,j\,k-1/2}\right) - \left(v_{i\,j\,k+1/2} - v_{i\,j\,k-1/2}\right)\right] \\
&\qquad = -2\frac{\mu_{i+1/2\,j\,k}}{\Delta^2}\left[\cos\kappa_x - 1\right] u_{i+1/2\,j\,k}
\end{aligned}
\tag{A.19}
$$

$$
\begin{aligned}
Re\,(D_x)_{i+1/2\,j\,k} = &\left\{\frac{2\mu_{i+1/2\,j\,k}}{\Delta^2}\left[\cos\kappa_x - 1\right] + \frac{\mu_{i+1/2\,j+1\,k}+2\mu_{i+1/2\,j\,k}+\mu_{i+1/2\,j-1\,k}}{2\Delta^2}\left[\cos\kappa_z - 1\right] + \frac{\mu_{i+1/2\,j\,k+1}+2\mu_{i+1/2\,j\,k}+\mu_{i+1/2\,j\,k-1}}{2\Delta^2}\left[\cos\kappa_y - 1\right]\right\} u_{i+1/2\,j\,k} \\
&+ \mathrm{i}\left\{\frac{\mu_{i+1\,j\,k}-\mu_{i\,j\,k}}{\Delta^2}\sin\kappa_x + \frac{\mu_{i+1/2\,j+1\,k}-\mu_{i+1/2\,j-1\,k}}{2\Delta^2}\sin\kappa_z + \frac{\mu_{i+1/2\,j\,k+1}-\mu_{i+1/2\,j\,k-1}}{2\Delta^2}\sin\kappa_y\right\} u_{i+1/2\,j\,k} \\
&+ \frac{\mu_{i+1/2\,j+1\,k}-\mu_{i+1/2\,j\,k}}{\Delta}\left(\frac{w_{i+1\,j+1/2\,k}-w_{i\,j+1/2\,k}}{2\Delta}\right) + \frac{\mu_{i+1/2\,j\,k}-\mu_{i+1/2\,j-1\,k}}{\Delta}\left(\frac{w_{i+1\,j-1/2\,k}-w_{i\,j-1/2\,k}}{2\Delta}\right) \\
&+ \frac{\mu_{i+1/2\,j\,k+1}-\mu_{i+1/2\,j\,k}}{\Delta}\left(\frac{v_{i+1\,j\,k+1/2}-v_{i\,j\,k+1/2}}{2\Delta}\right) + \frac{\mu_{i+1/2\,j\,k}-\mu_{i+1/2\,j\,k-1}}{\Delta}\left(\frac{v_{i+1\,j\,k-1/2}-v_{i\,j\,k-1/2}}{2\Delta}\right)
\end{aligned}
\tag{A.20}
$$

where $\nu \equiv \mu/\rho$. Putting this value of C into (A.11) we obtain,

$$\Delta t \leq \frac{1}{6} Re\, \Delta^2\, \nu^{-1}\,, \tag{A.17}$$

which, as expected, is equivalent to (2.11).

A.2.2 Multi-phase stability criteria

We now consider the same most unstable mode ($\kappa_x = \kappa_z = \kappa_y = \pi$), but now for a multi-phase flow where μ and ρ change across cells. Grouping the terms of $(D_x)_{i+1/2,j,k}$ that multiply $u_{i+1/2,j,k}$, (A.20) becomes

$$(D_x)_{i+1/2\,j\,k} = -\frac{1}{Re}\left\{\frac{8\mu_{i+1/2\,j\,k}+\mu_{i+1/2\,j+1\,k}+\mu_{i+1/2\,j-1\,k}+\mu_{i+1/2\,j\,k+1}+\mu_{i+1/2\,j\,k-1}}{\Delta^2}\right\} u_{i+1/2\,j\,k} + \ldots\,. \tag{A.21}$$

To make this similar to (A.15), we can define an effective dynamic viscosity

$$\tilde{\mu}_{i+1/2\,j\,k} \equiv \frac{8\mu_{i+1/2\,j\,k} + \mu_{i+1/2\,j+1\,k} + \mu_{i+1/2\,j-1\,k} + \mu_{i+1/2\,j\,k+1} + \mu_{i+1/2\,j\,k-1}}{12}\,, \tag{A.22}$$

and write

$$(D_x)_{i+1/2\,j\,k} = -12\frac{\tilde{\mu}_{i+1/2\,j\,k}}{Re\Delta^2} u_{i+1/2\,j\,k} + \ldots\,. \tag{A.23}$$

Following the same steps as the single-phase criteria, we obtain

$$\Delta t \leq \frac{1}{6} Re\, \Delta^2\, \tilde{\nu}^{-1}\,, \tag{A.24}$$

where the effective (kinematic) viscosity for the u-grid is defined

$$\tilde{\nu}_{i+1/2\,j\,k} \equiv \frac{2}{3}\left[\frac{\mu_{i+1/2\,j\,k}}{\rho_{i+1/2\,j\,k}}\right] + \frac{1}{12}\left[\frac{\mu_{i+1/2\,j+1\,k} + \mu_{i+1/2\,j-1\,k} + \mu_{i+1/2\,j\,k+1} + \mu_{i+1/2\,j\,k-1}}{\rho_{i+1/2\,j\,k}}\right]. \tag{A.25}$$

Equivalent definitions for $\tilde{\nu}_{i\,j+1/2\,k}$ and $\tilde{\nu}_{i\,j\,k+1/2}$ can be obtained by swapping indices.

A.3 Bounding effective viscosity

Before seeking the maximum possible value of $\tilde{\nu}$, we illustrate that it could be larger than the $\nu = \mu/\rho$ of either individual fluid. Consider an interface which is normal to z. The VOF field is

$$f_{i\,j\,k} = \begin{cases} 0 & j \geq 1 \\ 1 & j < 1 \end{cases}.$$

Using (2.15), we obtain (nondimensionalized) density and viscosity

$$\rho_{i\,j\,k} = \begin{cases} \lambda & j \geq 1 \\ 1 & j < 1 \end{cases}, \qquad \mu_{i\,j\,k} = \begin{cases} \eta & j \geq 1 \\ 1 & j < 1 \end{cases}.$$

Using (A.25) we obtain

$$\tilde{\nu}_{i+1/2\,j\,k} = \begin{cases} (\eta/\lambda) & j \geq 2 \\ \frac{11}{12}(\eta/\lambda) + \frac{1}{12}(1/\lambda) & j = 1 \\ \frac{11}{12}(1/1) + \frac{1}{12}(\eta/1) & j = 0 \\ (1/1) & j \leq -1 \end{cases}.$$

For air-water ($\lambda = 0.00123$, $\eta = 0.0159$), this corresponds to $\tilde{\nu}_{i+1/2\,j\,k} \approx 80$ at $j = 1$, much larger than either $\nu = \lambda/\eta \approx 13$ in air or $\nu = 1$ in water. As discovered by Matthew Coogan (personal communication, October 2025), DNS of a nearly quiescent air-water free surface confirms that, if one uses $\tilde{\nu} \sim \lambda/\eta$ in (A.24) there is a numerical instability and, as expected, the u velocity in the first cell of air above the interface grows exponentially.

As illustrated in the example, the viscosity in $j \pm 1$ and $k \pm 1$ cells are not necessarily related to the density in the cell at $j\,k$. This means that in general the second term in (A.25) can, at best, be bounded by

$$\frac{1}{12}\left[\frac{\mu_{i+1/2\,j-1\,k} + \mu_{i+1/2\,j+1\,k} + \mu_{i+1/2\,j\,k-1} + \mu_{i+1/2\,j\,k+1}}{\rho_{i+1/2\,j\,k}}\right] \leq \frac{1}{3}\frac{\max\left[1,\eta\right]}{\min\left[1,\lambda\right]}. \tag{A.26}$$

For the first term, the fact that the same interpolation is used for $\mu_{i+1/2\,j\,k}$ as $\rho_{i+1/2\,j\,k}$ does mean their quotient is bounded between the μ/ρ of each individual fluid. In summary, (A.25) can be bounded by

$$\tilde{\nu} \leq \frac{2}{3}\max\left[1,\frac{\eta}{\lambda}\right] + \frac{1}{3}\frac{\max\left[1,\eta\right]}{\min\left[1,\lambda\right]}. \tag{A.27}$$

A.4 Implementation options

Applying (A.27) to (A.24), one obtains a timestep restriction

$$\Delta t \leq \frac{1}{6}\,Re\,\Delta^2\left(\frac{2}{3}\max\left[1,\frac{\eta}{\lambda}\right] + \frac{1}{3}\frac{\max\left[1,\eta\right]}{\min\left[1,\lambda\right]}\right)^{-1}. \tag{A.28}$$

For air and water where $\eta \ll 1$, this condition is about 20 times more restrictive compared to (incorrectly) using $\tilde{\nu} \sim \eta/\lambda$ in (A.24). For many of the simulations in this thesis, this would make (A.28) the most restrictive timestep criteria, more restrictive than (2.13) driven by CVOF.

In practice, it is common to apply some smoothing to the VOF field before calculating ρ and μ (e.g., Tryggvason *et al.*, 2011, §7.1.4), which decreases the sharp changes in μ that make such large $\tilde{\nu}$ possible. In theory one could determine how smoothing links viscosity in

adjacent cells and derive a tighter bound on $\tilde{\nu}$ for when smoothing is used. Alternatively, one could calculate the true maximum value of $\tilde{\nu}$ before each timestep using (A.25) (and similar equations for $\tilde{\nu}_{i\,j+1/2\,k}$ and $\tilde{\nu}_{i\,j\,k+1/2}$), and apply the dynamic timestep restriction

$$\Delta t \leq \frac{1}{6} Re\, \Delta^2 \ (\max\,[\tilde{\nu}])^{-1} \ . \tag{A.29}$$

For a third approach, we first note that due to smoothing of the VOF field as well as turbulence (physically) smoothing the VOF field, in many simulations it is rare to see the strong viscosity gradients that lead to $\tilde{\nu} \gg \max[1, \eta/\lambda]$. While not guaranteed to be stable, in practice one could use a safety factor FS and set

$$\Delta t \leq \frac{1}{6} Re\, \Delta^2 \ (\mathrm{FS}\ \max[1, \eta/\lambda])^{-1} \ . \tag{A.30}$$

The simulations in this thesis use FS = 1 (i.e., no safety factor), and only rarely become unstable. However, in general the author suggests using (A.29).

Appendix B

Approximating Long Snapshot Intervals using Volume Tracking Matrix Multiplication

As introduced in §3.2.1, the volume tracking matrix (VTM) $\mathbf{A}^{(n\to n+1)}$ describes the movement of air between bubbles with volumes $\mathbf{v}^n$ at time t^n and bubbles with volumes $\mathbf{v}^{n+1}$ at time $t^{n+1} = t^n + \Delta t_s$

$$\mathbf{v}^{n+1} = \mathbf{A}^{(n\to n+1)}\mathbf{v}^n\,, \tag{3.11}$$

where Δt_s is the snapshot interval. The VTM is obtained from the vector color function $\mathbf{c}^n$, which is evolved using ELA, a method to solve

$$\frac{\partial \mathbf{c}^n}{\partial t} + \mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{c}^n = \mathbf{0} \tag{3.2}$$

in the discrete form

$$\mathbf{s}^n_{ijk}(t) \equiv \frac{\int_{\Omega_{ijk}} \mathbf{c}^n(\mathbf{x},t)\ \mathrm{d}V}{\Delta\Omega_{ijk}}\,. \tag{3.19}$$

While we do not find the memory or computational cost of ELA prohibitive in practice, in this appendix we discuss matrix multiplication as a way to, if desired, decrease the cost with a trade off in accuracy.

B.1 Memory cost of Eulerian label advection

While there is limited additional computational effort required for ELA due to the reuse of the flux information from cVOF (see §3.4.2), there is a memory cost related to storing the source vector $\mathbf{s}^n$. $\mathbf{s}^n$ is generally very sparse, so rather than attempting to store the entire length for each cell, we store the index and value $\{\ell, s_\ell\}$ of each non-zero entry of $\mathbf{s}^n$ in a cell. The same approach applies to storing $\tilde{\mathbf{c}}$ during advection (see (3.25)). Thus, the memory cost scales with the number of non-zero entries in each cell, $\mathrm{nnz}[\mathbf{s}^n_{ijk}]$.

