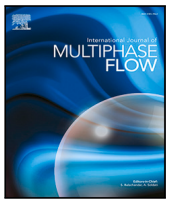
International Journal of Multiphase Flow 202 (2026) 105840

Contents lists available at ScienceDirect

International Journal of Multiphase Flow

journal homepage: www.elsevier.com/locate/ijmulflow



Research Paper

Connected-component labeling: A tool for the numerical investigation of buoyancy-driven bubbly flows

Ali Fakhreddine [a], Krishnan Mahesh [b,c,*]

[a] Department of Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, MN, USA
[b] Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, MI, USA
[c] Center for Naval Research and Education, University of Michigan, Ann Arbor, MI, USA

ARTICLE INFO

Keywords:
Connected-component labeling
VOF
DNS
Bubbly flow

ABSTRACT

We investigate the motion of buoyant bubbles rising in a quiescent liquid using a Volume-of-Fluid Connected-Component Labeling (VOFCCL) technique. Although the Eulerian VOF framework excels at interface-capturing and provides valuable insight into the behavior of the dispersed phase, it inherently loses local detail due to the spatial-averaging required for volume-fraction-dependent quantities. To address the loss of this topological metadata, we consider a Lagrangian tagging approach that recovers key information about individual bubbles without modifying the underlying multiphase solver. Building on the connected-component labeling algorithm of Herrmann (2010), we implement and extend a VOFCCL method to identify and track bubbles in VOF simulations. The extended algorithm makes use of the color function C to adapt the original Level-set-based method to VOF, and incorporates an additional retagging procedure that streamlines the post-processing of bubble information. Verification of the VOFCCL algorithm is first carried out, followed by validation of the two-phase flow solver through single-bubble rise simulations at different Morton numbers (Mo). Finally, we demonstrate the viability of the method as a tool to analyze bubbly flows and extract bubble-resolved statistical quantities such as bubble count evolution and spatial distribution, which are otherwise inaccessible quantities in traditional VOF analyses.

1. Introduction

Multiphase flows can be defined as systems in which different phases, or fluid and solid phases, are simultaneously present (Prosperetti and Tryggvason, 2009). Such flows are abundant in nature and in man-made applications. In technology, some of the applications include fluidized bed reactors, oil/gas transport in pipe lines, and diesel spray systems (Mudde, 2005). In nature, cloud formation, cavitation in the vascular tissues of plants, and sediment transport in rivers are all instances of multiphase flow phenomena. Understanding the dynamics of such flows is of paramount engineering and scientific importance and the literature is extensive. In liquid–gas systems, a flow pattern of particular interest is buoyancy-driven bubbly flows. Buoyancy-driven bubbly flows (BBF) are a unique type of two-phase flow characterized by the motion of gas/vapor bubbles through a liquid due to buoyant forces. These flows are encountered in a handful of industrial applications (e.g. aeration of water treatment systems, cooling of nuclear reactors) and have encouraged a large number of numerical and experimental investigations (Bhaga and Weber, 1981; Bröder and Sommerfeld, 2007; Bunner and Tryggvason, 2002a; Hua and Lou, 2007; Balcázar et al., 2015). Despite these efforts, our current understanding is still limited by the analysis tools available, both experimentally and numerically (Tryggvason et al., 2011). In experiments, the use of laser-based techniques within a bubble swarm is limited to very dilute flows due to the dispersed nature of the bubbly flow (Ravisankar and Zenit, 2024). Additionally, hot-wire-based techniques similar to those in Mendez-Diaz et al. (2013) and Alméras et al. (2017) require comprehensive signal processing which can become demanding at high volume fractions, and also physically invasive inside the carrier fluid. In simulations, the analysis is a trade-off between accuracy and computational efficiency based on the nature of the numerical method used. Therefore, improving methodologies and devising new complementary analysis tools is essential for a detailed understanding of bubbly flows in practical applications where length scales are small and time scales are short.

When simulating multiphase flows, the Navier–Stokes equations are coupled with a method to represent the interface or distribution of constituent phases. The strategies employed for this purpose are abundant and are typically categorized as Interface-Capturing (IC),

* Corresponding author.
E-mail address: krmahesh@umich.edu (K. Mahesh).

https://doi.org/10.1016/j.ijmultiphaseflow.2026.105840
Received 28 January 2026; Received in revised form 4 June 2026; Accepted 6 July 2026
Available online 17 July 2026

Interface-Tracking (IT), Euler–Lagrange (EL), or Euler–Euler (EE) methods. Amongst the interface-capturing methods, namely the volume-of-fluid (VOF) method (Hirt and Nichols, 1981), Level-set Method (LSM) (Osher and Sethian, 1988), and Phase-Field (PF) method (Jacqmin, 1999), VOF remains one of the most common methods used to track interfaces due to its inherent ability to maintain a sharp interface representation and conserve mass — if the numerics permit. The use of VOF, LSM, or any other Eulerian method often provides limited knowledge about the dispersed phase from the statistical point of view. Nevertheless, indicator-function-based analyses have proven valuable for characterizing the structure of liquid–gas flows. For example, Thiesset et al. (2021) employed two-point statistics of the liquid-phase indicator function to relate the spatial organization of the liquid field to interfacial geometrical properties, and also to investigate transport processes. Such approaches provide a field-based statistical description of the dispersed phase. For bubbly flows, however, a comprehensive analysis requires more attention since bubble-resolved quantities of practical interest (e.g. bubble count, positions, sizes, inter-bubble spacing) are not recoverable when computing spatial averages using the interface indicator function. Hence, the Eulerian description by itself is not sufficient, and a complementary Lagrangian framework is needed for a more complete understanding. Some of the earlier work investigating the coupling between Eulerian interface capturing and Lagrangian point-particle tracking is that of Herrmann (2010). They devised a method to track the dynamically evolving phase interface of an atomizing liquid jet using LSM in regions of the flow where the interface is sufficiently resolved. In parts of the flow where the phase interface geometry can no longer be resolved adequately, a Lagrangian point-particle approach was used to describe the small, separated liquid structures. This approach, the Connected Component Labeling (CCL) method, uses the sign of the indicator function ($\phi < 0$ for LSM) to identify the separated structures. The coupling between the Eulerian description and the Lagrangian counterpart is achieved through an efficient parallel algorithm that removes separated, under-resolved liquid structures from the Eulerian frame of reference and inserts them in the Lagrangian one whilst preserving their position, mass, momentum, and low-order shape. More recently, Hendrickson et al. (2020) suggested two modifications to the CCL method that improve its robustness and accuracy for VOF. Their Informed Component Labeling algorithm (ICL) incorporates the normal and uses multilevel thresholding to improve and refine connectivity decisions for components with spacing just larger than the grid size. In another recent work, Gao et al. (2021) presented a Lagrangian bubble tracking technique referred to as the Optimal Network (ON) method that also uses LSM. Their technique is based on the idea of establishing a network of mappings between bubbles identified between adjacent time instants. The mappings are determined by selecting the minimum from a set of pseudo-distance errors, which are themselves based on constraints imposed on bubble position, velocity, and volume between adjacent time instants. The method is made suitable for tracking bubble history through the established mappings. Beyond interface-tracking, such Eulerian–Lagrangian formulations provide a natural framework for extracting bubble-resolved statistical information that is difficult to obtain from an Eulerian description alone.

Motivated by these considerations, the present work builds on the connected-component labeling algorithm of Herrmann (2010) to perform direct numerical simulations (DNS) of dilute buoyancy-driven bubbly flows within a volume-of-fluid (VOF) framework. The demonstrated approach, referred to here as VOFCCL, adapts the original methodology to a directionally split VOF formulation and incorporates a retagging procedure that ensures consistent identification and processing of individual bubbles over time. This enables the introduction of a Lagrangian representation of the dispersed phase in regions where an Eulerian description alone is insufficient, while requiring no modification to the underlying fluid solver and relying solely on the color function C. The proposed framework provides the necessary infrastructure to support a diverse array of research areas. For instance, there is a growing interest in studying the effect of complex bubble deformations on flow evolution as well as bubble–bubble interactions (Yuan et al., 2020, 2021; Zhou et al., 2026); VOFCCL would enable the Lagrangian analysis of comprehensive datasets generated from such studies.

In the present study, VOFCCL is used to demonstrate the extraction of bubble location and inter-bubble distance statistics; more generally, the framework naturally allows access to other bubble-resolved statistical quantities, such as bubble size, velocity, and bubble–wall interaction statistics, when required. The paper is organized as follows: Section 2 describes the numerical methodology, including the governing equations and the VOF formulation; Section 3 presents the VOFCCL algorithm and verification test cases; Section 4 validates the fluid solver using the buoyancy-driven motion of a single bubble; Section 5 reports DNS of buoyancy-driven bubbly flows in the spherical regime and their analysis using VOFCCL; finally, conclusions are drawn in Section 6.

2. Numerical methodology

2.1. Governing equations

The governing equations are solved using the finite-volume algorithm developed by Mahesh et al. (2004) for the incompressible Navier–Stokes (NS) equations. The algorithm is robust and emphasizes discrete kinetic energy conservation in the inviscid limit which enables the simulation of high-Re flows without the need for added numerical dissipation. The in-house solver utilized is based on a one-fluid formulation of the conservation of mass and momentum to describe the dynamics of a two-phase flow such that,

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial (u_i u_j)}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{1}{\rho}\frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right] + \frac{F_{st,i}}{\rho} + \frac{F_{b,i}}{\rho} \tag{2}$$

where t is time, $x_i = (x, y, z)$ denotes the Cartesian coordinate axes, $u_i = (u, v, w)$ denotes the velocity, p is pressure, ρ is density, and μ is the dynamic viscosity. Additionally, $F_{st,i}$ and $F_{b,i}$ represent the surface tension force and body force per unit volume, respectively. The surface tension force is modeled as a continuum surface force as proposed by (Brackbill et al., 1992)

$$F_{st,i} = \sigma\kappa\frac{\partial C}{\partial x_i} \tag{3}$$

where σ is the surface tension constant, κ is the local interface curvature, and C is the interface indicator function which is often referred to as the color function in the volume-of-fluid (VOF) method. Since the two-phase system is assumed to be immiscible with a jump discontinuity at the interface, the density and viscosity can be evaluated as

$$\rho = \rho_1 C + \rho_2(1 - C) \tag{4}$$

$$\mu = \mu_1 C + \mu_2(1 - C). \tag{5}$$

Here, the subscripts 1 and 2 denote phase 1 and phase 2, respectively. In the context of multiphase flow, the form of Eq. (2) does not give direct insight into the relative importance between the different terms. Therefore, re-writing Eq. (2) using dimensionless groups is necessary. For buoyancy-driven flows, the following dimensionless characteristic variables are introduced:

$$\rho^* = \frac{\rho}{\rho_1}; \quad \mu^* = \frac{\mu}{\mu_1}; \quad \kappa^* = \frac{\kappa}{D^{-1}} \quad .$$

$$\tau^* = \frac{t}{D^{1/2}g^{-1/2}}; \quad x^* = \frac{x}{D}; \quad u^* = \frac{u}{(gD)^{1/2}}; \quad p^* = \frac{p}{\rho_1 g D};$$

Since the driving body force is gravity, $F_{b,i}$ is equal to ρg_i. To prevent the entire flow from accelerating, an additional body force defined by

$\rho_0 g_i$ is required as stated by Bunner and Tryggvason (2002a) where ρ_0 is the space-averaged density at $t = 0$ such that

$$\rho_0 = \frac{1}{V}\int_\Omega \Big(\rho_1 C + \rho_2(1-C)\Big)\Big|_{t=0} dV \quad . \tag{6}$$

The added term is analogous to the pressure gradient generated by the base of a flow container, which balances the total gravitational force on the fluid. After modifying the body force term and omitting the superscript $(*)$, the non-dimensionalized momentum equation (formerly Eq. (2)) is given by:

$$\frac{\partial u_i}{\partial t} + \frac{\partial (u_i u_j)}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re}\frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right] + \frac{1}{Bo}F_{st,i} + \left(1 - \frac{\rho_0}{\rho}\right)g_i \tag{7}$$

where

$$Re = \frac{\rho_1 g^{1/2} D^{3/2}}{\mu_1} \quad \text{and} \quad Bo = \frac{\rho_1 g D^2}{\sigma}$$

are the Reynolds number and Bond number (also known as Eötvös number Eo), respectively. Based on Eq. (7), the buoyancy-driven motion of bubbles can be characterized by four dimensionless parameters – the density ratio $\eta_\rho = (\rho_1/\rho_2)$ between the carrier phase and the dispersed phase, the viscosity ratio $\eta_\mu = (\mu_1/\mu_2)$, the Reynolds number, and the Bond number. An additional dimensionless number, the Morton number,

$$Mo = \frac{g\mu_1^4 \Delta\rho}{\rho_1^2 \sigma^3} = \frac{Bo^3}{Re^4}$$

is often considered alongside the Bond number to characterize the shapes of bubbles moving in the continuous phase. It is also practical in selecting simulation parameters in what follows. It is worth mentioning that experimental results on bubbles rising in a quiescent liquid are mostly presented using a Reynolds number that uses bubble terminal velocity U_T as the velocity scale such that $Re = \rho_1 D U_T/\mu_1$ (Hua and Lou, 2007).

A predictor–corrector, fractional step method is used to solve Eq. (7) where the velocities are first predicted using the momentum equation and then corrected using the pressure gradient obtained from solving the Poisson equation. The non-linear convective term is denoted by NL, the viscous term is denoted by $VISC$, and the body force term is denoted by BF. Explicit time advancement is performed using the Adams–Bashforth scheme which is $O(\Delta t^2)$. The predicted velocities $\hat{u}_i$ at the cell centers are first obtained from the previous time steps k and $k-1$:

$$\frac{u_i^* - u_i^k}{\Delta t} = \frac{1}{2}\Big[3(NL + VISC + BF)^k - (NL + VISC + BF)^{k-1}\Big] \quad , \tag{8}$$

$$\frac{\tilde{u}_i - u_i^k}{\Delta t} = \frac{1}{\rho^{k+1}}\sigma\kappa^{k+1}\frac{\partial C^{k+1}}{\partial x_i} \quad , \tag{9}$$

such that $\hat{u}_i = u_i^* + \tilde{u}_i$. The respective face velocities are then obtained by interpolation using symmetric averaging $O(\Delta x^2)$ to obtain the predicted face-normal velocities:

$$v_n^* = \left(\frac{u^*_{i,icv1} + u^*_{i,icv2}}{2}\right)n_i \quad \text{and} \quad \tilde{v}_n = \frac{1}{\rho_f^{k+1}}\sigma\kappa_f^{k+1}\frac{\partial C^{k+1}}{\partial x_i} \quad , \tag{10}$$

such that $\hat{v}_n = v_n^* + \tilde{v}_n$, where v_n points from control volume $icv1$ to $icv2$ in the direction of the face-normal $\mathbf{n}$. Note the distinction in velocity fields at the predictor step where the surface tension contribution is not simply added to NL, $VISC$, and BF. This is owed to the fact that simple averaging will cause an imbalance between the surface tension forces and the pressure gradient which inevitably leads to the generation of non-physical spurious currents, also known as parasitic currents. Spurious currents violate conservation principles and corrupt the phase interface over time, therefore, the above method ensures proper balance. The corrector step

$$\frac{u_i^{k+1} - \hat{u}_i}{\delta t} = -\frac{1}{\rho}\frac{\partial p^{k+1}}{\partial x_i} \quad , \tag{11}$$

is then projected onto the face-normal as:

$$\frac{v_n^{k+1} - \hat{v}_n}{\delta t} = -\frac{1}{\rho_f}\frac{\partial p^{k+1}}{\partial n} \quad . \tag{12}$$

The continuity equation imposes the incompressibility constraint

$$\sum_{\text{faces of cv}} v_n^{k+1} A_f = 0 \quad , \tag{13}$$

where A_f is the face area. Hence, integrating Eq. (12) over the computational domain and substituting Eq. (13) in the integrated equation yields the Poisson equation for p^{k+1}:

$$\sum_{\text{faces of cv}} \frac{\delta t}{\rho_f}\frac{\partial p^{k+1}}{\partial n} A_f = \sum_{\text{faces of cv}} \hat{v}_n A_f \tag{14}$$

which is solved using a pre-conditioned biconjugate gradient stabilized method (BiCGSTAB). In the presence of non-symmetric linear systems that are often encountered in multiphase flow problems, BiCGSTAB has been recommended for the solution of the Poisson equation or, more generally, the Helmholtz equation (Prosperetti and Tryggvason, 2009).

2.2. Volume-of-fluid method

Given a fluid boundary consisting of two phases sharing an interface at a particular location $\mathbf{x}$, the phase indicator function $H(\mathbf{x}, t)$ is defined as

$$H(\mathbf{x}, t) = \begin{cases} 1 & \text{if } \mathbf{x} \text{ is in fluid 1} \\ 0 & \text{if } \mathbf{x} \text{ is in fluid 2} \end{cases} \tag{15}$$

As the interface moves and deforms, the shape of each fluid changes, however, each fluid parcel is assumed to be immiscible in the absence of phase change so that the material derivative of H remains zero:

$$\frac{DH}{Dt} = \frac{\partial H}{\partial t} + \mathbf{u}\cdot\nabla H = 0 \quad . \tag{16}$$

Assuming incompressible conditions ($\nabla\cdot\mathbf{u} = 0$), Eq. (16) reduces to

$$\frac{\partial H}{\partial t} + \nabla\cdot(\mathbf{u}H) = 0 \quad . \tag{17}$$

The color function C represents a discrete version of the characteristic function H and is used to define the spatial average of H in each computational cell such that

$$C_{i,j,k} = \frac{1}{\delta x \delta y \delta z}\int_\Omega H(\mathbf{x}, t)\, dV \tag{18}$$

where Ω represents the computational domain, and the subscripts i, j, and k denote the index locations of the control volume on the Cartesian grid in the x, y, and z–directions, respectively. This results in a distribution of volume fraction values that range between $0 < C < 1$. In geometric VOF, the interface is approximated geometrically in a computational cell. The method is summarized in two fundamental steps – the reconstruction step and the advection step.

I. **Reconstruction step**: The most commonly used reconstruction approach is the piecewise linear interface calculation (PLIC) scheme (Debar, 1974), where the equation of the interface plane in three dimensions is written as

$$\mathbf{m}\cdot\mathbf{x} = m_x x + m_y y + m_z z = \beta \tag{19}$$

and β is adjusted iteratively until the area under the interface is equal to $\delta_x\delta_y\delta_z C_{i,j,k}$.

II. **Advection step**: Given the velocity field of the reconstructed interface, C is shared across neighboring cells in a manner that guarantees volume conservation.

In the current work, the interface normal $\mathbf{m}$ is estimated using Young's differencing method (Youngs, 1984).

3. Connected Component Labeling (CCL)

The purpose of the labeling procedure in this work is to identify each bubble suspended in the carrier liquid and mark it with a unique identifier. The base algorithm of Herrmann (2010) is adapted to the directionally-split VOF framework by replacing the level-set function ϕ in the original method by the color function C, and modifying interface identification criteria. Additions to the algorithm like the "retagging" procedure are discussed later in this section, and verification cases are presented last with the relevant details. To simplify the discussion of the tagging algorithm, details of the serial implementation are presented first followed by the parallel implementation.

3.1. Single block algorithm (serial)

First, one needs to search for untagged nodes in the computational domain that meet the criterion for $C = 0.5$. In other words, we are looking for the cells that represent the liquid–gas interface in order to initiate the tagging procedure of the closed volume associated with the first cell that we identified as an interfacial cell. After an interfacial cell is found, the algorithm grows into the gas in two layers denoted as S and C until the continuous structure is filled with a number id – essentially, CCL is a flood-fill algorithm. Once the last cell of the continuous structure is filled, the global id is incremented by 1 and the algorithm continues to search the computational domain using the updated global id. The algorithm exits when there are no more untagged interfacial cells left. The steps of the serial algorithm are visually demonstrated in Fig. 1.

3.2. Multi-block algorithm (parallel)

In the multi-block algorithm, the labeling procedure is done independently in each block (i.e. processor). If a bubble is located at the intersection of more than one block, the base algorithm produces multiple tag ids that erroneously label different parts of the same bubble. Therefore, the parallel algorithm involves an additional necessary step that unifies the different tags of a bubble shared across multiple processor boundaries through merging. This step produces a single id that is unique to the bubble regardless of its location with respect to the computational blocks. A schematic example of the merging step is shown in Fig. 2.

As Fig. 2(a) shows, the tagging process is carried out in each block, independently. The structures away from processor boundaries all have a uniform flood fill, on the other hand, the structure located at the intersection of processors has four different values of tag id for each of its four different parts, which is incorrect and necessitates merging. The merging procedure is an iterative process by which the values of the id in the ghost cells of a processor boundary are compared to those of the neighboring processor. If the values in the ghost cells are identical, that means that the parts crossing the processor boundary belong to the same entity, and there is no need to proceed. However, if the ghost cell values mismatch, the smallest of the two ids is chosen to fill the structure with the higher id. The process continues $locally$ on each processor until all parts of the suspended structure have the same identifier, and $globally$ until the tag id no longer changes in the entire computational domain. Note that the merging process is done cyclically in x, then in y, then in the z–direction. According to Herrmann (2010), only 2 to 6 of the described "redos" are needed to reach a uniformly labeled field. The steps of the full multi-block tagging algorithm (tagging+merging) are visually demonstrated in Fig. 3.

Upon the completion of the merging step, there is no guarantee that the maximum id value globally reflects the total number of distinct suspended entities present. This is illustrated in Fig. 2(b) where we see exactly 4 bubbles present but $MAX(id)$ is equal to 7. This numbering scheme can become inefficient when collecting statistical information about the dispersed phase in a fluid flow. For that reason and to make post-processing straightforward, a parallel bubble-retag algorithm was developed following the merging step to produce a continuous array such that $id = [1, 2, 3, \ldots, N]$. For example, if the result of tagging 5 bubbles is $id = [1, 5, 11, 21, 3]$ the corrected set would become $id = [1, 2, 3, 4, 5]$ after bubble-retag.

Since the fluid solver is based on structured grids, field information at a general location is stored using ijk indexing. Hence, the algorithm starts by converting the three-dimensional id array into a one-dimensional list — this makes the algorithm scalable and more computationally efficient. Once the 1D id list is created, an ordered list representing the index location of each element in the id list is initialized. The id list is sorted in ascending order to determine the number of different tag id values in each processor. Simultaneously, the index list is updated based on the new sorting of the id list. The sorted id list is then made unique to remove the redundant elements and a new index list is created simultaneously. Following that step, a map is created between the old index list and the new one to connect the identifier values to their respective order in the original list. Once the number of different tag id values present in each processor is determined, a displacement list is created that includes those numbers. For example, if $Rank\ 0$ has $id = [1, 2, 2, 3, 4, 4, 4]$, $Rank\ 1$ has $id = [3, 4, 5, 5, 5, 11, 11]$, and $Rank\ 2$ has $id = [3, 5, 9, 9, 12, 13, 13, 13]$, the displacement list is $[4, 4, 5]$ where the length of this list is equal to the number of processors used in a given simulation. The displacement list is later used to offset the id values in the proper order. Finally, the one-dimensional list is converted back to its original 3D ijk structure. A detailed description of the retagging steps is shown in Algorithm 1 and visually illustrated in Fig. 10. It is worth noting that "flattening" the three-dimensional arrays into one-dimensional lists and the corresponding computational sorting is performed using efficient recursive operations that have a negligible effect on the overall communication overhead.

3.3. Verification tests

The tagging procedure was tested using different dispersed phase configurations to ensure generality with respect to bubble shape and distribution, and to verify the parallel implementation. For the quantitative scaling study of the base algorithm, the reader is referred to Herrmann (2010). The verification tests are all done in the absence of any artificial flow (static). This is because in a full multiphase simulation where only ensemble averages are of interest, each labeling step produces a Lagrangian snapshot of the flow at a given time — each snapshot different from the one before it, therefore maintaining bubble history is not relevant for the current application.

3.3.1. Ellipsoids

The first test case is an array of three disconnected ellipsoids that are initialized in a unit cube domain resolved by as $128 \times 128 \times 128$ grid. The general equation used for the generation of the geometry is:

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} + \frac{(z - z_0)^2}{c^2} = 1 \tag{20}$$

where (x_0, y_0, z_0) are the coordinates of the ellipsoid center, and a, b and c are the semi-axis lengths in the x, y and z-directions, respectively. For all three ellipsoids, $x_0 = 0.5$, $y_0 = 0.5$, $a = 0.4$, $b = 0.08$, and $c = 0.08$. The location is varied in the z-direction such that $z_0 = [0.25, 0.5, 0.75]$. The ellipsoids are purposefully placed at processor intersection boundaries to verify the merging algorithm — 32 processors were used to run this calculation.

As Fig. 4(b) shows, the iso-surface of C is colored by three different ids after tagging — each color denotes a distinct suspended body. Fig. 5 shows 2D slices of the tag id iso-surface to further confirm the uniform tagging of the interior area of the ellipsoids across multiple processors. Note that the geometries were also rotated cyclically in different orientations (not shown here) to ensure directional invariance of the algorithm, all of which confirmed the same result. VOFCCL inherently handles extreme aspect ratio changes, similar to the "cusped" shapes investigated in the 3D DNS of Yuan et al. (2020), by mapping the exact topological changes captured by the Eulerian framework.

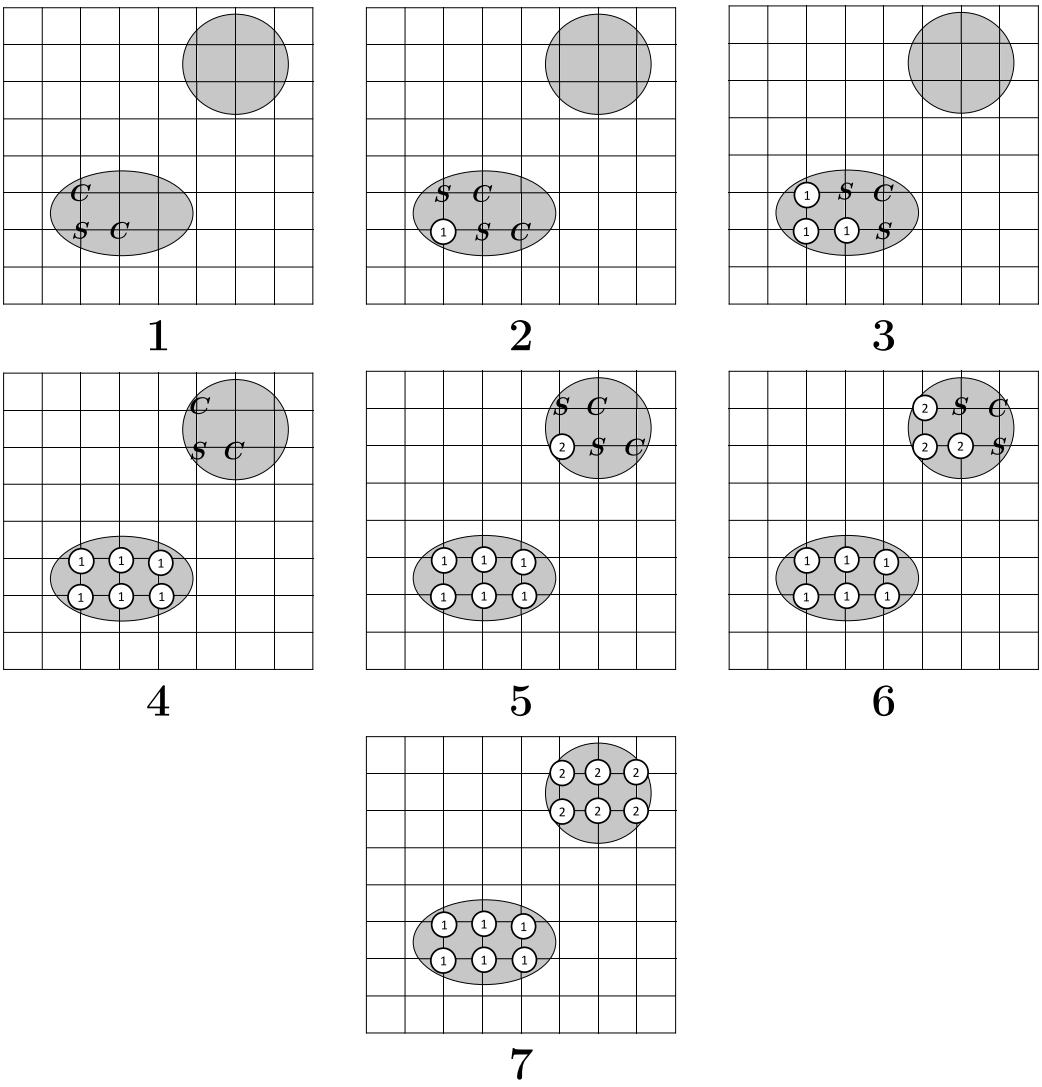


Fig. 1. Visual algorithm of the serial bubble tagging process.

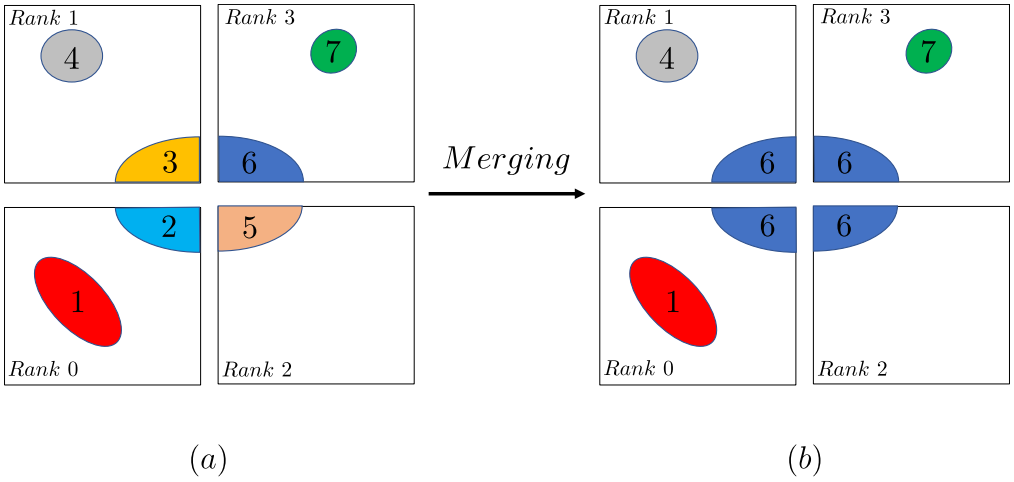


Fig. 2. Merging of marker-id of a structure shared across multiple processors – (a) before merging (b) after merging.