To bound $\mathrm{nnz}[\mathbf{s}^n_{ijk}]$, we first note the total size of $\mathbf{s}^n_{ijk}$ is equal to the number of bubbles at t^n, M^n. Assuming the CCL method used will not generate more than one label per

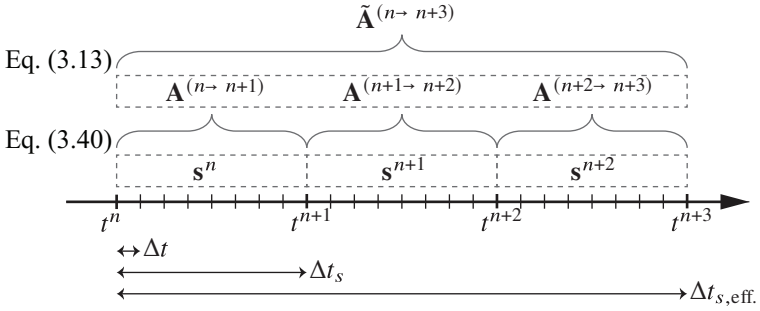


Figure B-1: Illustration of the flow of information using ELA and matrix multiplication with $K = 8$ and $N = 3$. After using ELA over K simulation time steps, the information in $\mathbf{s}^n$ is summarized in a VTM using (3.40). The information in N of these VTMs is summarized in an effective VTM using (3.13).

cell and noting that the Courant restriction limits advection of $\mathbf{s}^n$ to immediate neighbors each simulation time step Δt, over a snapshot interval involving K simulations time steps, $\Delta t_s = K\Delta t$,

$$\text{nnz}(\mathbf{s}^n_{ijk}) \leq \min\left\{(1+2K)^{\mathcal{N}}\,, \quad M^n\right\}\,, \tag{B.1}$$

where $\mathcal{N}$ is the dimensions of the simulation. Although the scaling in our simulations is less (Gaylo *et al.*, 2022, Fig. 12), the upper bound suggests that memory requirements scale with K^3, or Δt_s^3, for a three-dimensional simulation. Because numerical limits (see §3.5.3) as well as physical reasons (see §5.6) typically impose a minimum desirable value of Δt_s, there may be a desire to approximate long Δt_s without increasing K, which can be done through VTM multiplication.

B.2 Source of error with VTM multiplication

As mentioned in §3.2.1, the left-stochastic nature of the VTM means that over N snapshot intervals,

$$\mathbf{v}^{n+N} = \left[\prod_{m=n+N-1}^{n} \mathbf{A}^{(m\to m+1)}\right]\mathbf{v}^n\,, \tag{3.12}$$

Which defines an effective VTM

$$\tilde{\mathbf{A}}^{(n\to n+N)} \equiv \prod_{m=n+N-1}^{n} \mathbf{A}^{(m\to m+1)}\,, \tag{3.13}$$

with an effective snapshot interval $\Delta t_{s,\text{eff.}} = N\Delta t_s$. Therefore,

$$\Delta t_{s,\text{eff.}} = NK\Delta t\,. \tag{B.2}$$

With Δt constrained by the DNS solver, the effective snapshot interval can be adjusted through the true snapshot interval during the simulation (K) or through multiplying the matrices in post-processing (N), as illustrated in figure B-1.

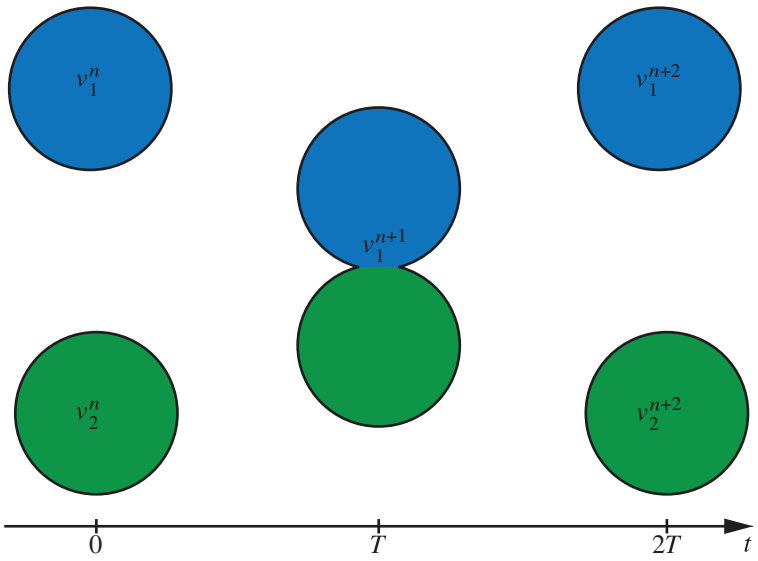


Figure B-2: Illustration of two bubbles at $t = 0$ whose dark fluid briefly coalesces into one bubble at $t = T$ but then fragments along the boundary between the two original bubbles such that at $t = T$ all of the volume from one original bubble is in only one of the final bubbles.

Although the effective VTM describes the same transition from $\mathbf{v}^n$ to $\mathbf{v}^{n+N}$, we note that there is a loss of information when using matrix multiplication versus a larger K. Through the vector color function $\mathbf{c}^n$, the source of each particle of air is known. However, when $\mathbf{A}$ is calculated using (3.5), we integrate $\mathbf{c}^n$ over the volume of each bubble, losing the spatial distribution of $\mathbf{c}^n$ within the bubble. In effect, matrix multiplication assumes that after each true snapshot interval Δt_s, $\mathbf{c}^n$ is homogeneous within a bubble, increasing the apparent entropy. This is consistent with the stochastic interpretation of the VTM: it provides a probability given only that a particle is within a bubble, not the specific location of the particle within the bubble. However, this implied diffusion is inconsistent with (3.2), which comes from the VTM providing less information than $\mathbf{s}^n$. A result of this loss of information is that effective VTMs depend on N and are thus not unique.

As an example, consider the case illustrated in figure B-2. Two bubbles of equal volumes $v_2^n = v_1^n$ briefly coalesce at $t = T$ to form a single bubble of volume $v_1^{n+1} = v_1^n + v_2^n$. At $t = 2T$ the single bubble fragments along the original boundary such that no dark fluid was mixed. Here we consider connectedness provided by CCL a given but note that such an event could be either physical or spurious, as the connectedness accuracy for closely passing bubbles depends on the CCL method and its parameters (Chan *et al.*, 2021*a*). First, consider the case where $\Delta t_s = 2T$. Eulerian volume tracking produces the correct VTM:

$$\mathbf{A}^{(0\to 2T)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{B.3}$$

Even if this were a spurious event, ELA with $\Delta t_s \geq 2T$ would correctly identify no exchange of volume and no cycles. Second, consider the case where $\Delta t_s = T$, with $N = 2$ to get

	$\tau = \Delta t_s/t_b$	$N = \Delta t_{s,\text{eff.}}/\Delta t_s$	$\tau_{\text{eff.}} = \Delta t_{s,\text{eff.}}/t_b$
Case Lb	0.1000	1	0.1000
Case S1	0.0500	2	0.1000
Case S2	0.0250	4	0.1000
Case S3	0.0125	8	0.1000
Case S4	0.0063	16	0.1000
Case S5	0.0031	32	0.1000
Case S6	0.0016	64	0.1000

Table B-1: Summary of how we apply matrix multiplication to the simulations from table 3-1 to obtain the same effective snapshot intervals.

$\Delta t_{s,\text{eff.}} = 2T$:

$$\tilde{\mathbf{A}}^{(0\to 2T)} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} . \tag{B.4}$$

While $\tilde{\mathbf{A}}^{(0\to 2T)}$ still satisfies (3.11), the loss of spatial information when the vector color function is collapsed into a VTM at $t = T$ creates a more diffuse tracking matrix. This loss of information means that there is no reliable way to decrease the diffusive error apart from decreasing N or introducing more information. Keeping N constant, one would have to make assumptions about the underlying flow to determine a likely evolutionary path from the possible events present in the diffuse effective VTM. This is equivalent to a (only slightly more constrained) Lagrangian tracking approach and would have the same challenges.

We note that, based on the interpretation of the VTM as a graph (see §3.3.2), it can be shown that for the special case where $\tilde{\mathbf{A}}$ (or a connected component of it) found using $N > 1$ has no cycles, it must be equal to the tracking matrix $\mathbf{A}'$ (or a corresponding connected component of it) found using a larger true snapshot interval and no multiplication over the same time period ($K' = NK$).

B.3 Performance of matrix multiplication

To study the effects of approximating long snapshot intervals through matrix multiplication, we use otherwise identical simulations with different Δt_s shown in table 3-1. We then multiply the resulting VTMs using (3.13) to achieve the same $\Delta t_{s,\text{eff.}}/t_b = 0.1$ (see table B-1). Defining n' to index the effective snapshot interval, i.e., $t^{n'+1} = t^{n'} + \Delta t_{s,\text{eff.}}$, each simulation and subsequent matrix multiplication generates VTMs describing

$$\mathbf{v}^{n'+1} = \tilde{\mathbf{A}}^{(n'\to n'+1)} \mathbf{v}^{n'} . \tag{B.5}$$

For case Lb, $(\Delta t_s)_{\text{Lb}} = (\Delta t_{s,\text{eff.}})_{\text{Lb}}$, so no matrix multiplication is necessary and we use $(\mathbf{A})_{\text{Lb}}$ as a reference. The difference introduced by matrix multiplication for the other snapshot interval cases (Case S1, Case S2, etc.) is

$$\mathbf{D}^{(n'\to n'+1)} \equiv \tilde{\mathbf{A}}^{(n'\to n'+1)} - \left[\mathbf{A}^{(n'\to n'+1)}\right]_{\text{Case Lb}} . \tag{B.6}$$

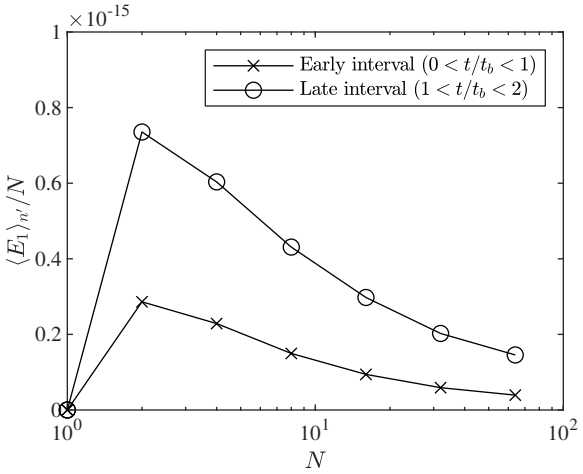


Figure B-3: The average normalized volume conservation error due to matrix multiplication, E_1, for HIT simulations with different true snapshot intervals Δt_s but the same effective snapshot interval $\Delta t_{s,\text{eff.}} = N\Delta t_s$ (see table B-1).

For analysis, we remove columns of $\mathbf{D}^{(n'\to n'+1)}$ and entries of $\mathbf{v}^{n'}$ relating to under-resolved parent bubbles, $v_i^{n'} < v_{\text{res}}$.

We first confirm that matrix multiplication is volume conservative. The normalized volume conservation error for each case is

$$\frac{\left\|\tilde{\mathbf{A}}^{(n'\to n'+1)}\mathbf{v}^{n'} - \mathbf{v}^{n'+1}\right\|_1}{\left\|\mathbf{v}^{n'+1}\right\|_1}$$

By subtracting the error from Case Lb, we obtain the volume conservation error due to matrix multiplication only:

$$(E_1)^{n'} = \frac{\left\|\mathbf{D}^{(n'\to n'+1)}\mathbf{v}^{n'}\right\|_1}{\left\|\mathbf{v}^{n'+1}\right\|_1}\,. \tag{B.7}$$

Figure B-3 shows the average E_1 error for the early and late time period normalized by the number of matrix multiplication operations N. The $O(10^{-15})$ normalized error for all simulations shows that matrix multiplication is volume conservative to machine precision.

While matrix multiplication is volume conservative, as discussed in §B.2, there is a loss of accuracy. To quantify this, we examine how individual columns of the VTM differ due to matrix multiplication. Based on $\mathbf{D}^{n'} = \{(d_{ij})^{n'}\}$, we define the average difference per column for each matrix

$$(E_2)^{n'} = \frac{1}{2}\left\langle \sum_i \left|(d_{ij})^{n'}\right| \right\rangle_j\,. \tag{B.8}$$

Noting that (by volume conservation) $\sum_i d_{ij} = 0$ and that for any VTM $\sum_i a_{ij} = 1$, we see that the factor of $1/2$ guarantees $(E_2)^{n'} \in [0, 1]$. Figure B-4 shows that the trend of the growth of the average E_2 error for both the early and late time periods behaves similarly

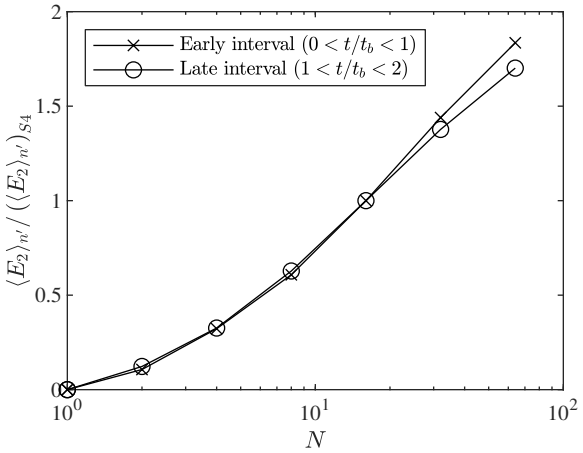


Figure B-4: The growth of the average E_2 error with N for HIT simulations with different true snapshot intervals Δt_s but the same effective snapshot interval $\Delta t_{s,\text{eff.}} = N\Delta t_s$ (see table B-1). Values are normalized by the value for $N = 16$ (case S4): $(\langle E_2 \rangle_{n'})_{\text{S4}} = 0.029$ for the early and $(\langle E_2 \rangle_{n'})_{\text{S4}} = 0.079$ for the late interval.

with changing N, apart from a scaling coefficient. For $N \gg 1$, E_2 exhibits approximately logarithmic growth. Although the similar behavior between early and later time periods suggests this logarithmic growth is independent to the complexity of the flow; we caution that this could also potentially depend on the type of CCL method used and the $\Delta t_{s,\text{eff.}}$ chosen, which we do investigate here. If logarithmic growth is generally true, it would mean that large Ns, and therefore small Ks, can be used with only moderate loss of accuracy.