3.3.2. Ordered bubble array

An array of equally spaced bubbles was initialized in a cubic domain for the second static tagging test case. The bubbles do not overlap, therefore the largest id value in the tagging list should be exactly equal to the number of bubbles present. The grid size is $256 \times 256 \times 256$ ($\approx$17 million cells) and the bubble diameter is $D = 0.04$. The problem was simulated for $N = 8^3$, 10^3, and 13^3, respectively. Fig. 6 shows the iso-surface of C on the top row and the corresponding tagged iso-surface on

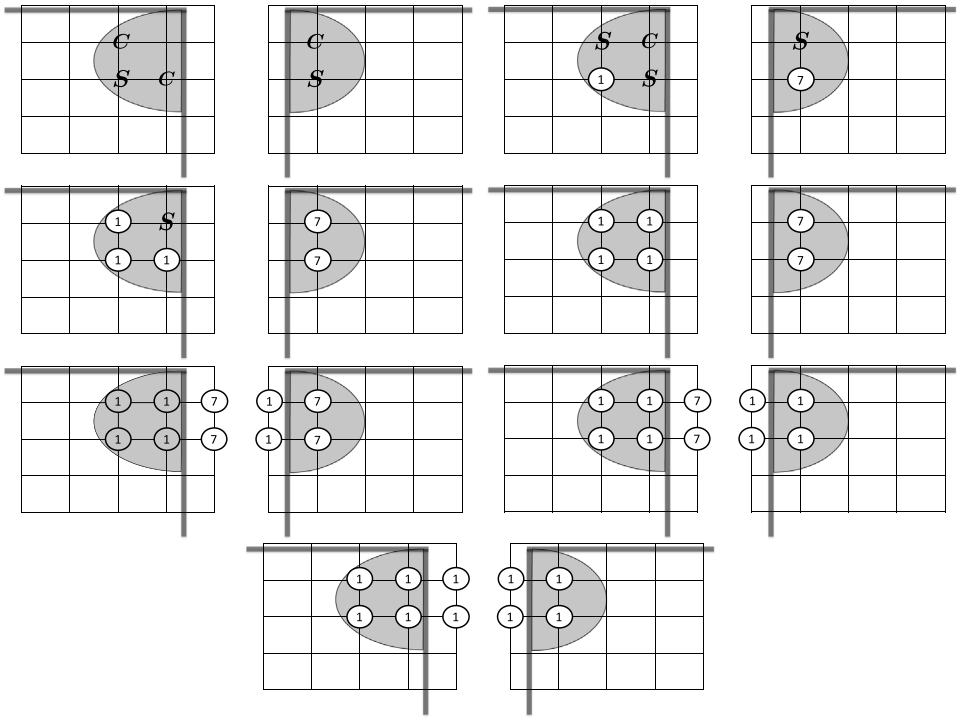


Fig. 3. Visual algorithm of the parallel bubble tagging process. Steps are ordered chronologically from left to right and top to bottom.

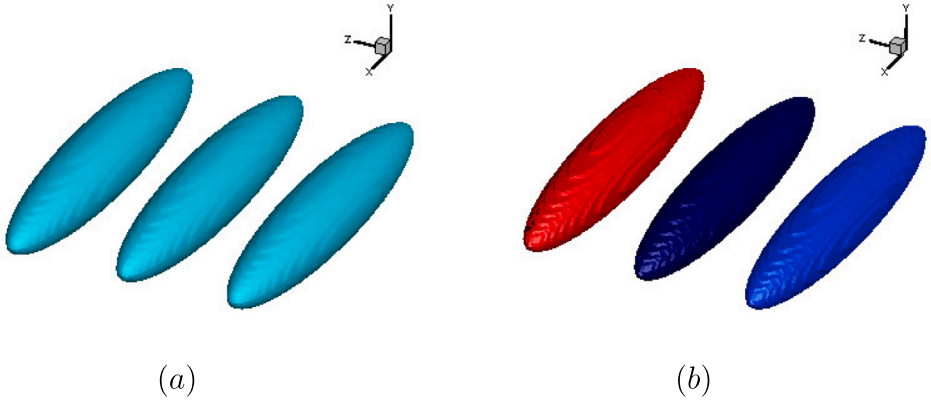


Fig. 4. Three static ellipsoids – (a) iso-surface of C and (b) iso-surface of C colored by tag id. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

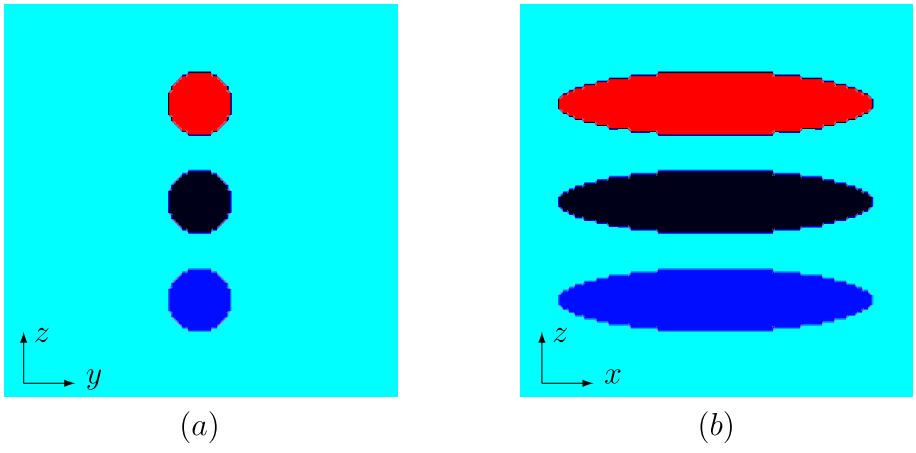


Fig. 5. Contours of tag id as seen from the (a) ZY-midplane and (b) ZX-midplane.

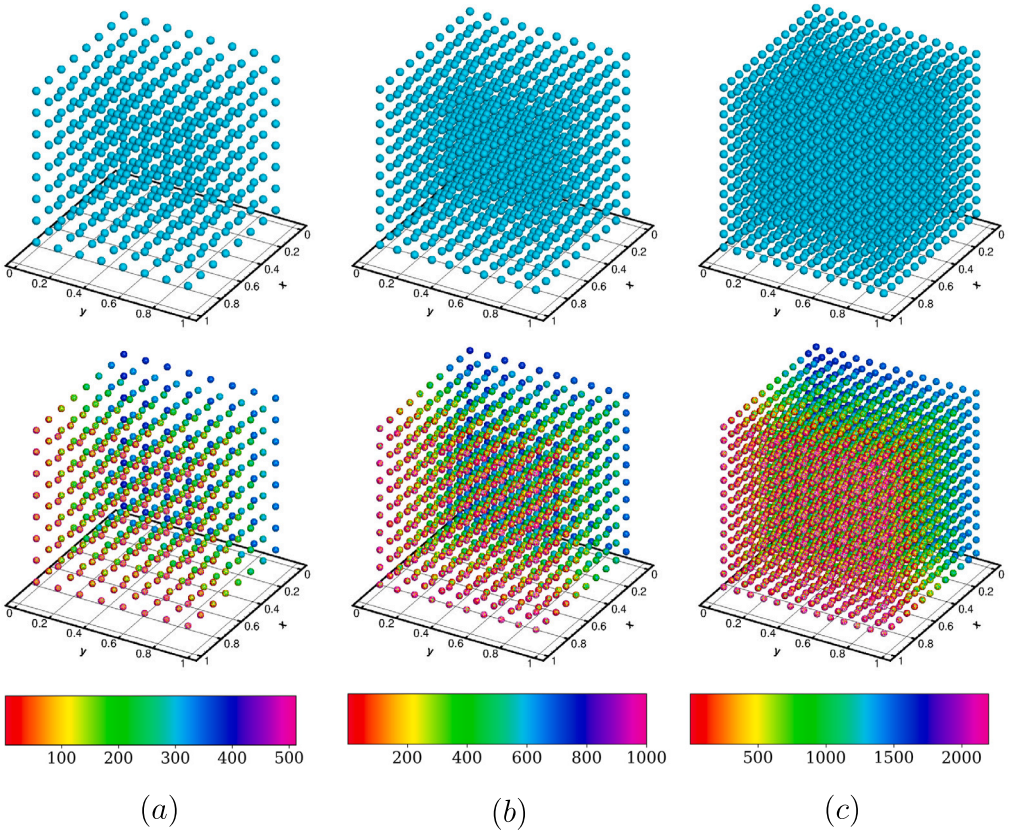


Fig. 6. Iso-surface of C (top row) and iso-surface of C colored by tag id (bottom row) for (a) $N = 8^3$, (b) $N = 10^3$, and (c) $N = 13^3$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the bottom row for all values of N. The colorbars represent bubble id value as different colors with the maximum matching the total number of bubbles — as expected.

3.3.3. Bubble cluster

For the final verification test, a bubble cluster was randomly distributed in a cubic domain resolved by a 256 × 256 × 256 grid. No constraint was applied to prohibit bubble overlap in the initialization process i.e. the bubbles can coincide or intersect partially. Three cases with different bubble numbers were simulated, $N = 500$, 1000, and 2000, respectively.

This configuration more closely mimics what is expected qualitatively in a large-scale two-phase system, where agglomeration and dispersion of the suspended phase is determined by the hydrodynamics rather than an artificial numerical constraint. Fig. 7 shows the color function (top row) and the corresponding tag id (bottom row) for the three bubble numbers simulated.

Note that for fixed domain extents, increasing the total number of bubbles increases the probability of overlap, and therefore the maximum tag id is not necessarily equal to the number of bubbles chosen at initialization. We have conducted a strong scaling study for the full tagging algorithm on a 256 × 256 × 256 grid and $N = 1000$ to demonstrate speedup with respect to the available number of cores. The algorithm exhibits near linear scaling up to 4000 cores beyond which communication overhead starts to dominate (see Fig. 8). For the same grid and domain size, the execution time (plotted in seconds in Fig. 9) was compared for increasing number of bubbles. The highest execution time recorded was 1.1443 s (averaged over 10 time steps) for the most number of bubbles ($N = 2000$). For the same computational domain, the typical time step in our in-house DNS multiphase flow solver consumes 5 s without tagging, therefore, the tagging procedure only adds a fractional contribution to the full simulation time step.

In this section, we have presented and verified the VOF-based component labeling procedure. The algorithm is implemented in parallel to run on multiple CPUs and scales well for large numbers of bubbles. Different shapes are initialized using the color function C; this is followed by labeling the shapes with an integer identifier referred to as tag id.

4. Buoyancy-driven motion of a single bubble

This section explores the rising motion of a single bubble in a quiescent liquid to validate the in-house fluid solver (see Section 2) for buoyancy-driven motion. This is done as a prerequisite for the investigation of a bubbly flow using VOFCCL in what follows.

Bubble shape, the rise velocity U_T, and the drag coefficient C_D are all a complex function of hydrodynamic, viscous, and interfacial forces (Balcázar et al., 2015). As discussed in Section 2.1, the relevant non-dimensional quantities are:

$$Re = \frac{\rho_1 g^{1/2} D^{3/2}}{\mu_1} \qquad Bo = \frac{\rho_1 g D^2}{\sigma} \qquad \eta_\rho = \frac{\rho_1}{\rho_2} \qquad \eta_\mu = \frac{\mu_1}{\mu_2} \quad ,$$

and the Morton number is expressed as $Mo = Bo^3/Re^4$. Four numerical simulations were performed for $Bo = 116$, $\eta_\rho = 100$, and $\eta_\mu = 100$. The Morton number was varied for each case such that $Mo = [848, 266, 41.1, 5.51]$, respectively. As the Mo decreases, inertial effects become more dominant, and the bubble shape deviates away from a sphere.

4.0.1. Computational domain

The buoyancy-driven motion of a single bubble is simulated inside a three-dimensional rectangular domain ($8D \times 8D \times 12D$) resolved by a non-uniform grid equal to 128 × 128 × 360 cells – D is the initial diameter of the bubble. The computational grid is made up of two

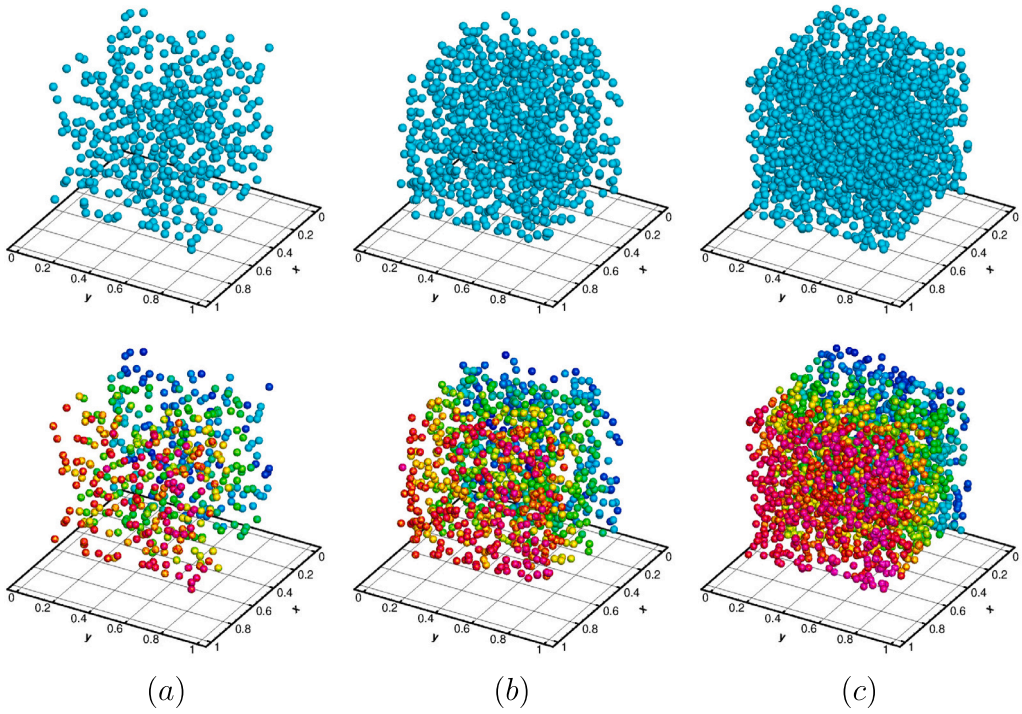


Fig. 7. Iso-surface of C (top row) and iso-surface of C colored by tag *id* (bottom row) for (a) $N = 500$, (b) $N = 1000$, and (c) $N = 2000$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

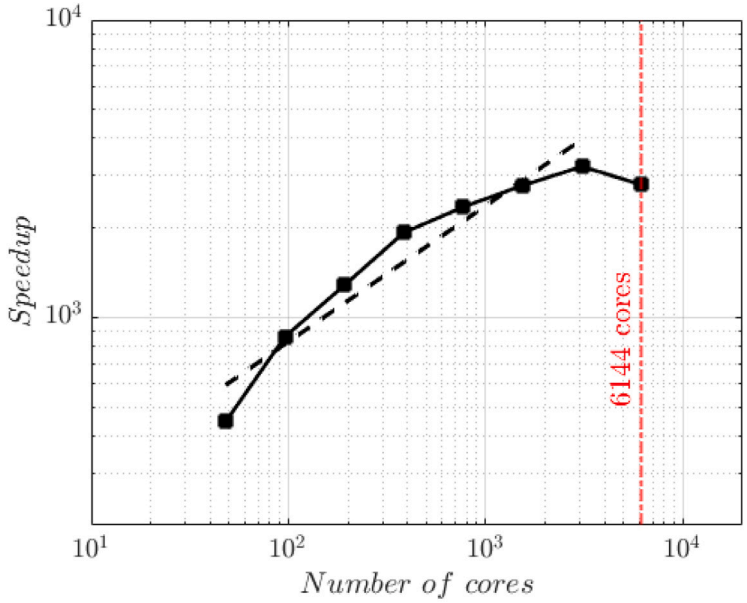


Fig. 8. Strong scaling of full tagging algorithm in a unit cube for $N = 1000$ and a mesh resolution of $256 \times 256 \times 256$.

parts: a central uniform part defined by a square of $L = 3D$, and a stretched part that extends from the edges of the square outwards to the domain boundaries (see Fig. 11). Bubble motion takes place inside the uniform portion of the grid. The choice of domain size and grid resolution is based on the convergence study of Balcázar et al. (2015) such that boundary effects are eliminated, and bubble shape and terminal velocity are not affected by further refinement (≈30 cells per diameter). Periodic boundary conditions were prescribed in all three Cartesian directions with gravity and buoyancy force acting in the $z-$direction, and the bubble is initialized at $(x_c, y_c, z_c) = (4D, 4D, 5D)$. Fig. 12 shows the initial and boundary conditions.

The terminal bubble shape is compared with the experimental results of Bhaga and Weber (1981) and the conservative level-set (CLS) method of Balcázar et al. (2015). Fig. 13 shows a side-by-side comparison of the iso-surface of C from NS-VOF, photographs from Bhaga and Weber (1981), and the iso-surface of ϕ from CLS.

The bubble shapes from our fluid solver compare well with the experimental results and CLS, except for $Mo = 5.51$ where a skirted bubble is visible. If we look at the interface profiles in Fig. 14(d), the skirt develops at a later stage of the rise, and the shape observed in the experiments and CLS more closely resembles the interface contour before the last. Since no information is available about the exact time

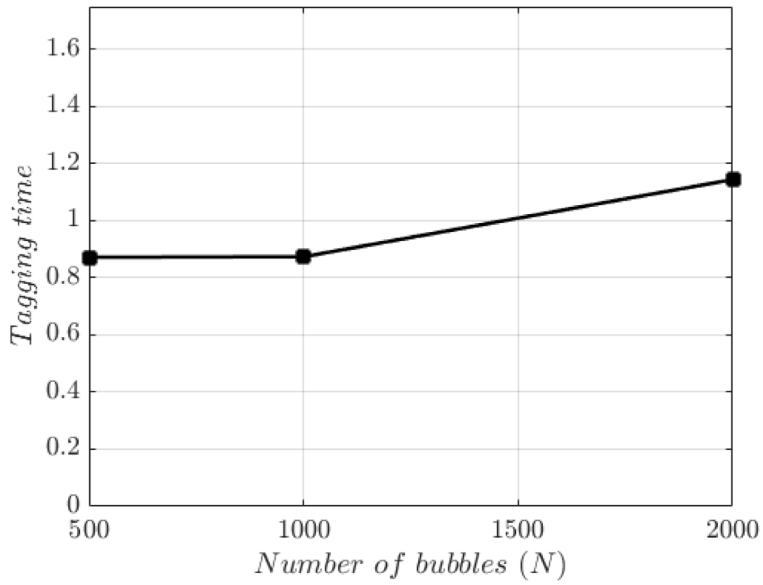


Fig. 9. Execution time of tagging algorithm with respect to number of bubbles.

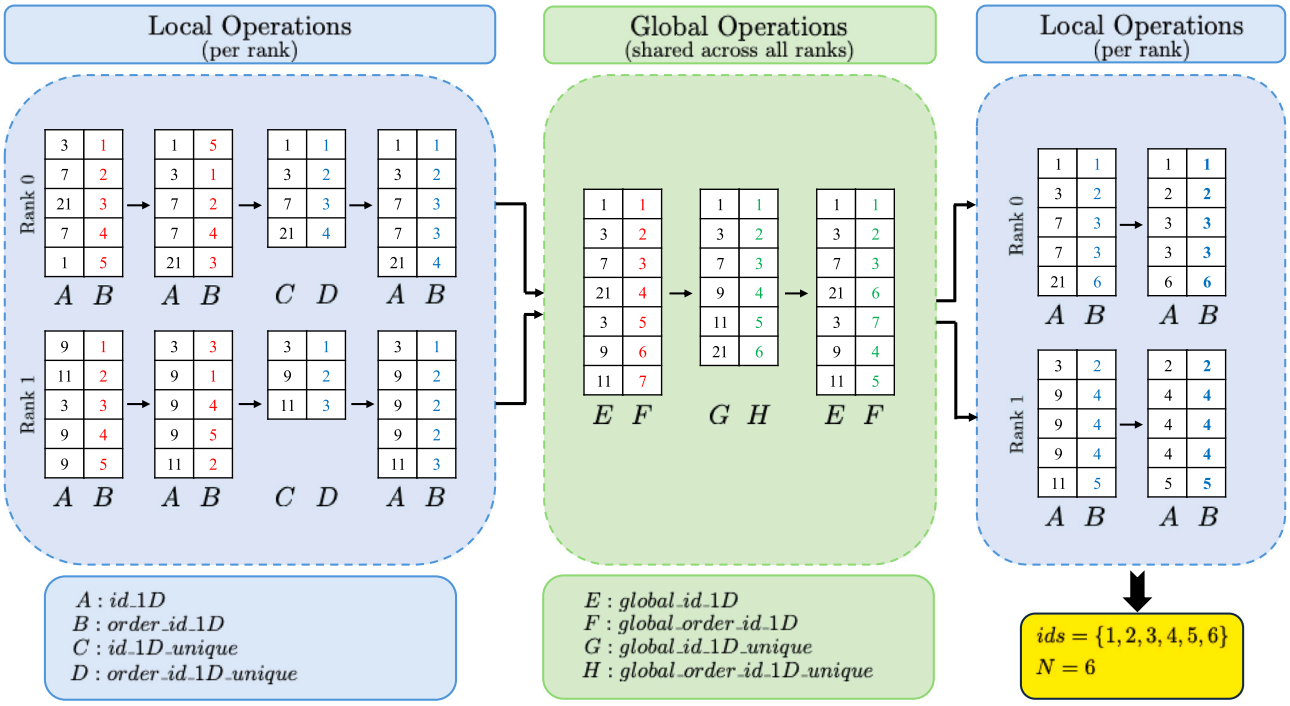


Fig. 10. Visual description of the data structure and procedure of bubble retagging.

the profiles were extracted in the literature, extra care should be practiced when deciding the instant at which the terminal shape is reached. This is especially the case for low Mo where viscous effects are not as strong.

The evolution of bubble Re versus t^* is shown in Fig. 15. The Re profiles for all cases agree well with Balcázar et al. (2015). Table 1 shows a quantitative comparison of the calculated terminal Re with the literature. The relative error in the Reynolds number is computed using $\epsilon_r = |Re_{exp} - Re_{sim}|Re_{exp}^{-1}$ and the values from NS-VOF are 6.96%, 1.28%, 1.73%, and 1.35%, respectively, with the highest error being for $Mo = 848$. For that Morton number, Hua et al. (2008) and Balcázar et al. (2015) report an error of 6.19% and 7.28%, respectively.

Table 1
Comparison of the terminal Re computed using NS-VOF against the experiments of Bhaga and Weber (1981), the FT method of Hua et al. (2008), and the CLS method of Balcázar et al. (2015).

Mo	Re			
	Bhaga and Weber (1981)	Hua et al. (2008)	Balcázar et al. (2015)	NS-VOF
848	2.47	2.317	2.29	2.298
266	3.57	3.621	3.54	3.616
41.1	7.16	7.0	6.94	7.036
5.51	13.3	13.17	12.86	13.12

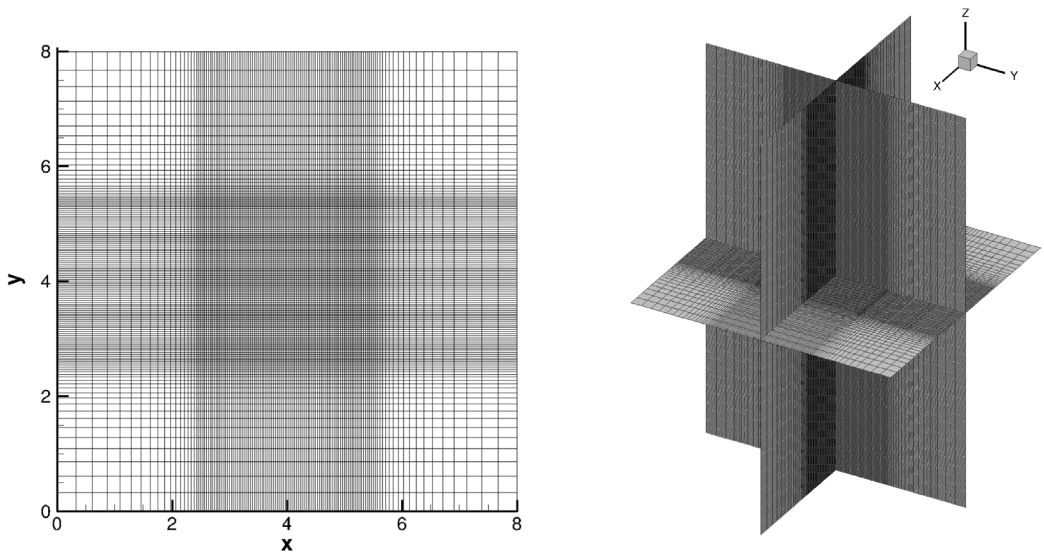


Fig. 11. Details of the computational grid for the single bubble simulations.

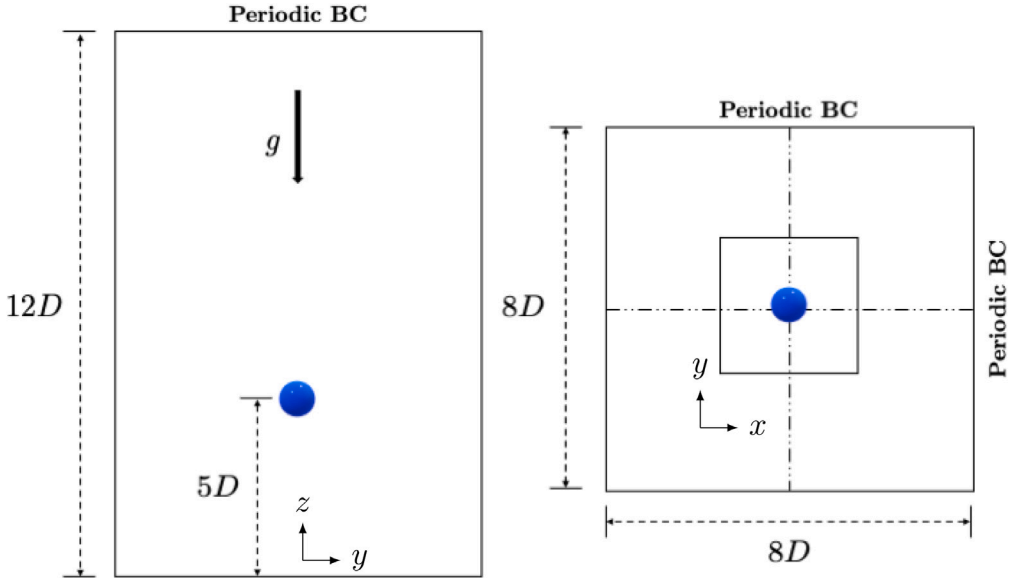


Fig. 12. Schematic of initial and boundary conditions.

The bubble drag coefficient C_D was computed from our simulations via a steady state balance in the vertical direction such that:

$$C_D = \frac{4(\rho_l - \rho_g)|\mathbf{g}|D}{3\rho_l U_T^2} \quad . \tag{21}$$

Bhaga and Weber (1981) proposed an experimental correlation between C_D and Re for fluids with Morton number $Mo > 4 \times 10^{-3}$ where

$$C_D = \left((2.67)^{0.9} + \left(\frac{16}{Re} \right)^{0.9} \right)^{1/0.9} \quad . \tag{22}$$

Additionally, Joseph (2003) derived a theoretical expression for C_D based on viscous potential flow theory such that

$$C_D = 0.445 \left(6 + \frac{32}{Re} \right) \quad . \tag{23}$$

Fig. 16 shows a comparison between C_D computed from NS-VOF and Eqs. (22) and (23). Each data point (solid) represents one Morton number, and the values are shown to agree well with the literature.

5. Buoyancy-driven motion of multiple bubbles

Although many experimental, theoretical, and numerical studies have been dedicated to understanding bubbly flow behavior, comprehensive knowledge is still lacking. When performing DNS of multi-fluid systems, VOF allows the processing of the dispersed phase in the Eulerian frame of reference — by construction. However, as mentioned earlier, many details about the complex flow are lost due to the spatial averaging involved in the computation of quantities that require a summation based on the color function C. This is where VOFCCL can be adopted as a practical tool that can aid in filling some of the gaps in bubbly flow knowledge. The latter is achieved by deriving a Lagrangian framework from the Eulerian counterpart with no alteration to the original numerical method used in simulating the flow.

The goal of this section is to demonstrate the feasibility of using VOFCCL for analyzing bubbly flows by computing dispersed phase quantities that are otherwise unavailable in the Eulerian description. The effects of density ratio η_ρ and initial bubble distribution on the evolution of a bubble swarm are investigated. The values chosen for Mo,

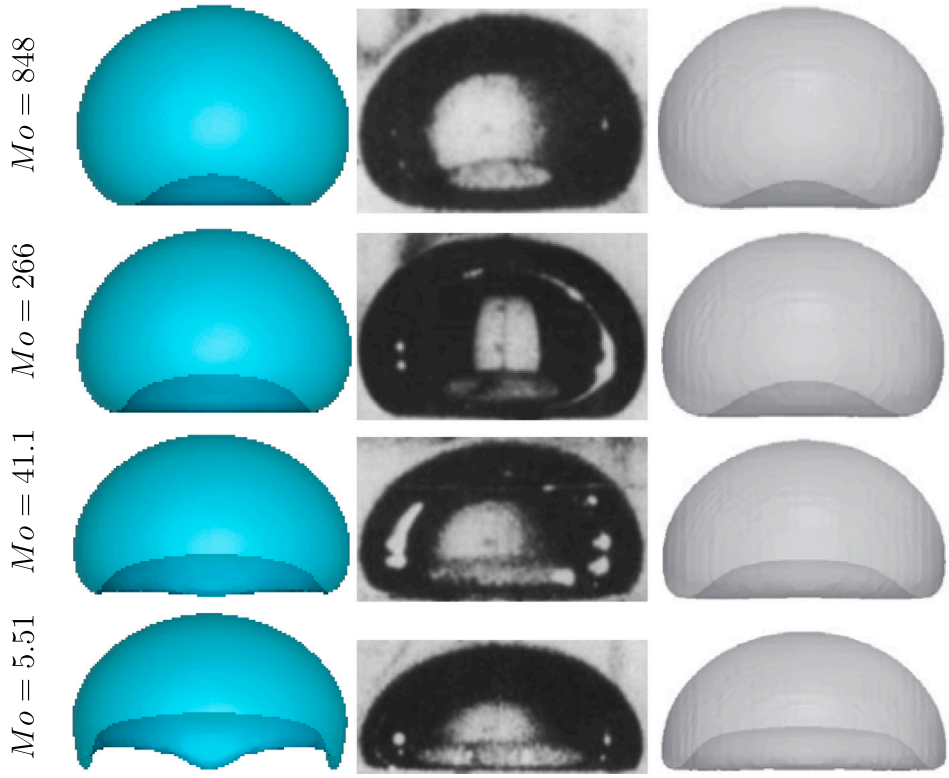


Fig. 13. Comparison of bubble terminal shape captured by NS-VOF (left), experiments of Bhaga and Weber (1981) (center), and the CLS method of Balcázar et al. (2015) (right).

Bo, and η_μ in all the following calculations are 5.9×10^{-4}, 2.5, and 10, respectively. In this parameter space, the bubbles are mostly spherical. The Reynolds number associated with the following simulations can be estimated directly from the Morton and Bond numbers such that:

$$Re_b = \left(\frac{Bo^3}{Mo}\right)^{\frac{1}{4}} \left(\frac{\Delta\rho}{\rho_l}\right)^{\frac{1}{4}} \tag{24}$$

where $\Delta\rho = \rho_l - \rho_g$ (subscripts $l = 1$ and $g = 2$). For the highest density ratio simulated ($\eta_\rho = 100$), Eq. (24) yields $Re_b \approx 13$ which puts the initially quiescent flow strictly in the laminar regime based on the corresponding terminal velocity of the bubbles. This is consistent with the parameter space of the DNS study by Bunner and Tryggvason (2002b) who simulated a homogeneous free array of up to 216 bubbles at a rise Reynolds number of about 36 and $\eta_\rho = 50$, where the carrier phase was treated as laminar throughout the study. For $Re_b < 20$, bubbles do not develop significant wakes (Blanco and Magnaudet, 1995; Mougin and Magnaudet, 2001), thereby removing the primary mechanism by which bubble-induced agitation is transferred to the carrier phase at higher Reynolds numbers.

5.1. Computational domain

The computational domain is a three-dimensional rectangular duct ($6.6D \times 6.6D \times 13.3D$) resolved by a $198 \times 198 \times 399$ uniform grid (≈ 16 million cells). The minimum grid spacing is chosen such that the initial bubble diameter is resolved by 30 cells. The boundary conditions are periodic in the direction of gravity and no-slip (wall) in all other directions. Thirty bubbles of size D were randomly initialized inside the computational domain with the following constraints:

- Minimum center-to-center distance between a bubble pair is $1.25D$.
- Minimum bubble center-to-wall distance is $1.5D$.