In Gaylo *et al.* (2022) there is additional analysis on how smaller K reduce $\max\{\text{nnz}(\mathbf{s}^n)\}$, the largest number of non-zero entry in any single cell anytime during the entire simulation. In the implementation of ELA used for that work, `ELA`[1], memory was pre-allocated equally among grid cells to store $\mathbf{s}^n$, so the memory requirement scaled with $\max\{\text{nnz}(\mathbf{s}^n)\}$. We now have a better implementation of ELA, `flexELA`[2], which uses dynamic memory allocation to store $\mathbf{s}^n$, so the memory requirement scales with the average $\langle \text{nnz}(\mathbf{s}^n) \rangle$ rather than the maximum. In our experience, this newer implementation eliminates memory as a significant barrier to using ELA.

[1] `https://github.com/dgaylo/ELA`
[2] `https://github.com/dgaylo/flexELA`

Appendix C

Proof of Eulerian Label Advection Volume Conservation

Using an approach based on the proof of cVOF volume conservation by Weymouth & Yue (2010), this appendix proves that the Courant restriction

$$\Delta t \sum_{d=1}^{\mathcal{N}} \left| \frac{u_d}{\Delta x_d} \right| < C \tag{2.13}$$

where $C = 1/2$ guarantees that there is no over- or under-filling of the vector source fraction field, i.e.,

$$0 \leq (s_l)^{(d)} \leq 1\,, \tag{3.33}$$

for any component l at any step $d = 1 \ldots \mathcal{N}$ of the operator-split advection. We have shown that the flux terms are conservative and the dilation terms sum to zero (see §3.4.2), so this proof of no over or under filling proves that ELA is volume conservative to machine precision.

C.1 Mathematical proof

cVOF with $C \leq 1/2$ guarantees $f^{(d)} \leq 1$ (Weymouth & Yue, 2010). By construction, ELA satisfies the consistency requirement, so (3.28) is always true and therefore $C \leq 1/2$ guarantees

$$\sum_{l} (s_l)^{(d)} \leq 1 \tag{C.1}$$

for any step d. Thus, at any step d, it is impossible for a component $(s_l)^{(d)} > 1$ without at least one other component $(s_l)^{(d)} < 0$. This means that proving no under-filling,

$$0 \leq (s_l)^{(d)}\,, \tag{C.2}$$

is sufficient to prove (3.33). It would also be sufficient to prove no over-filling, but we find proving no under-filling is the easier path.

To prove (C.2), for each operator-split step we consider the six possible combinations of the magnitude and sign of the velocity on the positive (u_r) and negative (u_l) side of the cell:

(a) $u_r \geq u_l > 0$ (c) $u_r > 0 > u_l$ (e) $0 > u_l \geq u_r$
(b) $u_l > u_r > 0$ (d) $u_l > 0 > u_r$ (f) $0 > u_r > u_l$

Case (e) is symmetric with case (a) and case (f) is symmetric with case (b). This leaves proving that cases a–d cannot under fill. For convenience we scale the velocities and fluxes to local Courant numbers,

$$u' = u\frac{\Delta t}{\Delta x_d}\,, \tag{C.3a}$$

$$\mathbf{F}' = \mathbf{F}\frac{\Delta t}{\Delta \Omega}\,. \tag{C.3b}$$

Thus, (3.25) becomes

$$\mathbf{s}^{(d)} - \mathbf{s}^{(d-1)} = \mathbf{F}'_{d+1/2} - \mathbf{F}'_{d-1/2} + \tilde{\mathbf{c}}\,\Delta u'_d \quad \text{for} \quad d \in 1 \ldots \mathcal{N}\,, \tag{C.4}$$

where $\Delta u'_d = u'_r - u'_l$. Substituting in (3.30) and (3.31) and introducing absolute values to illustrate the sign of each term,

$$\text{case (a)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} = -\hat{\mathbf{s}}_d^{(d-1)}\left|F'_{d+1/2}\right| + \hat{\mathbf{s}}_{d-1}^{(d-1)}\left|F'_{d-1/2}\right| + \hat{\mathbf{s}}_d^{(0)}\tilde{c}\left|\Delta u'_d\right|\,, \tag{C.5a}$$

$$\text{case (b)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} = -\hat{\mathbf{s}}_d^{(d-1)}\left|F'_{d+1/2}\right| + \hat{\mathbf{s}}_{d-1}^{(d-1)}\left|F'_{d-1/2}\right| - \hat{\mathbf{s}}_d^{(0)}\tilde{c}\left|\Delta u'_d\right|\,, \tag{C.5b}$$

$$\text{case (c)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} = -\hat{\mathbf{s}}_d^{(d-1)}\left|F'_{d+1/2}\right| - \hat{\mathbf{s}}_d^{(d-1)}\left|F'_{d-1/2}\right| + \hat{\mathbf{s}}_d^{(0)}\tilde{c}\left|\Delta u'_d\right|\,, \tag{C.5c}$$

$$\text{case (d)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} = +\hat{\mathbf{s}}_{d+1}^{(d-1)}\left|F'_{d+1/2}\right| + \hat{\mathbf{s}}_{d-1}^{(d-1)}\left|F'_{d-1/2}\right| - \hat{\mathbf{s}}_d^{(0)}\tilde{c}\left|\Delta u'_d\right|\,. \tag{C.5d}$$

Dropping the positive terms gives the inequalities,

$$\text{case (a)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} \geq \hat{\mathbf{s}}_d^{(d-1)}F'_{d+1/2}\,, \tag{C.6a}$$

$$\text{case (b)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} \geq \hat{\mathbf{s}}_d^{(d-1)}F'_{d+1/2} + \hat{\mathbf{s}}_d^{(0)}\tilde{c}\,\left(u'_r - u'_l\right)\,, \tag{C.6b}$$

$$\text{case (c)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} \geq \hat{\mathbf{s}}_d^{(d-1)}\Delta F'_d\,, \tag{C.6c}$$

$$\text{case (d)}: \quad \mathbf{s}_d^{(d)} - \mathbf{s}_d^{(d-1)} \geq \hat{\mathbf{s}}_d^{(0)}\tilde{c}\,\left(u'_r - u'_l\right)\,, \tag{C.6d}$$

where $\Delta F'_d = F'_{d+1/2} - F'_{d-1/2}$. Because the fluxes in ELA are calculated using up winding, the negative terms we are interested in depend on $\hat{\mathbf{s}}_d$ and not the value in neighboring cells ($\hat{\mathbf{s}}_{d-1}$ or $\hat{\mathbf{s}}_{d+1}$). Thus, as all values of $\mathbf{s}$ and $\hat{\mathbf{s}}$ are for the cell of interest, not its neighbors, we will drop the subscript index.

Cases (a) and (c)

We start by considering case (a) and case (c). For these cases, cVOF ensures (Weymouth & Yue, 2010)

$$F'_{d+1/2}, \Delta F'_d \geq -f^{(d-1)} \,. \tag{C.7}$$

Thus, both (C.6a) and (C.6c) can be written as a looser bound

$$\text{case (a), (c)}: \qquad \mathbf{s}^{(d)} - \mathbf{s}^{(d-1)} \geq -\hat{\mathbf{s}}^{(d-1)} f^{(d-1)} \,. \tag{C.8}$$

Using (3.28), we can rewrite (3.29), the definition of $\hat{\mathbf{s}}$, to give

$$\mathbf{s}^{(d-1)} = \hat{\mathbf{s}}^{(d-1)} f^{(d-1)} \,. \tag{C.9}$$

Substituting into (C.8), we have

$$\text{case (a), (c)}: \qquad \mathbf{s}^{(d)} \geq \mathbf{0} \,. \tag{C.10}$$

This shows that (C.7), which is guaranteed by cVOF for $C \leq 1/2$, also guarantees case (a) and case (c) cannot under fill.

Cases (b) and (d) without dilation

For case (b) and (d), we start with $\tilde{c} = 0$ (corresponding to $f^{(0)} \leq 1/2$). For case (b), (C.6b) with $\tilde{c} = 0$ reduces to (C.6a). As (C.7) also applies to case (b) (Weymouth & Yue, 2010), the previous proof for case (a) applies to case (b) with $\tilde{c} = 0$ and we obtain

$$\text{case (b) with } \tilde{c} = 0: \qquad \mathbf{s}^{(d)} \geq \mathbf{0} \,. \tag{C.11}$$

For case (d), (C.6d) with $\tilde{c} = 0$ reduces to

$$\text{case (d) with } \tilde{c} = 0: \qquad \mathbf{s}^{(d)} \geq \mathbf{s}^{(d-1)} \,. \tag{C.12}$$

As case (d) with $\tilde{c} = 0$ can only increase each component of $\mathbf{s}$ between steps, it cannot lead to under filling.

Cases (d) with dilation

This leaves us with case (b) and (d) with $\tilde{c} = 1$. These cases are challenging as we must ensure the dilation term, which is based on $\mathbf{s}^{(0)}$, does not subtract more than is present in $\mathbf{s}^{(d-1)}$. For case (d) we note that $\tilde{c} = 1$ means that $f^{(0)} > 1/2$. Thus, (C.9) provides the inequality

$$\hat{\mathbf{s}}^{(0)} \leq 2\mathbf{s}^{(0)} \,. \tag{C.13}$$

This allows us to write (C.6d) as a looser bound

$$\text{case (d) with } \tilde{c} = 1: \qquad \mathbf{s}^{(d)} - \mathbf{s}^{(d-1)} \geq 2\mathbf{s}^{(0)} \left(u'_r - u'_l\right) \,. \tag{C.14}$$

We need to ensure successive operations cannot under fill, so we sum this expression over $d \leq \mathcal{N}$ operations,

$$\text{case (d) with } \tilde{c} = 1: \qquad \mathbf{s}^{(d)} \geq \mathbf{s}^{(0)} \left[1 + 2 \sum_{d'=1}^{d} \min(\Delta u'_{d'}, 0) \right], \tag{C.15}$$

Thus, to ensure $\mathbf{s}^{(d)} \geq \mathbf{0}$ we require

$$\sum_{d=1}^{\mathcal{N}} \min(\Delta u'_d, 0) \geq -\frac{1}{2}. \tag{C.16}$$

Recalling that $\sum_{d=1}^{\mathcal{N}} \Delta u'_d = 0$ for a divergence-free flow, (C.16) is true for the Courant condition

$$C = \sum_{d=1}^{\mathcal{N}} \left| u'_d \right| \leq \frac{1}{2}, \tag{C.17}$$

recovering (2.13).