These constraints were imposed by generating the bubble locations externally using an in-house program that was developed for this work. The program penalizes center-to-center and center-to-wall distances, and adjusts the bubbles centers dynamically until the above conditions are met globally. The positions are then written out to a file and read into the fluid solver at initialization. This allows the simulations to be repeatable for the same bubble distribution. The overall volume fraction based on the bubble size relative to the computational domain is 2.71% which corresponds to a dilute bubbly flow. At a void fraction of 2.71% and $Re_b \sim O(10)$, the present simulations operate in the laminar bubbly flow regime, and transition to turbulence in the carrier phase is not expected. Fig. 17 shows the details of the computational setup for the bubbly flow simulations.

5.2. Effect of density ratio

Three density ratios are explored 10, 50, and 100, and all other physical parameters are fixed, as mentioned earlier. Since bubbly flow is not a fully developed flow regime, the simulations were stopped at $t^* = 45$ before complete agglomeration occurs. According to Hua and Lou (2007), the density ratio has little effect on the shape and terminal velocity of a single bubble rising in a $\eta_\rho > 50$ medium. This is not necessarily the case for the swarm due to non-linear interactions that occur as the flow evolves. Fig. 18(a) shows the iso-surface of C of the initialized swarm at $t^* = 0$. The corresponding Lagrangian representation through the tag id is shown in Fig. 18(b).

As individual bubbles begin to rise, each bubble evolves its own trailing wake that has the potential to entrain bubbles behind it. This phenomenon has been investigated in the literature both experimentally and numerically (Brereton and Korotney, 1991; Chakraborty et al., 2013; Balcázar et al., 2014). The bottom right of Fig. 19(a) shows the alignment of bubbles labeled 2 and 3 behind bubble 1. As the density ratio is increased, the bubble swarm accelerates further due the direct relationship between the buoyant force and η_ρ. Since we initialize with

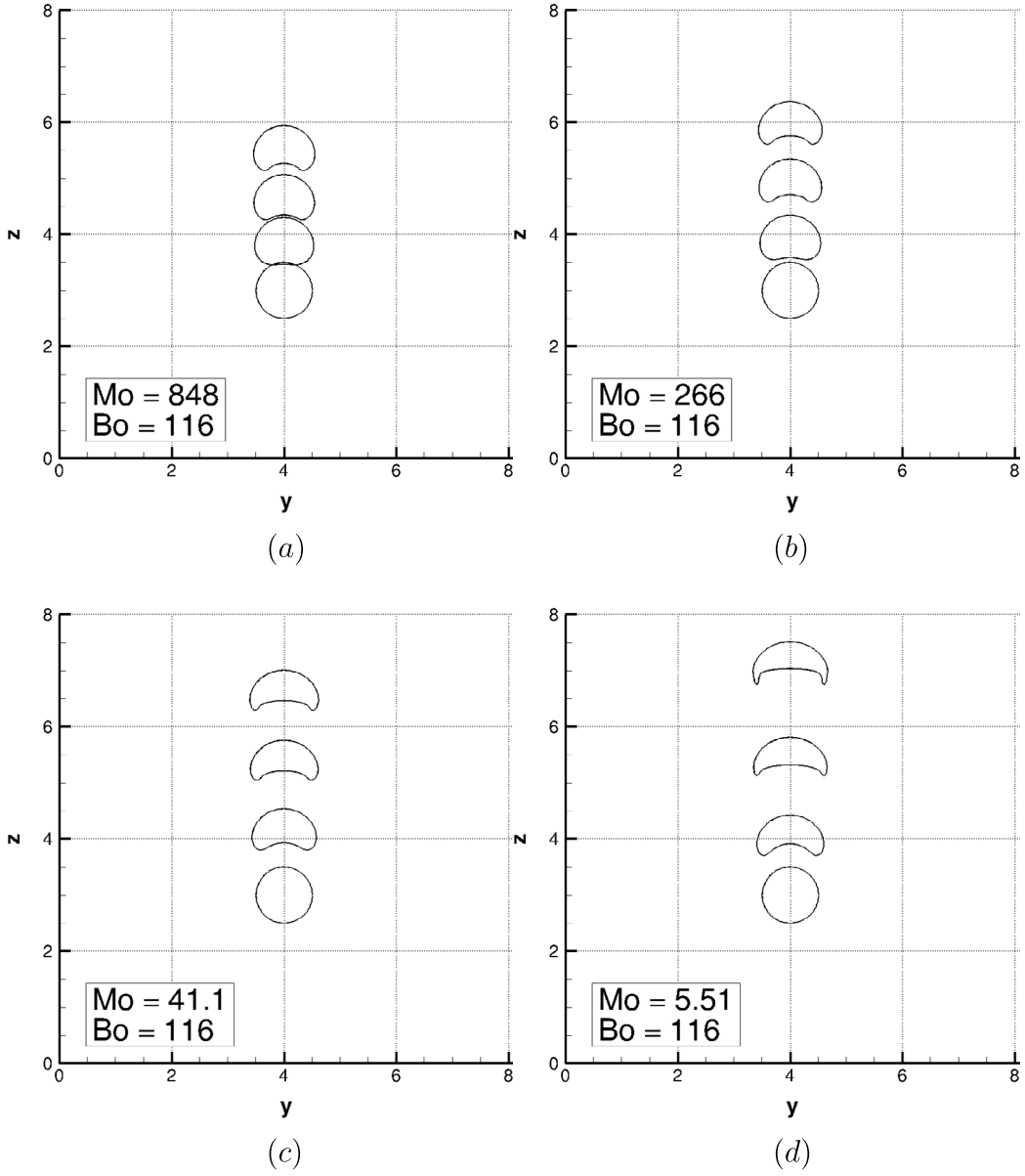


Fig. 14. The evolution of bubble shape for different values of Mo.

the same bubble distribution in every calculation, the slice extracts for the different density ratios shown in Fig. 19 give the impression of an evolving bubbly flow observed at different time instants when, in fact, all snapshots are at $t^* = 45$. At $\eta_\rho = 50$, bubble 2 is entrained by bubble 1 to form a larger bubble. Upon increasing the density ratio to 100, both bubbles 2 and 3 merge with the leading bubble (i.e. bubble 1) at $t^* = 45$.

Fig. 20 gives further insight into the nature of the flow developing around the rising bubbles where the contours of the component of the vorticity vector $\boldsymbol{\omega} \cdot \mathbf{e_x}$, which is normal to the $ZY-$plane, are shown. As a bubble in the swarm moves up, it displaces the liquid in its vicinity creating a rotational flow at the rim of the bubble. For a bubble Reynolds number $Re_b < 110$ and $Mo > 4 \times 10^{-3}$, a laminar toroidal wake pattern is observed behind the trailing surface of the bubble (Bhaga and Weber, 1981). As explained earlier, when the bubbles are conveniently aligned (see Fig. 20) and the rotational flow is strong enough to promote liquid film drainage between the surfaces that is faster than the vertical motion of the bubble pair, coalescence becomes more likely. As a rising bubble flattens and grows, the circulation around it increases and its zone of influence expands. In the cases where rising bubbles are not aligned vertically, ω_x can cause an off-center trailing bubble to tilt upwards towards the wake of the larger bubble rather than merge. In other words, the bubble is sufficiently far to stay locked but not absorbed (see bubble 4 in Fig. 20(c)).

In all three bubbly flow calculations, VOFCCL was used to sample the total number of bubbles N every $\Delta t^* = 0.05$. The maximum value of tag id reflects N, and therefore is stored temporally such that $N = f(t)$. Fig. 21 shows the evolution of with respect to t^* for the different density ratios considered. Up until $t^* = 5.6$, N exhibits identical variations for all η_ρ. Between $t^* = 0$ and $t^* = 13$, only N_{50} and N_{100} overlap with no observable deviation. Beyond $t^* = 13$, the decay profile becomes unique for each case which indicates different coalescence patterns. At the early stages of bubble rise, the bubble distribution is more homogeneous and the distance a bubble needs to travel before "feeling" another bubble is much shorter. During that phase, the liquid is still relatively quiescent and the effect of density ratio on global hydrodynamic variations is not dominant. As the larger bubbles form and gain more inertia, more liquid volume is displaced and consequently the effect of η_ρ becomes more pronounced.

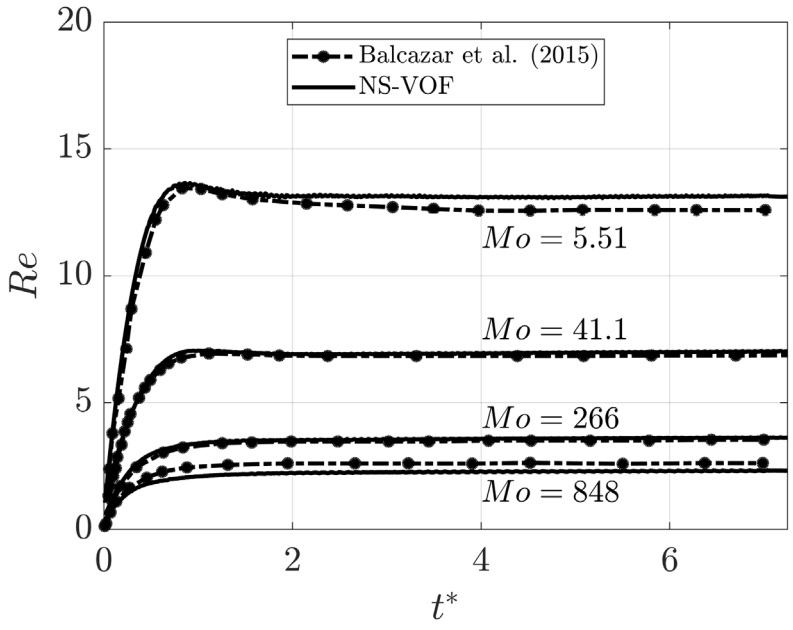


Fig. 15. Effect of Morton number (Mo) on terminal Reynolds number (Re).

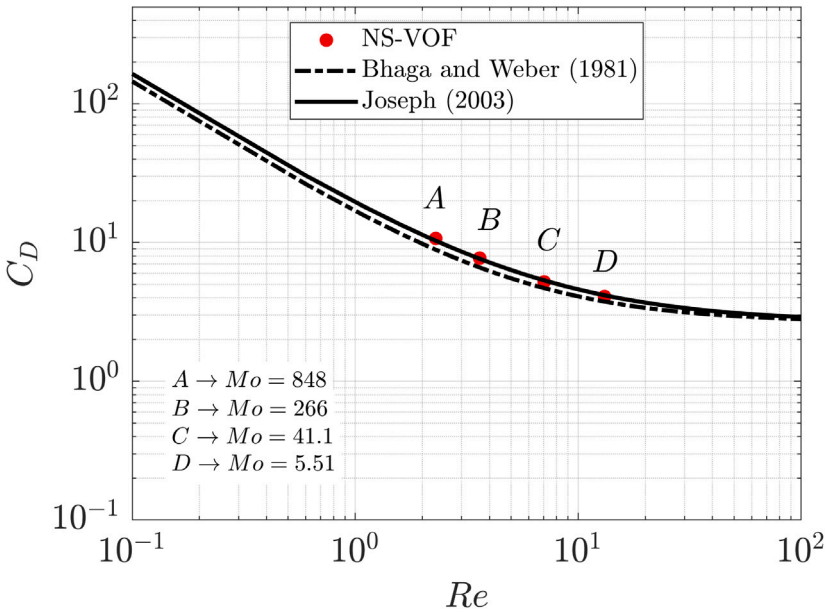


Fig. 16. Comparison of the numerical prediction of C_D versus Re against Eqs. (22) and (23).

VOFCCL was also utilized to compute the standard deviation of the distance between the bubbles as a metric for dispersed phase homogeneity. A small value of σ suggests that most bubbles are similarly spaced from one another i.e their distribution is more homogeneous. It should be noted that while this metric provides a computationally practical and robust measure of bubble distribution in this flow regime, it may become less representative for highly deformed or non-spherical bubbles, where the centroid location does not necessarily reflect the minimum separation between bubble interfaces. In such cases, surface-to-surface distances can provide a more physically meaningful measure of local interactions. Fig. 22 shows the effect of density ratio on the variation of σ_{ij} where subscripts i and j refer to the distance d_{ij} between any two bubbles. Similar to Fig. 21, overlap regions exist until $t^* = 3$ for all η_ρ, and until $t^* = 18$ for σ_{50} and σ_{100}. The standard derivation oscillates for all density ratios with an increasing phase shift from σ_{10} beyond $t^* = 18$. No homogeneous state is reached due to continuous coalescence.

In addition, the average radial distance of the bubble swarm from the domain center was calculated by sampling the distance between each bubble's centroid and $(x_c, y_c) = (L_x/2, L_y/2)$. Although this quantity can indeed be computed using the volume fraction only, it is included as evidence for the ability of VOFCCL to reproduce the generic spatial-averaged quantities as well as additional information that is otherwise unavailable using the color function C alone. Appendix shows a validation case for the use of VOFCCL to compute the bubble radius in Rayleigh–Plesset bubble collapse. Fig. 23 shows the variation of r_c with respect to time for all η_ρ. Up until $t^* = 25$, there is no significant effect of η_ρ on r_c, however beyond that time, the average radial distance of the swarm from the domain center overall increases as the density ratio increases. This means that bubbles migrate closer to the wall.

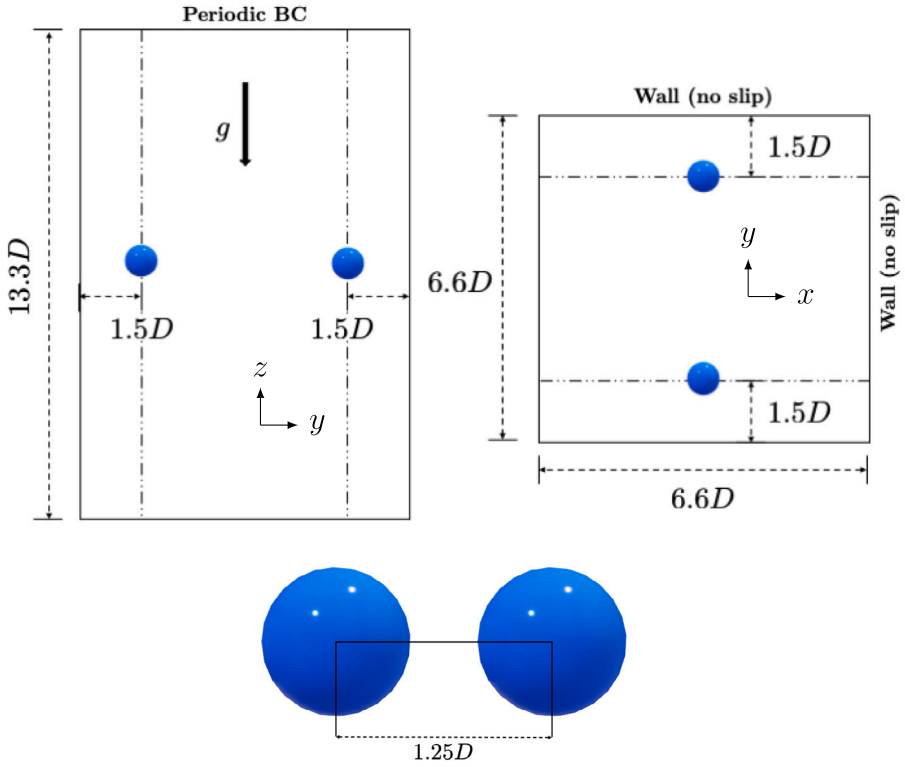


Fig. 17. Schematic of initial and boundary conditions for bubbly flow simulations.

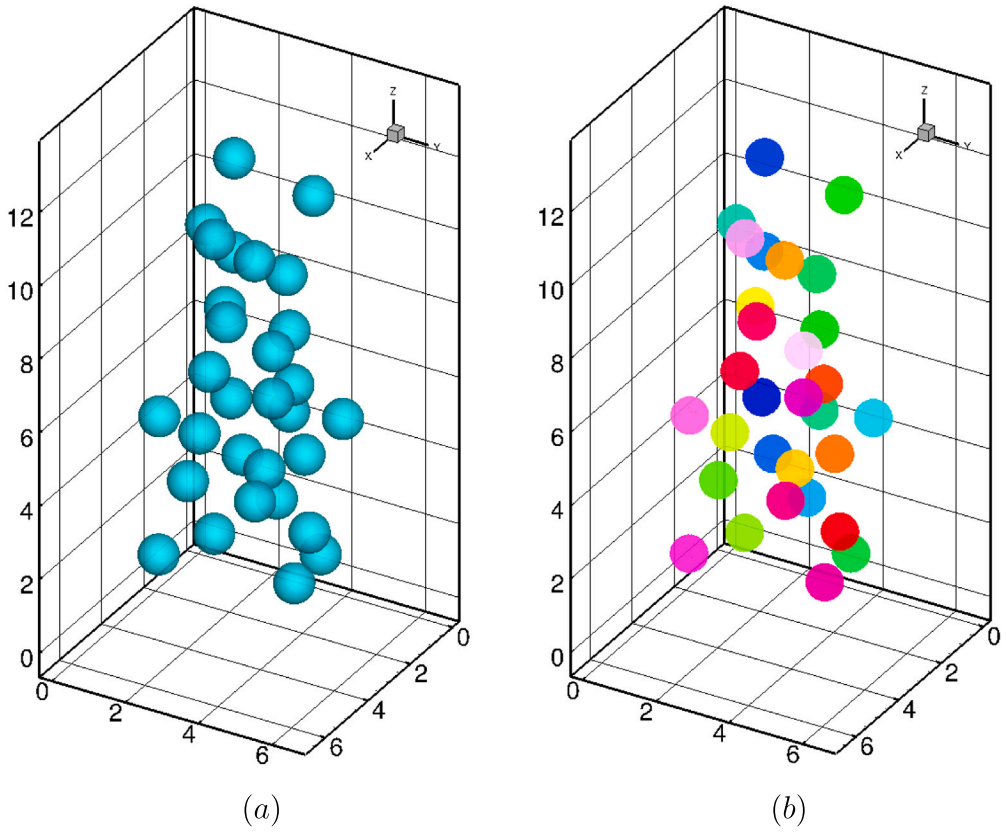


Fig. 18. Bubble swarm at $t^* = 0$, (a) shows the iso-surface of C and (b) shows the iso-surface of C colored by tag id. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

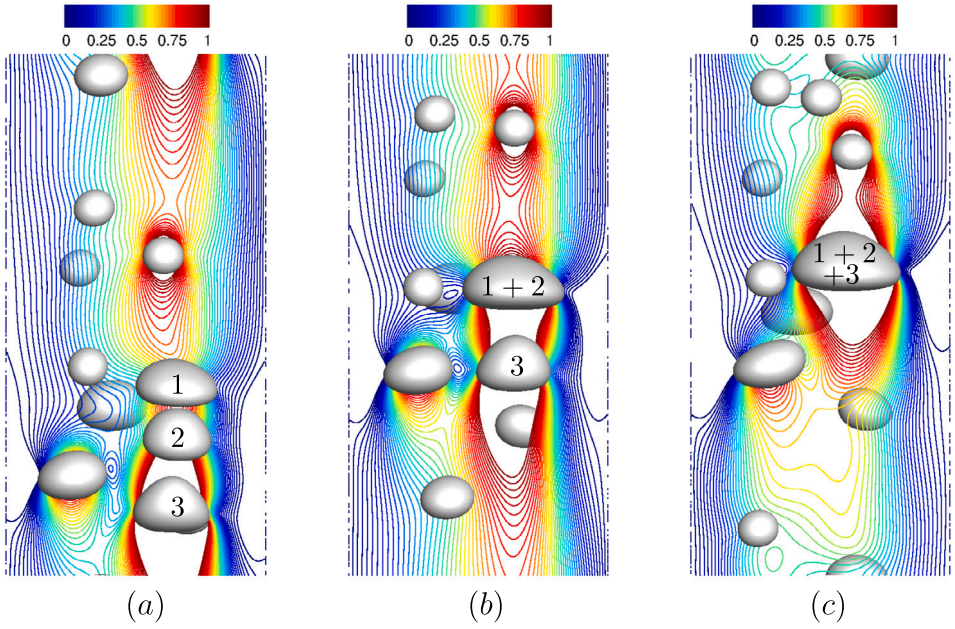


Fig. 19. Contours of vertical velocity w on the $ZY-$plane at $t^* = 45$ for η_ρ equal to (a) 10, (b) 50, and (c) 100.

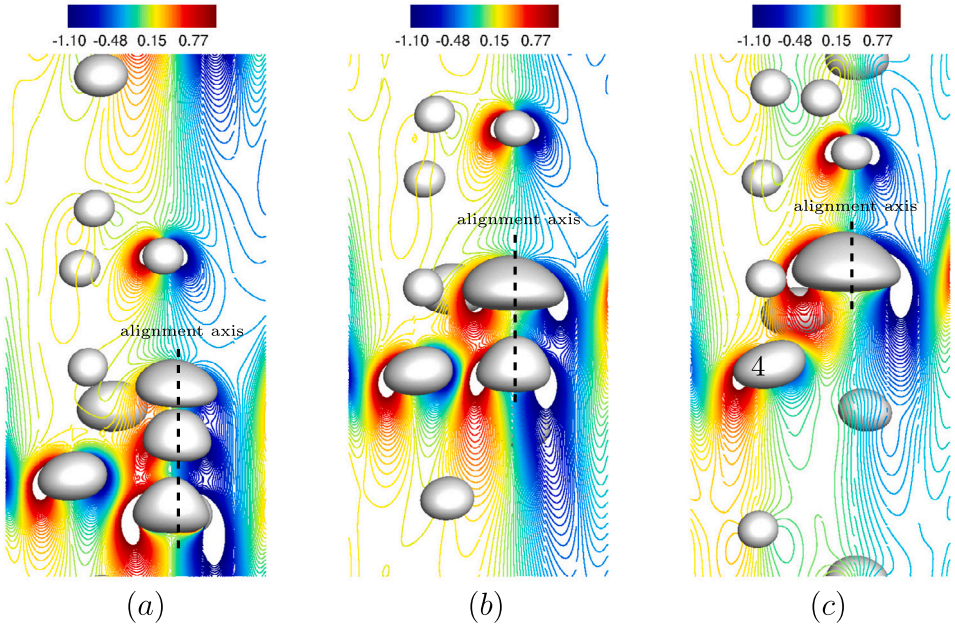


Fig. 20. Contours of the vorticity component $\boldsymbol{\omega} \cdot \mathbf{e}_x$ on the $ZY-$plane at $t^* = 45$ for η_ρ equal to (a) 10, (b) 50, and (c) 100.

To further verify the previous observation, we compute the variation of bubble radial distribution with respect to time. VOFCCL enables this calculation through the sampling of bubble *id*s in radial bins. If we compare $r-$distribution in the order of increasing density, the high bubble density tail labeled in Fig. 24 at $t^* = 25$ gradually faints. This indicates a lower concentration of bubbles near $r \approx 1.25$, which is consistent with the time evolution of r_c shown in Fig. 23.

5.3. Effect of bubble distribution

Six different initial bubble configurations were simulated to investigate the effect of bubble distribution on the evolution of the dilute bubbly flow. Fig. 25 shows the iso-surface of C colored by bubble *id* for all six different distributions. The Bond number Bo, Morton number Mo, η_ρ, η_μ were set to 2.5, 5.9×10^{-4}, 100, and 10, respectively, in all six calculations. Hence, the only influencing parameter is the initial locations of the bubbles inside the rectangular duct. The simulation time, computational domain, and the corresponding boundary conditions are the same as those in Section 5.2.

The bubble number was tracked in all configurations and plotted with respect to time, as shown in Fig. 26. By comparing Figs. 21 and 26, the profiles of N in the latter have a wider spread over multiple distributions, and also start to deviate from one another much earlier ($t^* = 2$).

If $N_i(t)$ in Fig. 26 is ensemble-averaged, a linear fit can be constructed for bubble number decay under the chosen flow parameters

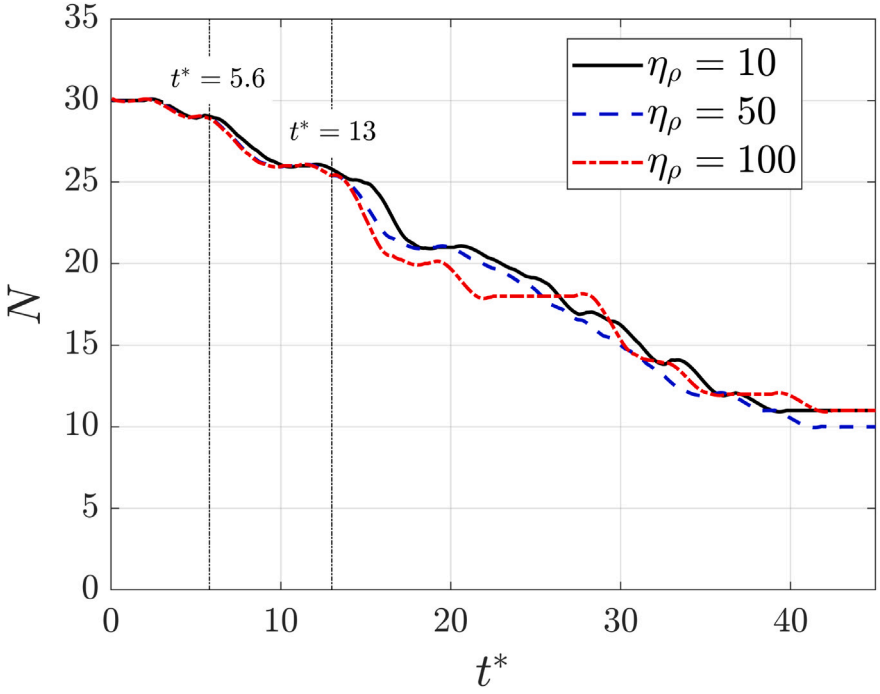


Fig. 21. The evolution of bubble number with time for different values of η_ρ.

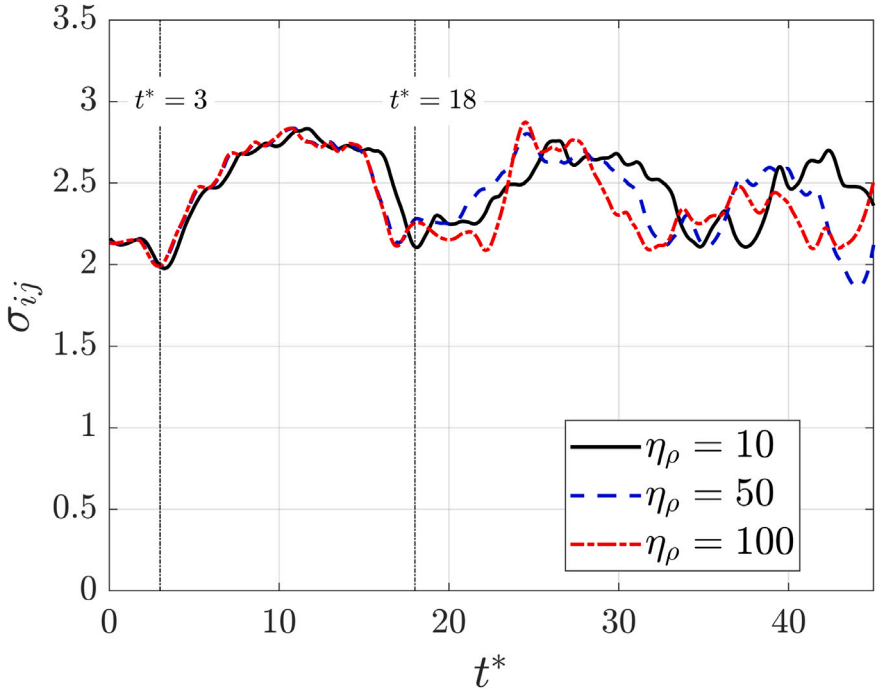


Fig. 22. The time evolution of the standard deviation of bubble-to-bubble distance for different values of η_ρ.

such that $N = -0.4269t + N_0$ with $R^2 = 0.9922 - N_0$ being the initial number of bubbles. Note that this relationship is not conclusive, of course, since the sample size (number of different distributions) is not high enough for a statistically significant result and additional simulations are needed to reach model convergence. Nonetheless, this exercise aims to demonstrate the power of using VOFCCL for deducing possible model relations from such analyses. The parameter σ_{ij} and average radial distance of the swarm (r_c) were also computed. Fig. 27 shows that all distributions exhibit comparable extrema in σ_{ij}, with no significant deviation from the computed mean $\bar{\sigma}$. This shows that initial bubble distribution has no observable effect on swarm homogeneity.

In Fig. 28, r_c is normalized such that $r_c^* = r_c/r_{c0}$ where r_{c0} is the average radial distance at $t^* = 0$. This normalization is necessary to consistently assess the spread in the profiles since each bubble distribution starts with a different r_{c0}, unlike the cases in Section 5.2.

With the exception of distribution 4 in Fig. 28 (solid magenta), $r_{c_i}^*(t)$ increases roughly at the same rate till $t^* = 13$. Beyond that time, the evolution is more stochastic and sharp changes in the swarm location

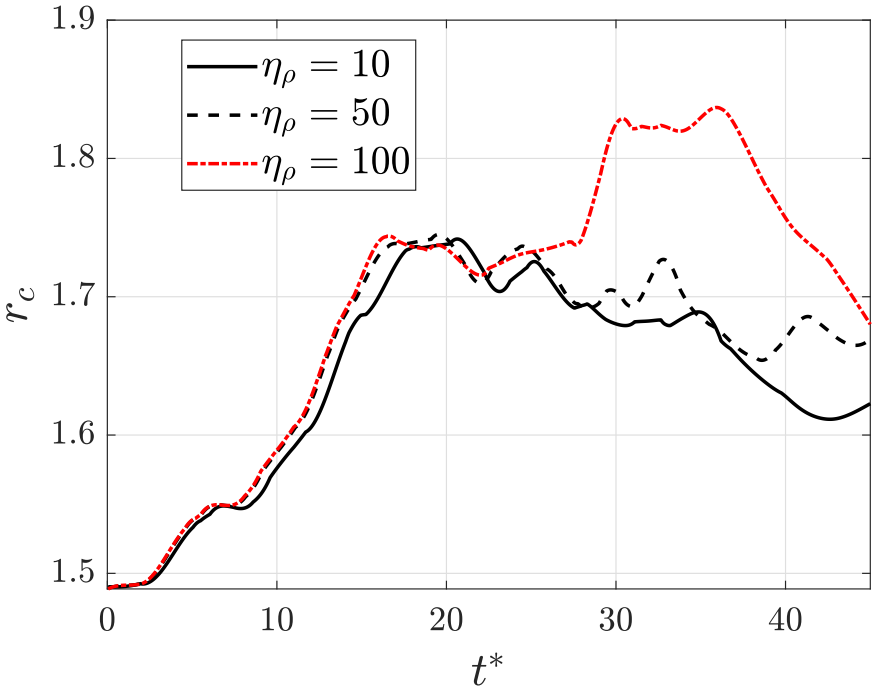


Fig. 23. The time evolution of swarm distance from the centerline for different values of η_ρ.