Cases (b) with dilation

We now consider case (b) with $\tilde{c} = 1$. For case (b) (in addition to (C.7)) cVOF with $C \leq 1/2$ ensures (Weymouth & Yue, 2010)

$$F'_{d+1/2}, \Delta F'_d \geq -u'_r. \tag{C.18}$$

This allows us to rewrite (C.6b) as a looser bound

$$\text{case (b) with } \tilde{c} = 1: \qquad \mathbf{s}^{(d)} - \mathbf{s}^{(d-1)} \geq -\hat{\mathbf{s}}^{(d-1)} u'_r + \hat{\mathbf{s}}^{(0)} \left(u'_r - u'_l \right), \tag{C.19}$$

For each component l of the vector equation, it is now necessary to further split case (b) into case (b.i), $(\hat{s}_l)^{(0)} < (\hat{s}_l)^{(d-1)}$ and case (b.ii), $(\hat{s}_l)^{(0)} \geq (\hat{s}_l)^{(d-1)}$. Recalling $u_l > u_r > 0$ for both, this allows us to write

$$\text{case (b.i) with } \tilde{c} = 1: \qquad (s_l)^{(d)} - (s_l)^{(d-1)} \geq -(\hat{s}_l)^{(d-1)} u'_l \tag{C.20a}$$

$$\text{case (b.ii) with } \tilde{c} = 1: \qquad (s_l)^{(d)} - (s_l)^{(d-1)} \geq -(\hat{s}_l)^{(0)} u'_l. \tag{C.20b}$$

Starting with case (b.ii), we use (C.13) to write

$$\text{case (b.ii) with } \tilde{c} = 1: \qquad (s_l)^{(d)} - (s_l)^{(d-1)} \geq -2(s_l)^{(0)} u'_l. \tag{C.21}$$

As with case (d), we sum over $d \leq \mathcal{N}$ and obtain

$$\text{case (b.ii) with } \tilde{c} = 1: \qquad (s_l)^{(d)} \geq (s_l)^{(0)} \left[1 - 2 \sum_{d'=1}^{d} \left| u'_{d'} \right| \right], \tag{C.22}$$

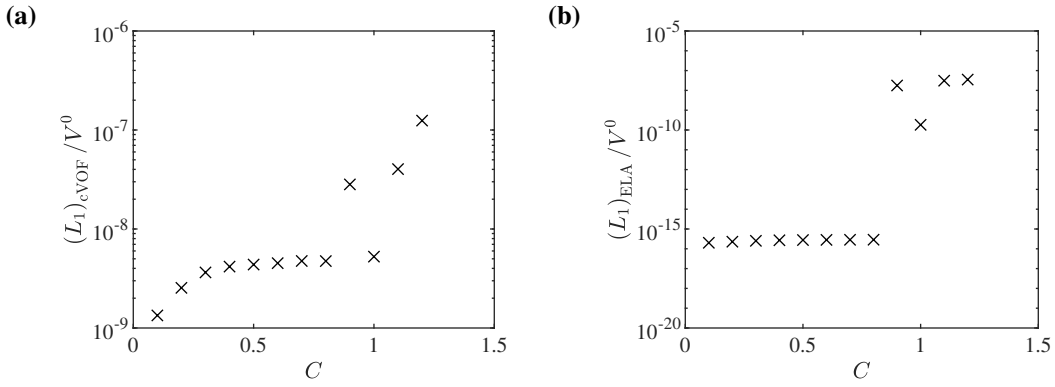


Figure C-1: Normalized L_1 volume-conservation error for (a) cVOF and (b) ELA as a function of the cVOF Courant restriction C.

where, to ensure the symmetric case (f) is also captured, $\left|u'_d\right| = \max\left(\left|u'_l\right|, \left|u'_r\right|\right)$. Thus, for $(s_l)^{(d)} \geq 0$, we require

$$1 - 2\sum_{d'=1}^{d} \left|u'_{d'}\right| \geq 0\,, \tag{C.23}$$

which is always satisfied given the Courant condition $C \leq 1/2$.

This leaves case (b.i). Using (C.9), (C.20a) can be rewritten

$$\text{case (b.i) with } \tilde{c} = 1: \qquad (s_l)^{(d)} \geq (s_l)^{(d-1)}\left[1 - \frac{u'_l}{f^{(d-1)}}\right]. \tag{C.24}$$

Thus we require $f^{(d-1)} \geq |u_d|$ for $(s_l)^{(d)} \geq 0$. The absolute value sign is introduced to ensure symmetry with case (f). Rewriting (2.40) for $\tilde{c} = 1$ in terms of the scaled velocities and fluxes gives

$$f^{(d)} = f^{(d-1)} + \Delta F'_d + \left(u'_r - u'_l\right)\,. \tag{C.25}$$

Using (C.18), $f^{(d)} \geq f^{(d-1)} - u'_l$. Summing (C.25) over $d - 1 \leq \mathcal{N} - 1$ and recalling we are interested in $f^{(0)} > 1/2$, it can be shown cVOF guarantees the bound

$$f^{(d-1)} \geq 1/2 - C + \left|u'_d\right|\,. \tag{C.26}$$

Thus $f^{(d-1)} \geq \left|u'_d\right|$ and therefore $\mathbf{s}^{(d)} \geq 0$ is true if $C \leq 1/2$, the same condition as (2.13).

C.2 Numerical validation

To validate the Courant restriction, we repeat the three-dimensional bubble fragmentation in HIT from §3.5 with a snapshot interval equal to simulation time step ($\Delta t_s = \Delta t$), without using (3.35), with a zero-threshold value $\epsilon = 10^{-8}$, and with a coarser grid of 128^3. Rather than $C = 1/2$, we repeat the simulation with a range of C. Although Δt is chosen dynamically based on (2.13) as well as numerical stability criteria (Campbell, 2014), we find that for these simulations (2.13) is the most restrictive condition.

The volume-error for this simulation over the time range $0 < t/t_b < 2$ as a function of C is shown in figure C-1. Note that an $O(\epsilon)$ error is expected for cVOF due to the filter (2.43). We see that cVOF is volume conservative for C as large as 0.8 and that ELA is also volume conservative over a similar range. Due to the way inequalities are simplified, it is not a surprise that the observed limit on C is larger than the theoretical one derived by Weymouth & Yue (2010) for cVOF and here for ELA. However, it is not clear how a larger theoretical limit could be proven for cVOF, let alone ELA. The larger observed limit may be due to the specific nature of the HIT flow we consider and may not be general, where the cVOF and ELA proofs only assume the flow is divergence free.

Appendix D

Correlations Between Density Fluctuations and Turbulence in Strong FST

In section 4.5.2 we consider the Reynolds stress term due to density in the momentum equation, $\overline{\rho' ww}$, and decompose it into three terms,

$$\overline{\rho' ww} = \overline{\rho'_0 ww} + \overline{\rho'_B ww} + \overline{\rho'_D ww}\,, \tag{4.25}$$

corresponding to the effects of the free surface, bubbles, and droplets respectively. In this appendix we consider the Pearson's correlation coefficient

$$P_{X,Y} \equiv \frac{\mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]}{\sqrt{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}\sqrt{\mathbb{E}[Y^2] - (\mathbb{E}[Y])^2}}\,, \tag{D.1}$$

which describe the covariance between X and Y normalized such that $P_{X,Y} \in [-1, 1]$. Here X will be the density fluctuations (ρ', ρ_0, etc.) and Y the strength of the vertical fluctuations (ww).

We start by deriving an expression for $P_{\rho',ww}$. First, we can define a fluctuation in the color function, $c' = c - \overline{c}$. Recalling that $\gamma = \overline{c}$, we can rewrite (4.23) as

$$\rho' = \Delta\rho\; c'\,, \tag{D.2}$$

Multiplying a variable by a constant does not change the Pearson's correlation coefficient, so $P_{\rho',ww} = P_{c',ww}$. Canceling terms in (D.1) where $\overline{c'} = 0$, we obtain

$$P_{\rho',ww} = \frac{\overline{c' ww}}{\sqrt{\overline{c'c'}}\,\sqrt{\overline{wwww} - \overline{ww}^2}}\,.$$

We can further simplify this expression because c is a Bernoulli random variable, meaning it

only takes values of 0 or 1. This leads to $\overline{c'c'} = \gamma(1-\gamma)$, and finally

$$P_{\rho',ww} = \frac{\overline{c'ww}}{\sqrt{\gamma(1-\gamma)}\sqrt{\overline{wwww} - \overline{ww}^2}}. \tag{D.3}$$

The equations for $P_{\rho'_0,ww}$ and $P_{\rho'_D,ww}$ are similar. Because c_B takes values of 0 or -1, we replace γ with $-\gamma_B$:

$$P_{\rho'_B,ww} = \frac{\overline{c'_B ww}}{\sqrt{-\gamma_B(1+\gamma_B)}\sqrt{\overline{wwww} - \overline{ww}^2}}. \tag{D.4}$$

For reference, Figure D-1 shows $\sigma_{ww} = \sqrt{\overline{wwww} - \overline{ww}^2}$ from the forced FST simulations analyzed in Chapter 4.

Figure D-2 shows the correlation coefficients from forced FST simulations analyzed in Chapter 4, specifically those in the strong FST regime ($Fr_T^2 > 0.1$), and we find that all correlation coefficients collapse well when plotted against z^*. As discussed in section 4.5.2, $\overline{\rho'_0 ww}$ is the dominant term contributing to $\overline{\rho' ww}$, so it is not a surprise that $P_{\rho',ww} \approx P_{\rho'_0,ww}$. Beneath $z^* \approx 0.2$, we see that $P_{\rho'_0,ww} < 0$. This means that the presence of air ($\rho'_0 < 0$) is correlated with $ww > \overline{ww}$ and the presence of water ($\rho'_0 > 0$) is correlated with $ww < \overline{ww}$. In other words, the magnitude of vertical velocity fluctuations is larger in air than water for $z^* < 0.2$. For $z^* > 0.2$ this reverses ($P_{\rho'_0,ww} > 0$), meaning that water is associated with larger vertical fluctuations. One interpretation is that the presence of water at these large z^* is associated with high energy splashing.

Notably, we find that $P_{\rho'_B,ww} \approx 0.005$ meaning there is not a strong correlation between the presence of a bubble and magnitude of vertical velocity. This is what one would expect from a passive scalar, suggesting in these simulations bubble advection by turbulence is dominant over buoyant rise. For what (small) correlation we do see, the sign ($P_{\rho'_B,ww} > 0$) means that the presence of a bubble ($\rho'_B < 0$) is correlated with a slightly decreased magnitude of the vertical fluctuations ($ww < \overline{ww}$).

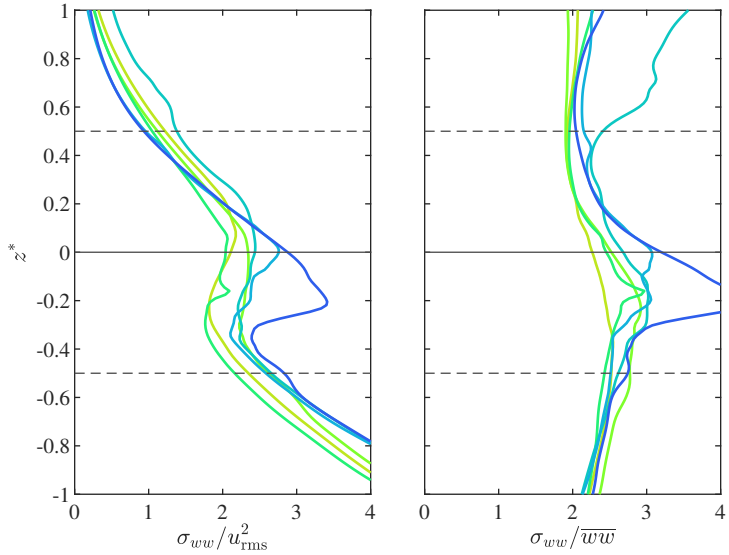


Figure D-1: Standard deviation of ww, σ_{ww}, as a function of depth, normalized by u^2_{rms} measured at $z^* = -0.5$ as well as $\overline{ww}$ measured at each depth for strong FST ($Fr^2_T > 0.1$, see figure 4-3a for color legend).

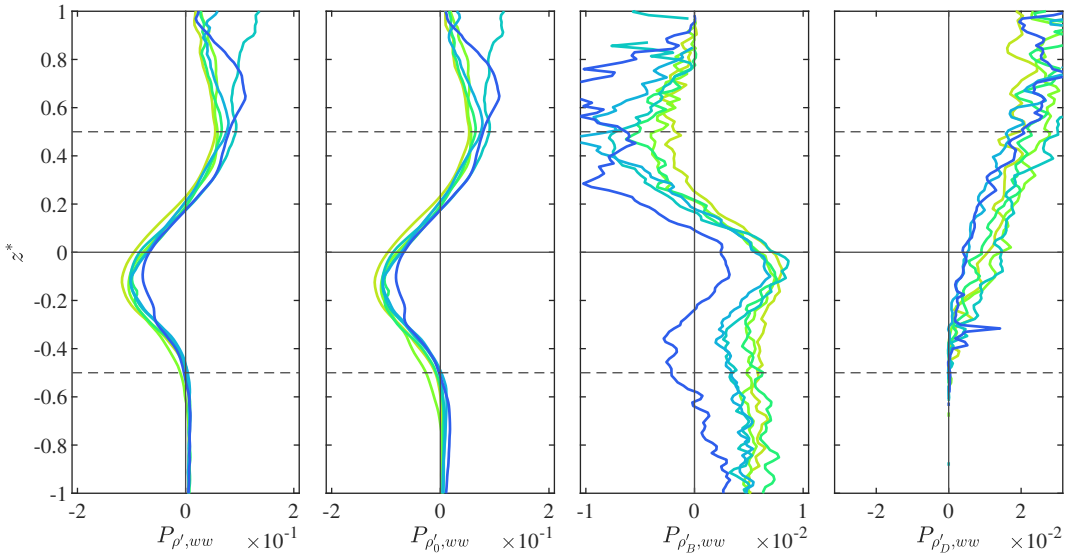
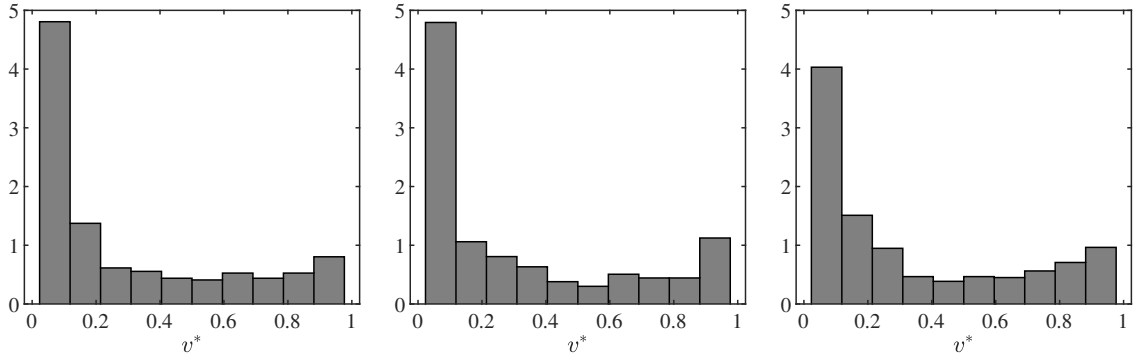


Figure D-2: Pearson's correlation coefficients corresponding to Reynolds stress terms in (4.25) for strong FST ($Fr^2_T > 0.1$, see figure 4-3a for color legend). Note the difference in horizontal scales.

Appendix E

Daughter Distributions from Bubble Fragmentation in Homogeneous Isotropic Turbulence



(a) We_B = 101–142, $\bar{m}$ = 3.37 **(b)** We_B = 50–71, $\bar{m}$ = 3.19 **(c)** We_B = 25–36, $\bar{m}$ = 3.00

(d) We_B = 13–18, $\bar{m}$ = 2.65 **(e)** We_B = 6.3–8.9, $\bar{m}$ = 2.55

Figure E-1: Measured daughter size distributions $f_V^*(v^*)$ and average number of daughter bubbles $\bar{m}$ for different ranges of bubble Weber numbers We_B, all measured using $T/t_\ell = 0.4$.

Appendix F

Instability of Horizontal Shear Flow with a Free Surface and Finite Depth

In this appendix, we extend "Instabilities of a horizontal shear flow with a free surface" by Longuet-Higgins (1998) to include the effects of finite depth. Our interest is the stability of the shear flow driven by (6.25), which can be modeled as perturbations from the mean flow

$$\bar{u}(z)/U = 1 - 0.9988 \ \mathrm{sech}(0.88137\, z/L)\,, \tag{F.1}$$

where the free surface is at $z = 0$. Longuet-Higgins (1998) argue that, for the purposes of stability analysis, one can consider a simpler mean flow,

$$\bar{u}(z) = \begin{cases} U & -H_1 < z \\ \Omega\,(H_2 + z) & -H_2 < z < -H_1 \\ 0 & z < -H_2 \end{cases}. \tag{F.2}$$

where $L = (H_1 + H_2)/2$, $\Omega \equiv U/2L$, and $h_1 = H_1/L$. Longuet-Higgins (1998) shows $h_1 = 0.1977$ makes the stability characteristics of (F.2) similar to that of (F.1). We extend the analysis by Longuet-Higgins (1998) to also include a bottom (where the vertical velocity must be zero) at finite depth $-H_3 < -H_2$. This introduces a third parameter $h_3 = H_3/L$, where $h_3 = 4$ in the shear flow simulations described in §6.4.

F.1 Linear dispersion relationship

For perturbations u and w in the lowest layer ($z < -H_2$), Longuet-Higgins (1998) consider

$$\begin{aligned} u(z < -H_2) &= -c \quad + \mathrm{i}k(E\mathrm{e}^{kz})e^{ikx}\,. \\ w(z < -H_2) &= \qquad\quad k(E\mathrm{e}^{kz})e^{\mathrm{i}kx} \end{aligned} \tag{F.3}$$

appropriate for infinite depth. For finite depth, we consider

$$\begin{aligned} u(z<-H_2) &= -c \quad +\mathrm{i}k(E\mathrm{e}^{kz} - F\mathrm{e}^{-kz})e^{\mathrm{i}kx}\,. \\ w(z<-H_2) &= \qquad\quad k(E\mathrm{e}^{kz} + F\mathrm{e}^{-kz})e^{\mathrm{i}kx} \end{aligned} \tag{F.4}$$

The bottom boundary condition ($w = 0$ at $z = -H_3$) gives $E\mathrm{e}^{-kH_3} + F\mathrm{e}^{kH_3} = 0$. We define $\lambda_3 \equiv \mathrm{e}^{-2kH_3}$ to express this boundary condition as $F = -E\lambda_3$ Substituting this into (F.4),

$$\begin{aligned} u(z<-H_2) &= -c \quad +\mathrm{i}kE(\mathrm{e}^{kz} + \lambda_3\mathrm{e}^{-kz})e^{\mathrm{i}kx}\,, \\ w(z<-H_2) &= \qquad\quad kE(\mathrm{e}^{kz} - \lambda_3\mathrm{e}^{-kz})e^{\mathrm{i}kx}\,. \end{aligned} \tag{F.5}$$

Now, following the same steps as Longuet-Higgins (1998), we define the velocity perturbation in the intermediate layer

$$\begin{aligned} u(-H_2<z<-H_1) &= \Omega(H_2+z) - c \quad +\mathrm{i}k(C\mathrm{e}^{kz} - D\mathrm{e}^{-kz})\mathrm{e}^{\mathrm{i}kx}\,, \\ w(-H_2<z<-H_1) &= \qquad\qquad\qquad\quad k(C\mathrm{e}^{kz} + D\mathrm{e}^{-kz})\mathrm{e}^{\mathrm{i}kx}\,. \end{aligned} \tag{F.6}$$

and in the upper layer,

$$\begin{aligned} u(z>-H_1) &= U - c \quad +\mathrm{i}k(A\mathrm{e}^{kz} - B\mathrm{e}^{-kz})\mathrm{e}^{\mathrm{i}kx}\,, \\ w(z>-H_1) &= \qquad\quad k(A\mathrm{e}^{kz} + B\mathrm{e}^{-kz})\mathrm{e}^{\mathrm{i}kx}\,. \end{aligned} \tag{F.7}$$

Applying linearized free-surface boundary conditions (e.g., $p = \rho g\eta$) at $z = 0$, gives

$$(U-c)^2 k(A-B) = \mathrm{g}(A+B)\,. \tag{F.8}$$

For $z = -H_1$ we specify continuity of velocity w and continuity of (linearized) force $\partial p/\partial x \approx \rho(\bar{u}-c)\partial w/\partial z - \rho w\,\mathrm{d}\bar{u}/\mathrm{d}z$, which gives the systems of equations

$$\begin{bmatrix} (A\mathrm{e}^{-kH_1} + B\mathrm{e}^{kH_1}) & = & (C\mathrm{e}^{-kH_1} + D\mathrm{e}^{kH_1}) \\ (U-c)k(A\mathrm{e}^{-kH_1} - B\mathrm{e}^{kH_1}) & = & (U-c)k(C\mathrm{e}^{-kH_1} - D\mathrm{e}^{kH_1}) - \Omega(C\mathrm{e}^{-kH_1} + D\mathrm{e}^{kH_1}) \end{bmatrix}. \tag{F.9}$$

Doing the same for $z = -H_2$,

$$\begin{bmatrix} (C\mathrm{e}^{-kH_2} + D\mathrm{e}^{kH_2}) & = & E(\mathrm{e}^{-kH_2} - \lambda_3\mathrm{e}^{kH_2}) \\ (0-c)k(C\mathrm{e}^{-kH_2} - D\mathrm{e}^{kH_2}) - \Omega(C\mathrm{e}^{-kH_1} + D\mathrm{e}^{kH_1}) & = & (0-c)kE(\mathrm{e}^{-kH_2} + \lambda_3\mathrm{e}^{kH_2}) \end{bmatrix}. \tag{F.10}$$

We note that finite depth makes the last equation this last equation different than Longuet-Higgins (1998).

Combining (F.8), (F.9), and (F.8) into a single system of equations, we have

$$\begin{bmatrix} Z^2(A-B) & = & A+B \\ \lambda_1 A + B & = & \lambda_1 C + D \\ Z(\lambda_1 A - B) & = & Z(\lambda_1 C - D) + \beta(\lambda_1 C + D) \\ \lambda_2 C + D & = & (\lambda_2 - \lambda_3)E \\ (Z+q)(\lambda_2 C - D) + \beta(C\lambda_2 + D) & = & (Z+q)E(\lambda_2 + \lambda_3) \end{bmatrix}, \tag{F.11}$$

where for clarity we have introduced the following notation:

$$\lambda_1 = \mathrm{e}^{-2kH_1}\,; \quad \lambda_2 = \mathrm{e}^{-2kH_2}\,; \quad c_0 = (\mathrm{g}/k)^{1/2}\,; \quad Z = \frac{c-U}{c_0}, \quad \beta = \frac{\Omega}{kc_0}\,; \quad q = U/c_0\,.$$

This is the same as Longuet-Higgins (1998), apart from the second to last and last equation which now include the effect of $\lambda_3 \neq 0$. To eliminate E from the system of equations, we rearrange the second to last equation in (F.11) to obtain

$$E(\lambda_2 + \lambda_3) = (\lambda_2 C + D)\,\Gamma\,, \tag{F.12}$$

where $\Gamma = (\lambda_2 + \lambda_3)\,/(\lambda_2 - \lambda_3)$. We substitute this into the last equation in (F.11) to obtain

$$\begin{bmatrix} Z^2(A-B) & = & A+B \\ \lambda_1 A + B & = & \lambda_1 C + D \\ Z(\lambda_1 A - B) & = & Z(\lambda_1 C - D) + \beta(\lambda_1 C + D) \\ (Z+q)(\lambda_2 C - D) + \beta(C\lambda_2 + D) & = & (\lambda_2 C + D)\Gamma \end{bmatrix}. \tag{F.13}$$

After some algebra, this system of equations can be represented by the matrix equation

$$\begin{bmatrix} (Z^2-1) & -(Z^2+1) & 0 & 0 \\ \lambda_1 & 1 & -\lambda_1 & -1 \\ \lambda_1 Z & -Z & -\lambda_1(Z+\beta) & Z-\beta \\ 0 & 0 & \lambda_2\,[(Z+q)(\Gamma-1)-\beta] & [(Z+q)(1+\Gamma)-\beta] \end{bmatrix} \begin{Bmatrix} A \\ B \\ C \\ D \end{Bmatrix} = \mathbf{0}\,. \tag{F.14}$$

Thus, setting the determinant of the matrix to zero gives Z. After some manipulation,

$$\begin{vmatrix} (Z^2-1) & 2 & 0 & 0 \\ \lambda_1 & -(1+\lambda_1) & -\lambda_1 & -1 \\ 0 & 2Z & -\lambda_1\beta & 2Z-\beta \\ 0 & 0 & \lambda_2\,[(Z+q)(\Gamma-1)-\beta] & (Z+q)(1+\Gamma)-\beta \end{vmatrix} = 0\,. \tag{F.15}$$

The result is a fourth order equation of the form

$$p_1 Z^4 + p_2 Z^3 + p_3 Z^2 + p_4 Z + p_5 = 0\,, \tag{F.16}$$

which can be solved for Z for a given k. From a Z, the frequency is

$$\omega = kc = kc_0(Z+q)\,. \tag{F.17}$$

For stability, our interest is how the imaginary part of (nondimensionalized) angular frequency $\sigma = \omega(L/U)$ depends on the (nondimensionalized) wave number $\kappa = kL$, for a given $Fr^2 = U^2/L\mathrm{g}$ and depth $h_3 = H_3/L$. For $Fr^2 = 5$, figure F-1 compares the infinite depth solution to the finite depth solution for $h_3 = 4$. This deep ($h_3 \gg 1$) but finite depth removes the $\kappa \ll 1$ instability present for infinite depth, but causes very little change to the $\kappa \sim 1$ instability. For this branch, we find the unstable wave numbers are $\kappa \in [0.6653, 1.2302]$.