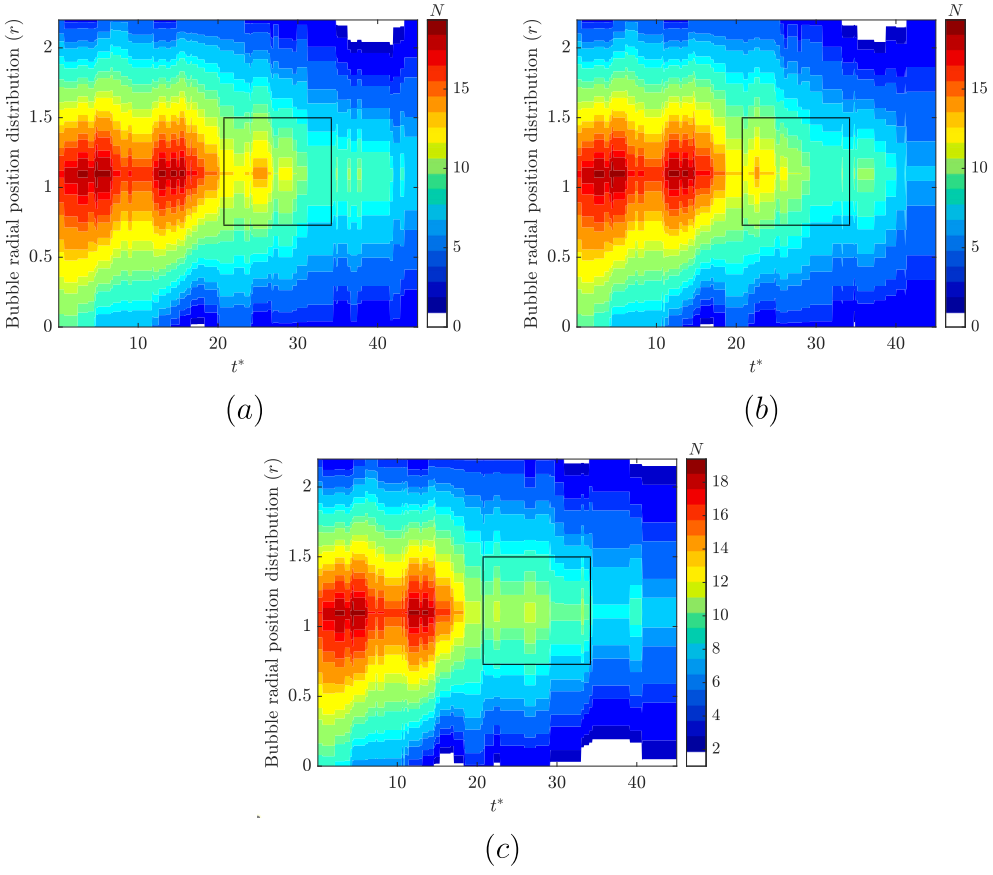


Fig. 24. Radial distribution of bubbles with time for (a) $\eta_\rho = 10$, (b) $\eta_\rho = 50$, and (c) $\eta_\rho = 100$.

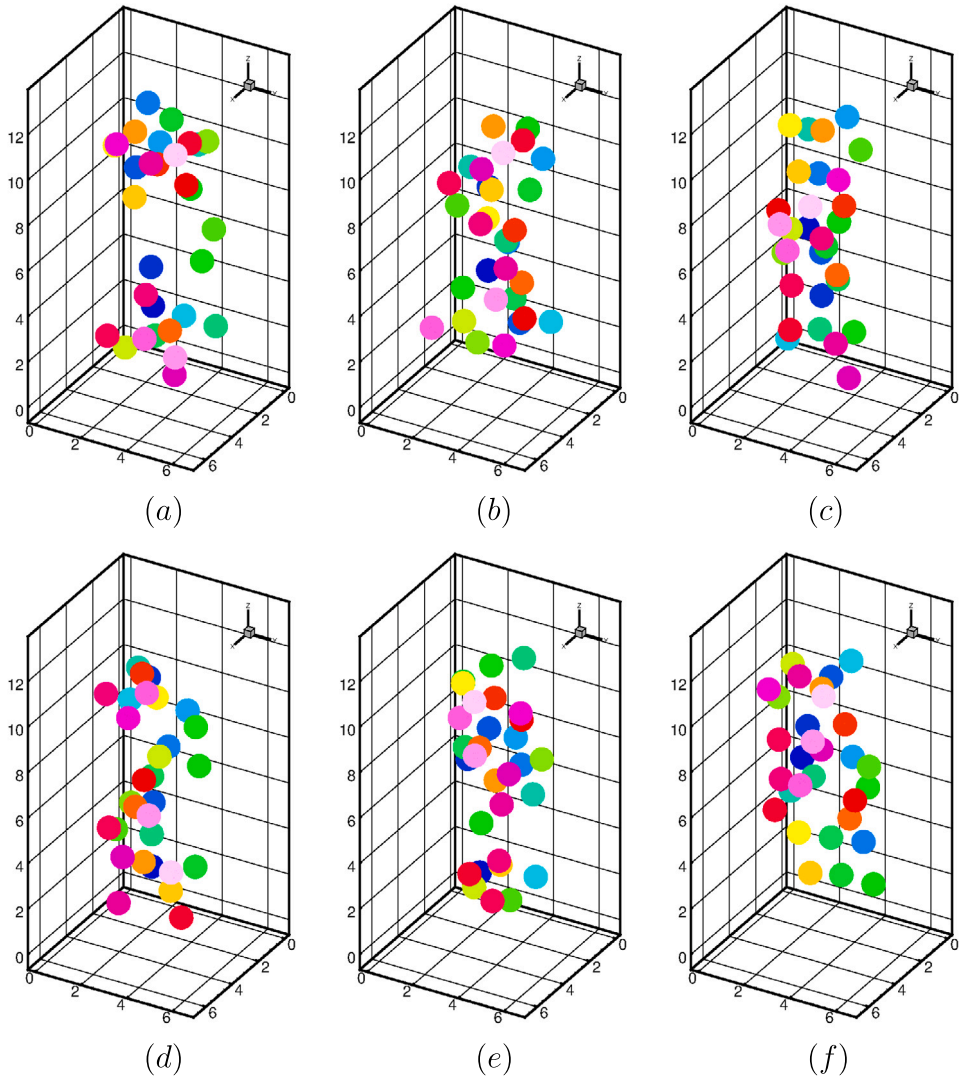


Fig. 25. Iso-surface of C colored by tag id for different initial bubble distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are visible across the different distributions. This naturally means that the initial distribution has a significant influence on the lateral movement of the swarm in the $XY-$plane and its proximity to the centerline, especially after the surrounding liquid is no longer quiescent. Interestingly, since the swarm centroid is the spatial mean of all individual bubble positions, its dynamics can be connected to the classical single-particle dispersion framework of Taylor (1922). At short times ($t^* \lesssim 13$), almost all realizations collapse onto a single trajectory, with r_c^* rising sharply and coherently. This reflects the collective ballistic acceleration of the swarm under buoyancy, during which individual bubble velocities are strongly correlated and the initial bubble distribution plays no role. Beyond $t^* \approx 15$, the realizations diverge considerably. Some configurations sustain centroid growth while others plateau or reverse (e.g. r_3 and r_5), indicating that hydrodynamic interactions begin to dominate swarm dynamics. This distribution-dependence at long times is analogous to the de-correlation phase in Taylor's framework, where $t \sim T_L$ and the velocity autocorrelation has decayed sufficiently that trajectories lose memory of their initial state.

As a closing remark for the framework explored and discussed in this section, experimental investigations of bubbly flows typically rely on conductivity probes, optical probes, or wire-mesh sensors to characterize swarm organization. Although these techniques do not provide complete three-dimensional interface information, they have been used to estimate quantities such as inter-bubble distances, local clustering behavior, and void fraction distribution. VOFCCL operates directly on the color function C and therefore provides access to similar spatial descriptors while retaining full three-dimensional information. We have considered three distinct bubble-resolved statistical quantities here: the instantaneous bubble count N, the standard deviation of bubble-to-bubble distance σ_{ij}, and the swarm centroid location r_c. These quantities respectively characterize population dynamics, spatial dispersion, and large-scale swarm motion. Additional statistics could be extracted from the same tag id field, including nearest-neighbor, size distributions, and bubble trajectory statistics, but such analyses are beyond the scope of the present work.

6. Conclusion

In this work, the VOF-based Connected-Component Labeling (VOFCCL) procedure is demonstrated as a practical diagnostic tool for analyzing bubbly flows in regimes where an Eulerian description alone is insufficient. The method is derived from the formulation of Herrmann (2010), reformulated for the VOF framework, and extended through a retagging procedure that enforces a contiguous bubble indexing from 1 to N. A series of verification test cases, representing distinct dispersed-phase configurations, is presented to assess the performance

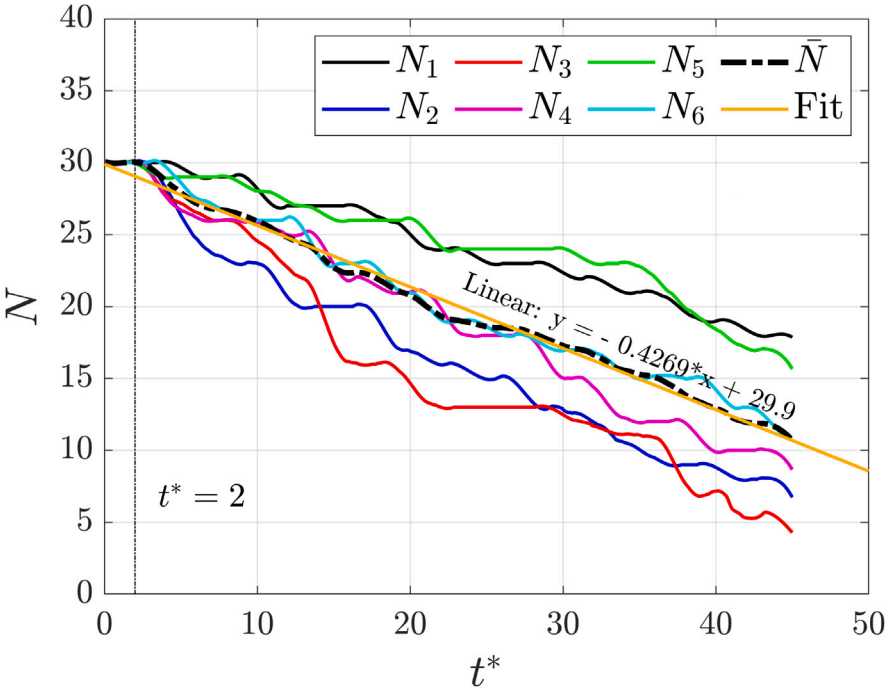


Fig. 26. The evolution of bubble number with time for different initial bubble distributions.

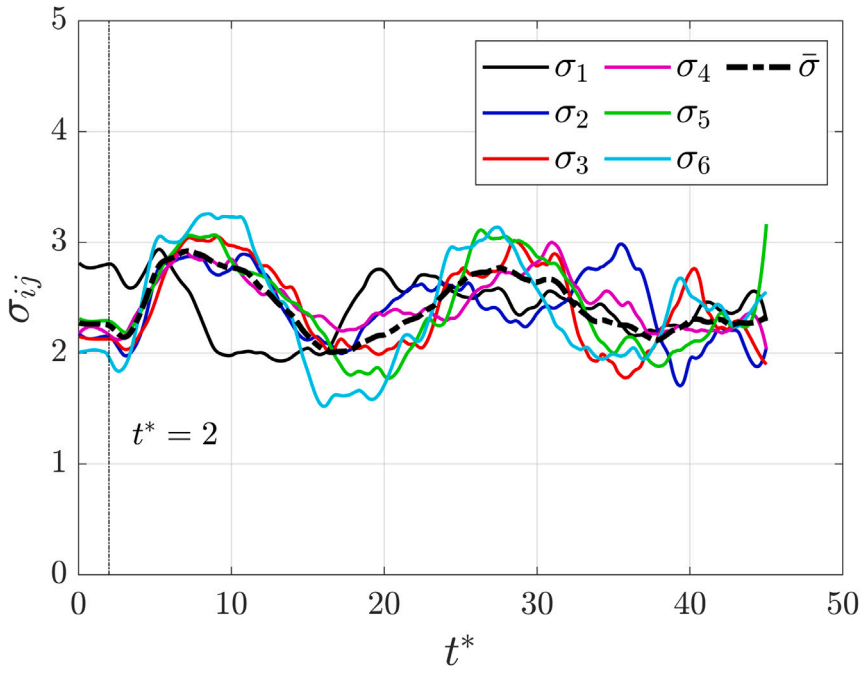


Fig. 27. The time evolution of the standard deviation of bubble-to-bubble distance for different initial bubble distributions.

and consistency of the algorithm. Direct numerical simulations (DNS) of buoyancy-driven bubble swarms in a quiescent liquid are then carried out, with VOFCCL employed in post-processing to extract dispersed-phase statistics. Prior to the swarm simulations, single-bubble rise was validated across a range of Morton numbers (Mo), and the predicted rise velocity and terminal shape were compared with experimental and numerical results from the literature. For the bubbly-flow cases, the nearly spherical regime was considered. Using the bubble tag id field provided by VOFCCL, we evaluated the temporal evolution of the bubble count, the standard deviation of inter-bubble separations, and the motion of the swarm centroid. The simulations examined the influence of the density ratio η_ρ and initial bubble distribution, showing that a hybrid Eulerian–Lagrangian formulation supports the development of model relations and provides a more rigorous basis for ensemble-averaged statistics. Overall, the findings underscore the promise of VOFCCL for enabling more representative dispersed-phase statistics and

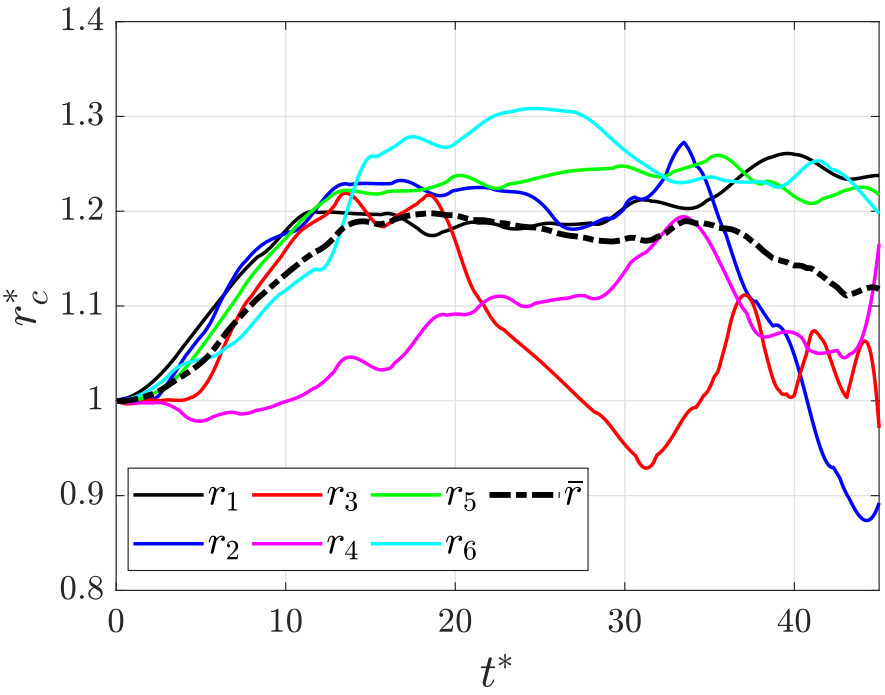


Fig. 28. The time evolution of swarm distance from the centerline for different initial bubble distributions.

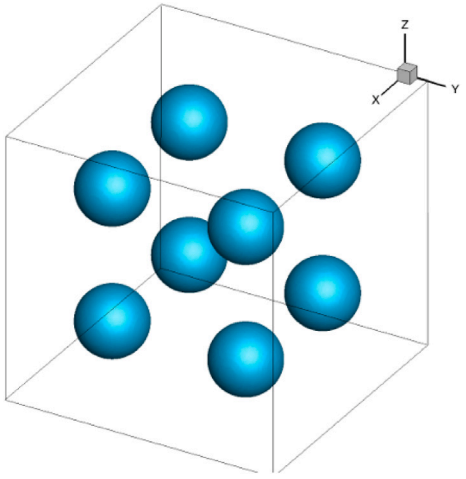


Fig. A.29. Iso-surface of C for eight initialized bubbles.

for supporting future model development in more complex bubbly-flow configurations.