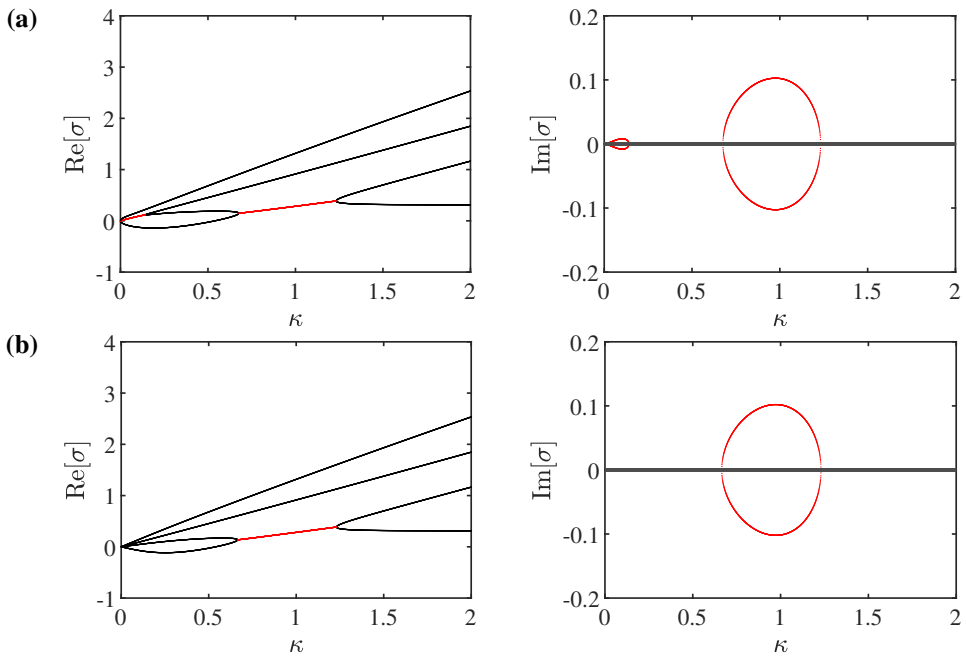


Figure F-1: Frequency $\mathrm{Re}[\sigma]$ and growth rate $\mathrm{Im}[\sigma]$ as a function of wave number κ for $h_1 = 0.1977$, $Fr^2 = 5$, and (a) $h_3 = \infty$; (b) $h_3 = 4$.

F.2 Results from large Froude number simulations

We now look for evidence of this shear flow instability in the DNS of large Fr. We perform a simulation at $Fr^2 = 5$ like that described in §6.4, but with the horizontal domain length increased by a factor of $8/3$ to capture possible longer waves. Because the free surface can be multi-valued, there is not a strictly well-defined definition of the wave height $\eta(x, y)$ based on the color function c. For an approximation $\hat{\eta}(x, y)$, we obtain the correct total gravitational potential energy if

$$\tfrac{1}{2}\hat{\eta}(x, y)^2 = \int_{z_0}^{z_1} cz \, \mathrm{d}z + z_0^2 \,, \tag{F.18}$$

where z_0 is the bottom of the domain, z_1 the top of the domain, and $c(x, y, z) = 1$ in water. This quadratic equation has two solutions of opposite signs, so we use the integral of c to guess the appropriate sign:

$$\hat{\eta}(x, y) \equiv \mathrm{sign}\left[\int_{z_0}^{z_1} cz \, \mathrm{d}z + z_0\right] \sqrt{2\int_{z_0}^{z_1} cz \, \mathrm{d}z + z_0^2} \,. \tag{F.19}$$

This always gives the correct $\hat{\eta}(x, y) = \eta(x, y)$ if the free surface is single valued.

Using $\hat{\eta}(x, y)$ sampled over $t \in [40, 70]$, we calculate the two-dimensional energy spectrum $E_{\eta\eta}(\boldsymbol{\kappa})$, normalized such that $\iint E_{\eta\eta}(\boldsymbol{\kappa}) \, \mathrm{d}\boldsymbol{\kappa} = 1$. For analysis, we split the wave

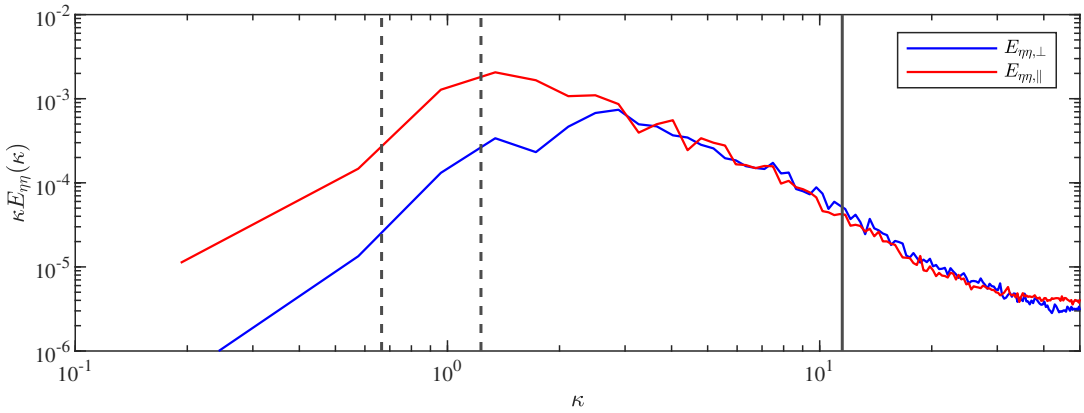


Figure F-2: Wave spectrum, split into shear-parallel and shear-perpendicular components, for a free-surface shear flow simulation at $Fr^2 = 5$, averaged over $t \in [40, 70]$. (- - - -) show the unstable wave numbers $\kappa \in [0.6653, 1.2302]$ from linear analysis and (——) shows $k_T = 2\pi/L_T$ associated with near-surface turbulence.

number space into energy parallel to the shear,

$$E_{\eta\eta,\parallel}(\kappa) \equiv 4 \int_0^{\pi/4} E_{\eta\eta}(\kappa\cos\theta, \kappa\sin\theta)\kappa \, \mathrm{d}\theta \,, \tag{F.20}$$

and perpendicular to the shear,

$$E_{\eta\eta,\perp}(\kappa) \equiv 4 \int_{\pi/4}^{\pi/2} E_{\eta\eta}(\kappa\cos\theta, \kappa\sin\theta)\kappa \, \mathrm{d}\theta \,, \tag{F.21}$$

shown in figure F-2. As expected, for long waves ($\kappa < 3$) there is significantly more energy in the parallel component than the perpendicular component. We see the majority of this energy is around wave numbers $\kappa = 1$–2, nearby the range predicted from linear analysis. For $\kappa > 3$, we see the wave spectrum becomes roughly isotropic. All these waves are much longer than the longest wave number associated with near-surface turbulence, $k_T = 2\pi/L_T$ ($L_T \approx 0.55L$ for this simulation).

Appendix G

Verification of Grid Independence for Free-surface Shear Flow

Yu *et al.* (2019) performed a grid convergence study for free-surface shear flow at the same Re and similar Fr as here, and confirmed turbulence and the bubble population are sufficiently resolved by the grid described in §6.4.1. In addition, we perform a convergence study of the entrainment size distribution $I(a)$ and degassing rate $\Lambda(a)$ measured by ELA, to confirm they are sufficiently resolved. We perform a set of 3 simulations at $Fr^2 = 15$ using the same method described in §6.4.1, but with a finer grid of $576^2 \times 384$. This gives $\Delta_{576} \approx 0.018$ and $\eta_T/\Delta_{576} \approx 2.0$, versus $\Delta_{384} \approx 0.027$ and $\eta_T/\Delta_{384} \approx 1.3$ from §6.4.1. For these simulations, we include all bubbles of radius larger than $a_{\text{res}} = 1.5\Delta_{576}$. For $t \in [40, 70]$, we obtain $\varepsilon \times 10^4 = 6.1$ and $u_{\text{rms}} = 0.074$, consistent with $Fr^2 = 15$ using Δ_{384} (see table 6-3).

Figure G-1 compares the measured entertainment size distribution $I(a)$, degassing rate $\Lambda(a)$, and bulk bubble size distribution $N(a)$ between the two grid resolutions. These correspond to figure 6-7a, figure 7-3, and figure 7-4a, and we note that the range of the horizontal axis has been extended in figure G-1. Because there are only 3 simulations at Δ_{576} (compared to 6 with Δ_{384}) there is more statistical variability, particularly for larger bubbles of which fewer are observed in each simulation. From figure G-1, it is clear, especially for the smallest bubbles where resolution would be a concern, that the results are consistent between the Δ_{576} and Δ_{384} grids. We conclude that DNS with the Δ_{384} grid described in §6.4.1 sufficiently resolves the relevant physics.

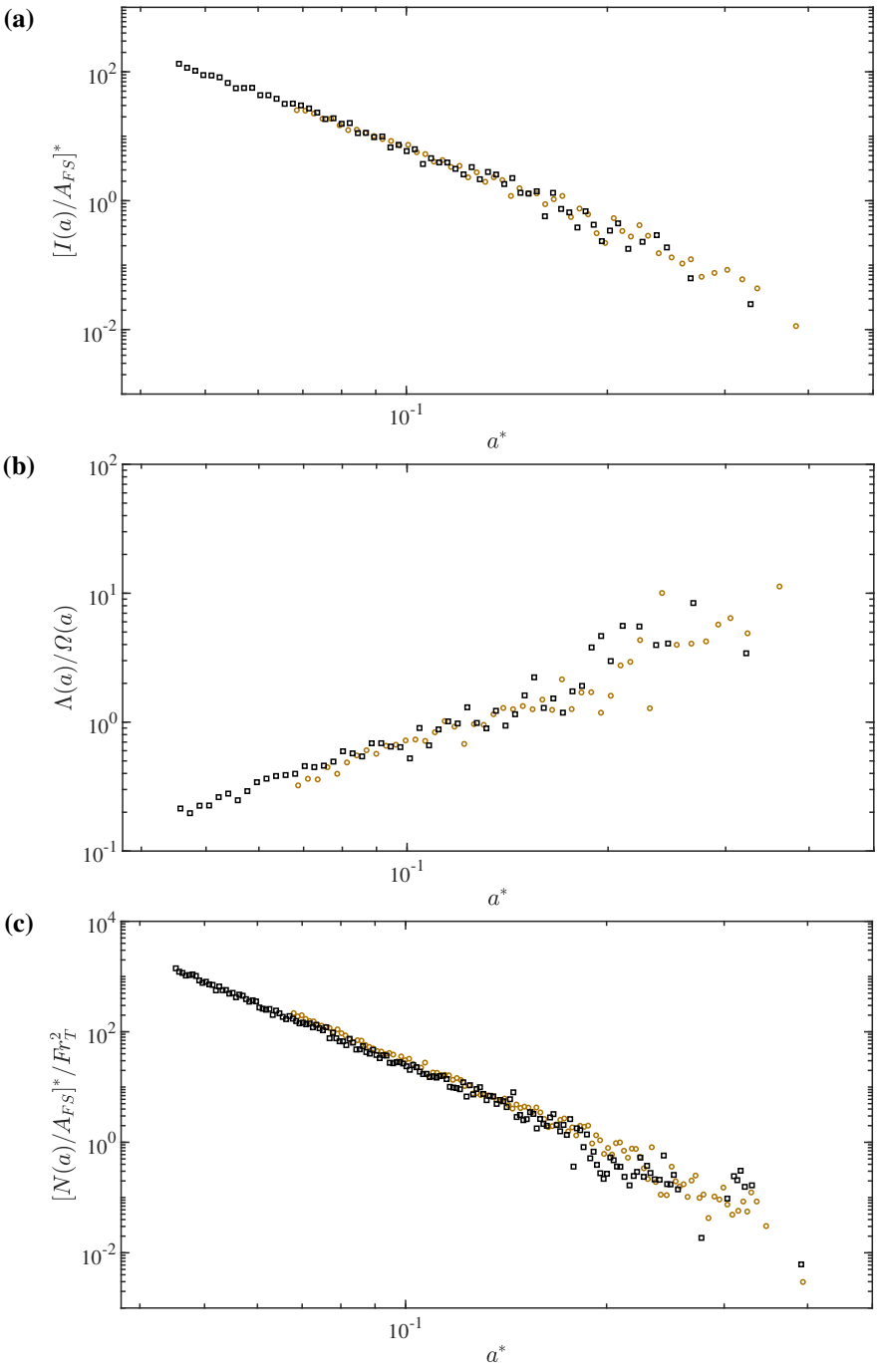


Figure G-1: Entrainment size distribution (a), degassing rate (b), and bubble size distribution (c) during $t \in [40, 70]$ for $Fr^2 = 15$ with ○, Δ_{384} and □, Δ_{576}. For consistency, turbulence values from Δ_{384} are used for non-dimensionalization and to calculate $\Omega(a)$ and $Fr_T^2 = 0.12$.

References

Baraldi, A., Dodd, M.S. & Ferrante, A. 2014 A mass-conserving volume-of-fluid method: Volume tracking and droplet surface-tension in incompressible isotropic turbulence. *Comput. Fluids* **96**, 322–337. doi: 10.1016/j.compfluid.2013.12.018.

Basak, Saswata, Bitencourt, Umberto C., Gao, Qiang, Deane, Grant B., Dale, M. & Shen, Lian 2026 A parallel algorithm for detection and tracking of non-binary fragmentation and coalescence of droplets and bubbles in numerical simulations of multiphase flows. *J. Comput. Phys.* **545**, 114441. doi: 10.1016/j.jcp.2025.114441.

Bertola, Numa, Wang, Hang & Chanson, Hubert 2018 A physical study of air–water flow in planar plunging water jet with large inflow distance. *Int. J. Multiphase Flow* **100**, 155–171. doi: 10.1016/j.ijmultiphaseflow.2017.12.015.

Brackbill, J.U, Kothe, D.B & Zemach, C. 1992 A continuum method for modeling surface tension. *J. Comput. Phys.* **100** (2), 335–354. doi: 10.1016/0021-9991(92)90240-Y.

Brocchini, M. 2002 Free surface boundary conditions at a bubbly/weakly splashing air–water interface. *Phys. Fluids* **14** (6), 1834–1840. doi: 10.1063/1.1475998.

Brocchini, M. & Peregrine, D. H. 1996 Integral flow properties of the swash zone and averaging. *J. Fluid Mech.* **317**, 241–273. doi: 10.1017/S0022112096000742.

Brocchini, M. & Peregrine, D. H. 2001*a* The dynamics of strong turbulence at free surfaces. part 1. description. *J. Fluid Mech.* **449**, 225–254. doi: 10.1017/S0022112001006012.

Brocchini, M. & Peregrine, D. H. 2001*b* The dynamics of strong turbulence at free surfaces. part 2. free-surface boundary conditions. *J. Fluid Mech.* **449**, 255–290. doi: 10.1017/S0022112001006024.

Calado, Andre & Balaras, Elias 2025 Interfacial deformation and energy exchange in strong free-surface turbulence, arXiv: 2506.10090.

Callaghan, Adrian H., Deane, Grant B. & Stokes, M. Dale 2013 Two regimes of laboratory whitecap foam decay: Bubble-plume controlled and surfactant stabilized. *J. Phys. Oceanogr.* **43**, 1114–1126. doi: 10.1175/JPO-D-12-0148.1.

Campbell, Bryce 2014 A Mechanistic Investigation of Nonlinear Interfacial Instabilities Leading to Slug Formation in Multiphase Flows. PhD thesis, MIT.

Castro, Alejandro M. & Carrica, Pablo M. 2013 Bubble size distribution prediction for large-scale ship flows: Model evaluation and numerical issues. *Int. J. Multiphase Flow* **57**, 131–150. doi: 10.1016/j.ijmultiphaseflow.2013.08.001.

Castro, Alejandro M., Li, Jiajia & Carrica, Pablo M. 2016 A mechanistic model of bubble entrainment in turbulent free surface flows. *Int. J. Multiphase Flow* **86**, 35–55. doi: 10.1016/j.ijmultiphaseflow.2016.07.005.

Chan, Wai Hong Ronald, Dodd, Michael S., Johnson, Perry L. & Moin, Parviz 2021*a* Identifying and tracking bubbles and drops in simulations: A toolbox for obtaining sizes, lineages, and breakup and coalescence statistics. *J. Comput. Phys.* **432**, 110156. doi: 10.1016/j.jcp.2021.110156.

Chan, Wai Hong Ronald, Johnson, Perry L. & Moin, Parviz 2021*b* The turbulent bubble break-up cascade. Part 1. Theoretical developments. *J. Fluid Mech.* **912**, A42. doi: 10.1017/jfm.2020.1083.

Chan, Wai Hong Ronald, Johnson, Perry L., Moin, Parviz & Urzay, Javier 2021*c* The turbulent bubble break-up cascade. Part 2. Numerical simulations of breaking waves. *J. Fluid Mech.* **912**, A43. doi: 10.1017/jfm.2020.1084.

Chanson, Hubert 1996 *Air bubble entrainment in free-surface turbulent shear flows*. Academic Press.

Chanson, Hubert & Toombes, Luke 2003 Strong interactions between free-surface aeration and turbulence in an open channel flow. *Exp. Therm. Fluid Sci.* **27** (5), 525–535. doi: 10.1016/S0894-1777(02)00266-2.

Clift, Roland, Grace, John R. & Weber, Martin E. 2013 *Bubbles, drops, and particles*. Dover Publ.

Davies, R. M. & Taylor, Geoffrey 1950 The Mechanics of Large Bubbles Rising through Extended Liquids and through Liquids in Tubes. *Proc. R. Soc. Lond. A* **200** (1062), 375–390.

Deane, Grant B., Preisig, James C. & Lavery, Andone C. 2013 The suspension of large bubbles near the sea surface by turbulence and their role in absorbing forward-scattered sound. *IEEE J. Ocean. Eng.* **38** (4), 632–641. doi: 10.1109/JOE.2013.2257573.

Deane, Grant B. & Stokes, M. Dale 2002 Scale dependence of bubble creation mechanisms in breaking waves. *Nature* **418**, 839–844. doi: 10.1038/nature00967.

Deike, Luc 2022 Mass Transfer at the Ocean–Atmosphere Interface: The Role of Wave Breaking, Droplets, and Bubbles. *Annu. Rev. Fluid Mech.* **54** (1), 191–224. doi: 10.1146/annurev-fluid-030121-014132.

Deike, Luc, Melville, W. Kendall & Popinet, Stéphane 2016 Air entrainment and bubble statistics in breaking waves. *J. Fluid Mech.* **801**, 91–129. doi: 10.1017/jfm.2016.372.

Diemer, R. B. & Olson, J. H. 2002 A moment methodology for coagulation and breakage problems: Part 3—generalized daughter distribution functions. *Chem. Eng. Sci.* **57**, 4187–4198. doi: 10.1016/S0009-2509(02)00366-4.

Dimas, Athanassios A. & Triantafyllou, George S. 1994 Nonlinear interaction of shear flow with a free surface. *J. Fluid Mech.* **260**, 211–246. doi: 10.1017/S0022112094003496.

Dommermuth, Douglas G., Sussman, Mark, Beck, Robert F., O'Shea, Thomas T., Wyatt, Donald C., Olson, Kevin & MacNeice, Peter 2004 The Numerical Simulation of Ship Waves Using Cartesian Grid Methods with Adaptive Mesh Refinement. In *25th Symposium on Naval Hydrodynamics*. St John's, arXiv: 1410.1942.

Falgout, Robert D., Jones, Jim E. & Yang, Ulrike Meier 2006 The design and implementation of hypre, a library of parallel high performance preconditioners. In *Numerical Solution of Partial Differential Equations on Parallel Computers*, pp. 267–294. Berlin, Heidelberg: Springer.

Falvey, Henry T. & Ervine, D. Alan 1988 Aeration in jets and high velocity flows. In *Model-Prototype Correlation of Hydraulic Structures*, pp. 25–55. ASCE.

Farmer, D. M., McNeil, C. L. & Johnson, B. D. 1993 Evidence for the importance of bubbles in increasing air–sea gas flux. *Nature* **361** (6413), 620–623. doi: 10.1038/361620a0.

Francois, Marianne M., Cummins, Sharen J., Dendy, Edward D., Kothe, Douglas B., Sicilian, James M. & Williams, Matthew W. 2006 A balanced-force algorithm for continuous and sharp interfacial surface tension models within a volume tracking framework. *J. Comput. Phys.* **213** (1), 141–173. doi: 10.1016/j.jcp.2005.08.004.

Gao, Qiang, Deane, Grant B., Liu, Han & Shen, Lian 2021 A robust and accurate technique for Lagrangian tracking of bubbles and detecting fragmentation and coalescence. *Int. J. Multiphase Flow* **135**, 103523. doi: 10.1016/j.ijmultiphaseflow.2020.103523.

Garrett, Chris, Li, Ming & Farmer, David 2000 The connection between bubble size spectra and energy dissipation rates in the upper ocean. *J. Phys. Oceanogr.* **30** (9), 2163–2171. doi: 10.1175/1520-0485(2000)030<2163:TCBBSS>2.0.CO;2.

Gaylo, Declan B., Hendrickson, Kelli & Yue, Dick K.P. 2021 Effects of power-law entrainment on bubble fragmentation cascades. *J. Fluid Mech.* **917**, R1. doi: 10.1017/jfm.2021.333.

Gaylo, Declan B., Hendrickson, Kelli & Yue, Dick K.P. 2022 An Eulerian label advection method for conservative volume-based tracking of bubbles/droplets. *J. Comput. Phys.* **470**, 111560. doi: doi.org/10.1016/j.jcp.2022.111560.

Gaylo, Declan B., Hendrickson, Kelli & Yue, Dick K.P. 2023 Fundamental time scales of bubble fragmentation in homogeneous isotropic turbulence. *J. Fluid Mech.* **962**, A25. doi: 10.1017/jfm.2023.281.

Gaylo, Declan B., Hendrickson, Kelli & Yue, Dick K.P. 2024 Effect of degassing on bubble populations in air-entraining free-surface turbulent flows. *J. Fluid Mech.* **995**, A12. doi: 10.1017/jfm.2024.780.

Gaylo, Declan B. & Yue, Dick K.P. 2025 Size distribution of large air bubbles entrained by strong free-surface turbulence. *J. Fluid Mech.* **1020**, A40. doi: 10.1017/jfm.2025.10669.

George, William K., Beuther, Paul D. & Arndt, Roger E. A. 1984 Pressure spectra in turbulent free shear flows. *J. Fluid Mech.* **148**, 155–191. doi: 10.1017/S0022112084002299.

Gulliver, John S. & Rindels, Alan J. 1993 Measurement of air-water oxygen transfer at hydraulic structures. *J. Hydraul. Eng.* **119** (3), 327–349. doi: 10.1061/(ASCE)0733-9429(1993)119:3(327).

Guo, Xin & Shen, Lian 2009 On the generation and maintenance of waves and turbulence in simulations of free-surface turbulence. *Comput. Fluids* **228** (19), 7313–7332. doi: 10.1016/j.jcp.2009.06.030.

Guo, Xin & Shen, Lian 2010 Interaction of a deformable free surface with statistically steady homogeneous turbulence. *J. Fluid Mech.* **658**, 33–62. doi: 10.1017/S0022112010001539.

Harlow, Francis H. & Welch, J. Eddie 1965 Numerical Calculation of Time-Dependent Viscous Incompressible Flow of Fluid with Free Surface. *The Physics of Fluids* **8** (12), 2182–2189. doi: 10.1063/1.1761178.

Harrison, Cyrus, Childs, Hank & Gaither, Kelly P. 2011 Data-Parallel Mesh Connected Components Labeling and Analysis. In *Eurographics Symposium on Parallel Graphics and Visualization* (ed. Torsten Kuhlen, Renato Pajarola & Kun Zhou). The Eurographics Association.

He, Lifeng, Chao, Yuyan & Suzuki, Kenji 2007 A Linear-Time Two-Scan Labeling Algorithm. In *2007 IEEE International Conference on Image Processing*, , vol. 5, pp. 241–244. IEEE.

Hendrickson, Kelli, Weymouth, Gabriel D., Yu, Xiangming & Yue, Dick K. P. 2019 Wake behind a three-dimensional dry transom stern. Part 1. Flow structure and large-scale air entrainment. *J. Fluid Mech.* **875**, 854–883. doi: 10.1017/jfm.2019.505.

Hendrickson, Kelli, Weymouth, Gabriel D. & Yue, Dick K.P. 2020 Informed component label algorithm for robust identification of connected components with volume-of-fluid method. *Comput. Fluids* **197**, 104373. doi: 10.1016/j.compfluid.2019.104373.

Hendrickson, Kelli & Yue, Dick K.P. 2019 Wake behind a three-dimensional dry transom stern. part 2. analysis and modelling of incompressible highly variable density turbulence. *J. Fluid Mech.* **875**, 884–913. doi: 10.1017/jfm.2019.506.

Herrmann, M. 2010 A parallel Eulerian interface tracking/Lagrangian point particle multi-scale coupling procedure. *J. Comput. Phys.* **229** (3), 745–759. doi: 10.1016/j.jcp.2009.10.009.

Hinze, J. O. 1955 Fundamentals of the hydrodynamic mechanism of splitting in dispersion processes. *AIChE J.* **1**, 289–295. doi: 10.1002/aic.690010303.

Iverson, J., Kamath, C. & Karypis, G. 2015 Evaluation of connected-component labeling algorithms for distributed-memory systems. *Parallel Comput.* **44**, 53–68. doi: 10.1016/j.parco.2015.02.005.

Keller, Robert J., Lai, K. K. & Wood, Ian R. 1974 Developing region in self-aerated flows. *J. Hydr. Div.* **100** (4), 553–568. doi: 10.1061/JYCEAJ.0003932.

Kiger, Kenneth T. & Duncan, James H. 2012 Air-entrainment mechanisms in plunging jets and breaking waves. *Annu. Rev. Fluid Mech.* **44** (1), 563–596. doi: 10.1146/annurev-fluid-122109-160724.

Killen, John Mark 1968 The surface characteristics of self aerated flow in steep channels. PhD thesis, University of Minnesota, Minneapolis, MN.

Kraichnan, Robert H. 1967 Inertial ranges in two-dimensional turbulence. *Phys. Fluids* **10** (7), 1417–1423. doi: 10.1063/1.1762301.

Kramer, Matthias & Valero, Daniel 2023 Linking turbulent waves and bubble diffusion in self-aerated open-channel flows: two-state air concentration. *J. Fluid Mech.* **966**, A37. doi: 10.1017/jfm.2023.440.

Lamarre, Eric & Melville, W. K. 1991 Air entrainment and dissipation in breaking waves. *Nature* **351** (6326), 469–472. doi: 10.1038/351469a0.

Langlois, Timothy R., Zheng, Changxi & James, Doug L. 2016 Toward animating water with complex acoustic bubbles. *ACM Trans. Graph.* **35** (4), 1–13. doi: 10.1145/2897824.2925904.

Liao, Yixiang & Lucas, Dirk 2009 A literature review of theoretical models for drop and bubble breakup in turbulent dispersions. *Chem. Eng. Sci.* **64**, 3389–3406. doi: 10.1016/j.ces.2009.04.026.

Longuet-Higgins, Michael S. 1998 Instabilities of a horizontal shear flow with a free surface. *J. Fluid Mech.* **364**, 147–162. doi: 10.1017/S0022112098008957.

Lundgren, T. S. 2003 Linearly forced isotropic turbulence. *Tech. Rep.*. Center for Turbulence Research, Stanford University.

Martínez-Bazán, C., Montañés, J. L. & Lasheras, J. C. 1999*a* On the breakup of an air bubble injected into a fully developed turbulent flow. part 1. breakup frequency. *J. Fluid Mech.* **401**, 157–182. doi: 10.1017/S0022112099006680.

Martínez-Bazán, C., Montañés, J. L. & Lasheras, J. C. 1999*b* On the breakup of an air bubble injected into a fully developed turbulent flow. part 2. size pdf of the resulting daughter bubbles. *J. Fluid Mech.* **401**, 183–207. doi: 10.1017/S0022112099006692.

Martínez-Bazán, C., Rodríguez-Rodríguez, J., Deane, G. B., Montaes, J. L. & Lasheras, J. C. 2010 Considerations on bubble fragmentation models. *J. Fluid Mech.* **661**, 159–177. doi: 10.1017/S0022112010003186.

Medwin, Herman & Beaky, Matthew M. 1989 Bubble sources of the knudsen sea noise spectra. *J. Acoust. Soc. Am.* **86** (3), 1124–1130. doi: 10.1121/1.398104.

Medwin, Herman & Clay, Clarence S. 1998 *Chapter 8 - Bubbles*, chap. Chapter 8 - Bubbles, pp. 287–347. San Diego: Academic Press.

Melville, W K 1996 The role of surface-wave breaking in air-sea interaction. *Annu. Rev. Fluid Mech.* **28** (1), 279–321. doi: 10.1146/annurev.fl.28.010196.001431.

Mendelson, Harvey D. 1967 The prediction of bubble terminal velocities from wave theory. *AIChE J.* **13** (2), 250–253. doi: https://doi.org/10.1002/aic.690130213.

NDRC 1946 Acoustic properties of wakes. In *Physics of Sound in the Sea*, , vol. 8, pp. 441–559. Washington, D. C.: Office of Scientific Research and Development.

Park, Sung Hoon, Park, Changhwan, Lee, JinYong & Lee, Byungchul 2017 A Simple Parameterization for the Rising Velocity of Bubbles in a Liquid Pool. *Nucl. Eng. Technol.* **49**, 692–699. doi: 10.1016/j.net.2016.12.006.

Patwardhan, A.W., Mali, R.G., Jadhao, S.B., Bhor, K.D., Padmakumar, G. & Vaidyanathan, G. 2012 Argon entrainment into liquid sodium in fast breeder reactor. *Nucl. Eng. Des.* **249**, 204–211. doi: 10.1016/j.nucengdes.2011.07.046.

Pope, Stephen B. 2000 *Turbulent Flows*. Cambridge: Cambridge University Press.

Popinet, Stéphane 2009 An accurate adaptive solver for surface-tension-driven interfacial flows. *J. Comput. Phys.* **228** (16), 5838–5866. doi: 10.1016/j.jcp.2009.04.042.

Popinet, Stéphane 2018 Numerical Models of Surface Tension. *Annu. Rev. Fluid Mech.* **50** (1), 49–75. doi: 10.1146/annurev-fluid-122316-045034.

Qi, Yinghe, Mohammad Masuk, Ashik Ullah & Ni, Rui 2020 Towards a model of bubble breakup in turbulence through experimental constraints. *Int. J. Multiphase Flow* **132**, 103397. doi: 10.1016/j.ijmultiphaseflow.2020.103397.

Qi, Yinghe, Tan, Shiyong, Corbitt, Noah, Urbanik, Carl, Salibindla, Ashwanth K. R. & Ni, Rui 2022 Fragmentation in turbulence by small eddies. *Nat. Commun.* **13** (1), 469. doi: 10.1038/s41467-022-28092-3.

Rein, Martin 1998 Turbulent open-channel flows: Drop-generation and self-aeration. *J. Hydraul. Eng.* **124** (1), 98–102. doi: 10.1061/(ASCE)0733-9429(1998)124:1(98).

Risso, Frédéric & Fabre, Jean 1998 Oscillations and breakup of a bubble immersed in a turbulent field. *J. Fluid Mech.* **372**, 323–355. doi: 10.1017/S0022112098002705.

Rivière, Aliénor, Mostert, Wouter, Perrard, Stéphane & Deike, Luc 2021 Sub-Hinze scale bubble production in turbulent bubble break-up. *J. Fluid Mech.* **917**, A40. doi: 10.1017/jfm.2021.243.

Rivière, Aliénor, Ruth, Daniel, Mostert, Wouter, Deike, Luc & Perrard, Stephane 2021 Origin of the sub-hinze scale bubble production in turbulence. In *APS Division of Fluid Dynamics Meeting Abstracts*, pp. P28–003.

Rivière, Aliénor, Ruth, Daniel J., Mostert, Wouter, Deike, Luc & Perrard, Stéphane 2022 Capillary driven fragmentation of large gas bubbles in turbulence. *Phys. Rev. Fluids* **7** (8), 083602. doi: 10.1103/PhysRevFluids.7.083602.

Rosales, Carlos & Meneveau, Charles 2005 Linear forcing in numerical simulations of isotropic turbulence: Physical space implementations and convergence properties. *Phys. Fluids* **17**, 095106. doi: 10.1063/1.2047568.

Rubel, Clark & Owkes, Mark 2019 Extraction of droplet genealogies from high-fidelity atomization simulations. *At. Sprays* **29**, 709–739. doi: 10.1615/AtomizSpr.2020031624.

Ruth, Daniel J. & Coletti, Filippo 2024 Structure and energy transfer in homogeneous turbulence below a free surface. *J. Fluid Mech.* **1001**, A46. doi: 10.1017/jfm.2024.1017.

Ruth, Daniel J., Vernet, Marlone, Perrard, Stéphane & Deike, Luc 2021 The effect of nonlinear drag on the rise velocity of bubbles in turbulence. *J. Fluid Mech.* **924**, A2. doi: 10.1017/jfm.2021.556.

Salibindla, Ashwanth K. R., Masuk, Ashik Ullah Mohammad, Tan, Shiyong & Ni, Rui 2020 Lift and drag coefficients of deformable bubbles in intense turbulence determined from bubble rise velocity. *J. Fluid Mech.* **894**, A20. doi: 10.1017/jfm.2020.244.

Scardovelli, Ruben & Zaleski, Stephane 2000 Analytical Relations Connecting Linear Interfaces and Volume Fractions in Rectangular Grids. *J. Comput. Phys.* **164** (1), 228–237. doi: 10.1006/jcph.2000.6567.

Shen, Lian, Triantafyllou, George S. & Yue, Dick K. P. 2000 Turbulent diffusion near a free surface. *J. Fluid Mech.* **407**, 145–166. doi: 10.1017/S0022112099007466.

Shen, Lian, Zhang, Xiang, Yue, Dick K. P. & Triantafyllou, George S. 1999 The surface layer for free-surface turbulent flows. *J. Fluid Mech.* **386**, 167–212. doi: 10.1017/S0022112099004590.

Solsvik, Jannike, Maaß, Sebastian & Jakobsen, Hugo A. 2016 Definition of the Single Drop Breakup Event. *Ind. Eng. Chem. Res.* **55** (10), 2872–2882. doi: 10.1021/acs.iecr.6b00591.

Sporleder, Federico, Borka, Zsolt, Solsvik, Jannike & Jakobsen, Hugo A. 2012 On the population balance equation. *Rev. Chem. Eng.* **28**, 149–169. doi: 10.1515/revce-2011-0013.

Thorpe, Stephen Austen 1982 On the clouds of bubbles formed by breaking wind-waves in deep water, and their role in air-sea gas transfer. *Philos. Trans. R. Soc. London, Ser. A* **304**, 155–210. doi: 10.1098/rsta.1982.0011.

Trevorrow, Mark V., Vagle, Svein & Farmer, David M. 1994 Acoustical measurements of microbubbles within ship wakes. *J. Acoust. Soc. Am* **95** (4), 1922–1930. doi: 10.1121/1.408706.

Tryggvason, Grétar, Scardovelli, Ruben & Zaleski, Stéphane 2011 *Direct Numerical Simulations of Gas–Liquid Multiphase Flows*. Cambridge University Press.

Tsouris, C. & Tavlarides, L. L. 1994 Breakage and coalescence models for drops in turbulent dispersions. *AIChE J.* **40**, 395–406. doi: 10.1002/aic.690400303.

Valentas, Kenneth J., Bilous, Olegh & Amundson, Neal R. 1966 Analysis of breakage in dispersed phase systems. *Ind. Eng. Chem. Fundamen.* **5**, 271–279. doi: 10.1021/i160018a019.

Vejražka, Jiří, Zedníková, Mária & Stanovský, Petr 2018 Experiments on breakup of bubbles in a turbulent flow. *AIChE J.* **64** (2), 740–757. doi: 10.1002/aic.15935.

Veron, Fabrice 2015 Ocean spray. *Annu. Rev. Fluid Mech.* **47** (1), 507–538. doi: 10.1146/annurev-fluid-010814-014651.

Wacławczyk, Tomasz 2021 Modeling of non-equilibrium effects in intermittency region between two phases. *Int. J. Multiph. Flow* **134**, 103459. doi: 10.1016/j.ijmultiphaseflow.2020.103459.

Wallace, D. W. R. & Wirick, C. D. 1992 Large air–sea gas fluxes associated with breaking waves. *Nature* **356** (6371), 694–696. doi: 10.1038/356694a0.

Wallis, Graham B. 1974 The terminal speed of single drops or bubbles in an infinite medium. *Int. J. Multiph. Flow* **1** (4), 491–511. doi: https://doi.org/10.1016/0301-9322(74)90003-2.

Wei, Wangru, Xu, Weilin, Deng, Jun, Tian, Zhong & Zhang, Faxing 2019 Bubble formation and scale dependence in free-surface air entrainment. *Sci. Rep.* **9** (1), 11008. doi: 10.1038/s41598-019-46883-5.

Weymouth, G.D. & Yue, Dick K.P. 2010 Conservative Volume-of-Fluid method for free-surface simulations on Cartesian-grids. *J. Comput. Phys.* **229** (8), 2853–2865. doi: 10.1016/j.jcp.2009.12.018.

Wilhelms, Steven C. & Gulliver, John S. 2005 Bubbles and waves description of self-aerated spillway flow. *J. Hydraul. Res.* **43** (5), 522–531. doi: 10.1080/00221680509500150.

Yu, Xiangming 2019 Theoretical and Numerical Study of Air Entrainment and Bubble Size Distribution in Strong Free-surface Turbulent Flow at Large Froude and Weber Number. PhD thesis, MIT.

Yu, Xiangming, Hendrickson, Kelli, Campbell, Bryce K. & Yue, Dick K.P. 2019 Numerical investigation of shear-flow free-surface turbulence and air entrainment at large Froude and Weber numbers. *J. Fluid Mech.* **880**, 209–238. doi: 10.1017/jfm.2019.695.

Yu, Xiangming, Hendrickson, Kelli & Yue, Dick K.P. 2020 Scale separation and dependence of entrainment bubble-size distribution in free-surface turbulence. *J. Fluid Mech.* **885**, R2. doi: 10.1017/jfm.2019.986.

Zabaleta, F., Bombardelli, F. A. & Márquez Damián, S. 2024 Numerical modeling of self-aerated flows: Turbulence modeling and the onset of air entrainment. *Phys. Fluids* **36** (4), 043337. doi: 10.1063/5.0197678.