CRediT authorship contribution statement

Ali Fakhreddine: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Krishnan Mahesh:** Writing – review & editing, Supervision, Software, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the United States Office of Naval Research (ONR) under ONR Grant N00014-17-1-2676 with Dr. Ki-Han Kim and Dr. Julie Young as the technical monitors. The computations were made possible through the Minnesota Supercomputing Institute (MSI) at the University of Minnesota.

Appendix. Artificial Rayleigh–Plesset bubble collapse

Three-dimensional Rayleigh–Plesset (RP) bubble collapse is modeled as an advection problem to assess VOFCCL's ability to compute geometric quantities (e.g. bubble radius) with reasonable accuracy. VOFCCL is based on a flood-fill algorithm, therefore the resulting integer field does not guarantee sharp interface representation. Consider the simplified RP equation such that:

$$R\ddot{R} + \frac{3}{2}\dot{R}^2 = \frac{p_{vap} - p_\infty}{\rho_l} \tag{A.1}$$

where $\dot{R}$ represents the rate of change of bubble radius with respect to time i.e. interface speed. With a constant value for p_∞, Eq. (A.1) can be integrated by multiplying through by $2R^2\dot{R}$ and forming the time derivatives, then integrating the equation with the boundary condition $\dot{R}(0) = 0$ (Brennen, 1995). Therefore,

$$\dot{R} = \sqrt{\frac{2(p_{vap} - p_\infty)}{3\rho_l}\left[1 - \left(\frac{R_0}{R}\right)^3\right]} \tag{A.2}$$

where R_0 is the initial radius and R is the time-varying radius to be computed for each bubble using the corresponding id. Note that Eq. (A.2) sets the basis for RP-based cavitation models (Sauer, 2000; Yuan et al., 2001). In the absence of a method to recover the cubic term in Eq. (A.2), this contribution to $\dot{R}$ is generally ignored and important bubble dynamics is lost. This is one area where VOFCCL can also be used to improve the accuracy of existing phase change models.

Since our current numerical method employs a directionally-split VOF method, we can decompose interfacial speed into its Cartesian components using the color function such that:

$$u_R = -\dot{R}\frac{\partial C}{\partial x}\frac{1}{|\nabla C|} \qquad v_R = -\dot{R}\frac{\partial C}{\partial y}\frac{1}{|\nabla C|} \qquad w_R = -\dot{R}\frac{\partial C}{\partial z}\frac{1}{|\nabla C|} \quad . \tag{A.3}$$

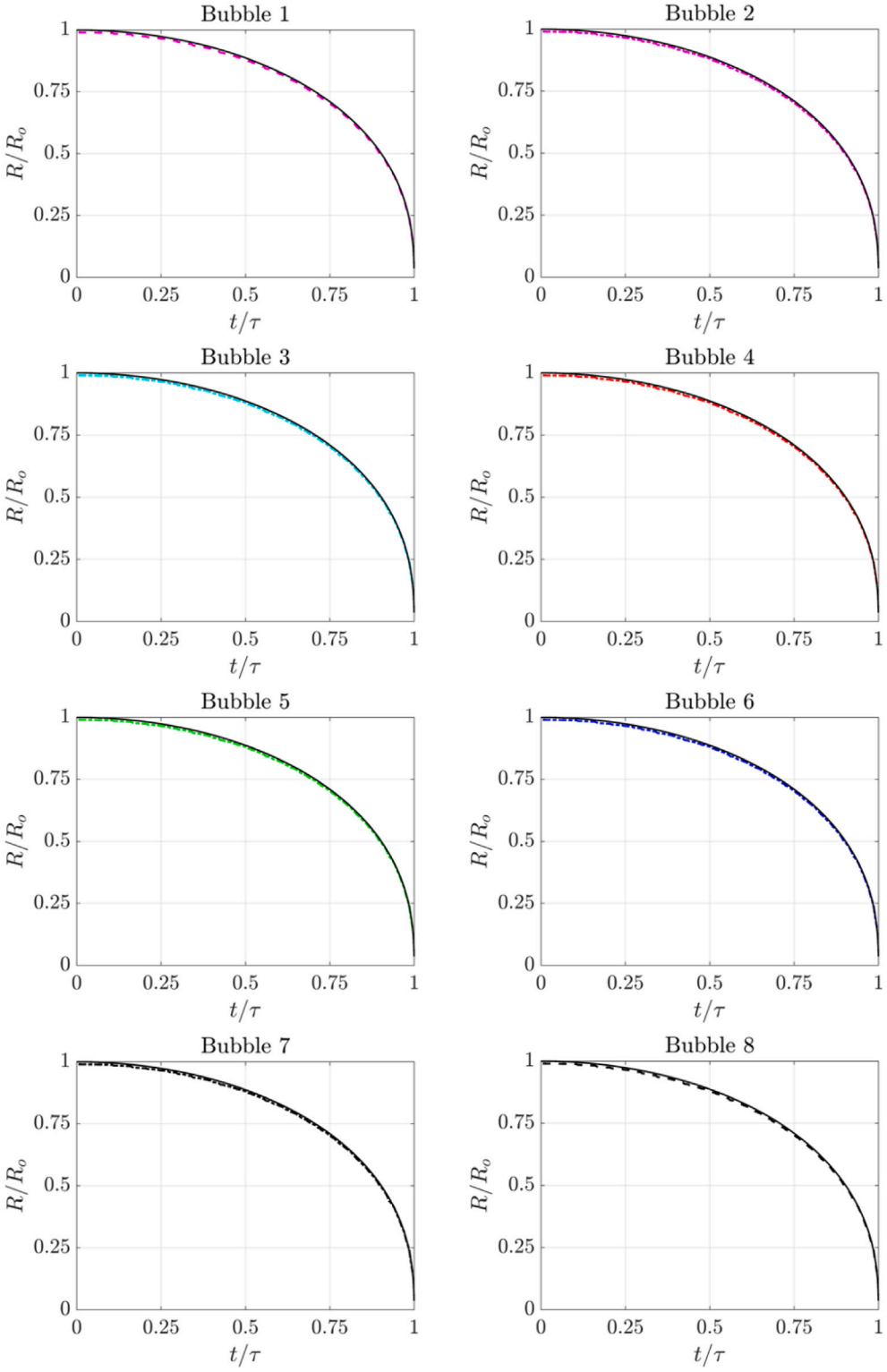


Fig. A.30. Evolution of bubble radius with time compared with the numerical solution of R using 4th−order Runge–Kutta (RK4) integration.

The tag *id* is used to prescribe the interfacial velocity of each of the 8 bubbles shown in Fig. A.29. The bubbles are initialized in a unit cube domain on a 128^3 grid and $R_0 = 0.075$.

As Fig. A.30 shows, the evolution of all eight radii agrees well with the numerical solution of Eq. (A.1) (solid black). This indicates that the computation of R using tag *id* can produce numerically accurate results for models that require geometric information about the dispersed phase.

In recent work, Nathan and Jain (2025) propose a volume-correction approach that incorporates an analytical approximation of

Algorithm 1 Pseudo-code for Bubble Retagging

1: Convert id_{ijk} to id_{1D} (one dimensional list).
2: Create an order list of id_{1D} (i.e. $order\_id_{1D}$) and initialize it with values going from 1 to size of id_{1D}.
3: Sort id_{1D} and $order\_id_{1D}$ such that the id-identifiers are in ascending order in each block.
4: Create a copy of id_{1D} such that $id_{1D}\_copy = id_{1D}$, and uniquify id_{1D}_copy to get id_{1D}_unique.
5: Initialize an order list $order\_id_{1D}\_unique$ similar to Step 2.
6: Create a map between $order\_id_{1D}$ and id_{1D}_unique using $order\_id_{1D}\_unique$ to connect unique id-identifiers to their respective order.
7: Gather new size of id_{1D}_copy from all blocks and store in a global list $size\_id_{1D}\_copy$.
8: Create a displacement list such that, $displacement\_list(1) = 0$
9: **for** $i = 2$ to $nproc$ **do**
10: $displacement\_list(i) = displacement\_list(i-1) + size\_id_{1D}\_copy(i-1)$
11: **end for**
12: Gather id_{1D}_unique from all blocks in a single list $global\_id_{1D}$ and initialize an order list $order\_global\_id_{1D}$ similar to Step 2.
13: Sort $global\_id_{1D}$ and $order\_global\_id_{1D}$ similar to Step 4.
14: Maintain copy of $global\_id_{1D}$ and uniquify the copy $global\_id_{1D}\_copy$ to get $global\_id_{1D}\_unique$.
15: Initialize an order list $order\_global\_id_{1D}\_unique$ similar to Step 2.
16: Create a map between $order\_global\_id_{1D}$ and $global\_id_{1D}\_unique$ using $order\_global\_id_{1D}\_unique$ to connect unique id-identifiers to their respective order.
17: **if** $global\_id_{1D}\_unique(1) = 0$ **then**
18: $order\_global\_id_{1D}(:) = order\_global\_id_{1D}(:) - 1$
19: **end if**
20: **for** $i = 1$ to size of id_{1D} **do**
21: $order\_id_{1D}(i) = order\_global\_id_{1D}(order\_id_{1D}(i) + displacement\_list(my\_rank + 1))$
22: **end for**
23: Set id_{1D} equal to $order\_id_{1D}$
24: Convert id_{1D} back to id_{ijk}
25: Update id-identifiers of ghost nodes in all blocks

the truncated volume in flood-fill methods to improve the accuracy of volume computation. This is especially useful when the interface is not well-resolved and is a potential improvement on the current method.

Data availability

Data will be made available on request.

References

Alméras, E., Mathai, V., Lohse, D., Sun, C., 2017. Experimental investigation of the turbulence induced by a bubble swarm rising within incident turbulence. J. Fluid Mech. 825, 1091–1112.

Balcázar, N., Jofre, L., Lehmkuhl, O., Castro, J., Rigola, J., 2014. A finite-volume/level-set method for simulating two-phase flows on unstructured grids. Int. J. Multiph. Flow 64, 55–72.

Balcázar, N., Lehmkuhl, O., Jofre, L., Oliva, L., 2015. Level-set simulations of buoyancy-driven motion of single and multiple bubbles. Int. J. Heat Fluid Flow 56, 91–107.

Bhaga, D., Weber, M.E., 1981. Bubbles in viscous liquids: shapes, wakes and velocities. J. Fluid Mech. 105, 61–85.

Blanco, A., Magnaudet, J., 1995. The structure of the axisymmetric high-Reynolds number flow around an ellipsoidal bubble of fixed shape. Phys. Fluids 7 (6), 1265–1274.

Brackbill, J.U., Kothe, D.B., Zemach, C., 1992. A continuum method for modeling surface tension. J. Comput. Phys. 100 (2), 335–354.

Brennen, C.E., 1995. Cavitation and Bubble Dynamics. Cambridge University Press, Cambridge, UK.

Brereton, G., Korotney, D., 1991. Coaxial and oblique coalescence of two rising bubbles. Dyn. Bubbles Vortices Near Free Surf. 119, 50–73.

Bröder, D., Sommerfeld, M., 2007. Planar shadow image velocimetry for the analysis of the hydrodynamics in bubbly flows. Meas. Sci. Technol. 18 (8), 2513.

Bunner, B., Tryggvason, G., 2002a. Dynamics of homogeneous bubbly flows Part 1. Rise velocity and microstructure of the bubbles. J. Fluid Mech. 466, 17–52.

Bunner, B., Tryggvason, G., 2002b. Dynamics of homogeneous bubbly flows. Part 2. Velocity fluctuations. J. Fluid Mech. 466, 53–84.

Chakraborty, I., Anderson, T., Bishnoi, H., Mostafa, I., 2013. A coupled level-set and volume-of-fluid method for the buoyant rise of gas bubbles in liquids. Int. J. Heat Mass Transfer 58 (1–2), 240–259.

Debar, R.B., 1974. Fundamentals of the KRAKEN Code. LLNL Tech. Rep. UCIR-760.

Gao, Q., Deane, G.B., Liu, H., Shen, L., 2021. A robust and accurate technique for Lagrangian tracking of bubbles and detecting fragmentation and coalescence. Int. J. Multiph. Flow 135, 103523.

Hendrickson, K., Weymouth, G.D., Yue, K.-P., 2020. Informed component label algorithm for robust identification of connected components with volume-of-fluid method. Comput. & Fluids 197, 104373.

Herrmann, M., 2010. A parallel Eulerian interface tracking/Lagrangian point particle multi-scale coupling procedure. J. Comput. Phys. 229 (3), 745–759.

Hirt, C.W., Nichols, B.D., 1981. Volume of fluid (VOF) method for the dynamics of free boundaries. J. Comput. Phys. 39, 201–225.

Hua, J., Lou, J., 2007. Numerical simulation of bubble rising in viscous liquid. J. Comput. Phys. 222 (2), 769–795.

Hua, J., Stene, J.F., Lin, P., 2008. Numerical simulation of 3D bubbles rising in viscous liquids using a front tracking method. J. Comput. Phys. (ISSN: 0021-9991) 227 (6), 3358–3382.

Jacqmin, D., 1999. Calculation of two-phase Navier–Stokes flows using phase-field modeling. J. Comput. Phys. 155 (1), 96–127.

Joseph, D.D., 2003. Rise velocity of a spherical cap bubble. J. Fluid Mech. 488, 213–223.

Mahesh, K., Constantinescu, G., Moin, P., 2004. A numerical method for large-eddy simulation in complex geometries. J. Comput. Phys. 197 (1), 215–240.

Mendez-Diaz, S., Serrano-Garcia, J.C., Zenit, R., Hernandez-Cordero, J.A., 2013. Power spectral distributions of pseudo-turbulent bubbly flows. Phys. Fluids 25 (4), 043303.

Mougin, G., Magnaudet, J., 2001. Path instability of a rising bubble. Phys. Rev. Lett. 88 (1), 014502.

Mudde, R.F., 2005. Gravity-driven bubbly flows. Annu. Rev. Fluid Mech. 37 (1), 393–423.

Nathan, P.J., Jain, S.S., 2025. Accurate calculation of bubble and droplet properties in diffuse-interface two-phase simulations. J. Comput. Phys. 538, 114190.

Osher, S., Sethian, J.A., 1988. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. J. Comput. Phys. 79, 12–49.

Prosperetti, A., Tryggvason, G., 2009. Computational Methods for Multiphase Flow. Cambridge University Press.

Ravisankar, M., Zenit, R., 2024. Velocity fluctuations for bubbly flows at small Re. J. Fluid Mech. 1001, A34.

Sauer, J., 2000. Instationär Kavitierende Strömungen - Ein neues Modell, Basierend auf Front Capturing (VoF) und Blasendynamik (Ph.D. thesis). Universität Karlsruhe (TH), http://dx.doi.org/10.5445/IR/3122000, Fak. f. Maschinenbau, Diss. v. 18.7.2000.

Taylor, G.I., 1922. Diffusion by continuous movements. Proc. Lond. Math. Soc. 20 (1), 196–212.

Thiesset, F., Ménard, T., Dumouchel, C., 2021. Space-scale-time dynamics of liquid–gas shear flow. J. Fluid Mech. 912, A35.

Tryggvason, G., Scardovelli, R., Zaleski, S., 2011. Direct Numerical Simulations of Gas–Liquid Multiphase Flows. Cambridge University Press, ISBN: 9781139496704.

Youngs, D.L., 1984. An interface tracking method for a 3D eulerian hydrodynamics code. At. Weapons Res. Establ. (AWRE) Tech. Rep. 44 (92), 35.

Yuan, W., Sauer, J., Schnerr, G.H., 2001. Modelling and computation of unsteady cavitation flows in injection nozzles. J. Mech. Ind. 2, 383–394.

Yuan, W., Zhang, M., Khoo, B.C., Phan-Thien, N., 2020. Dynamics and deformation of a three-dimensional bubble rising in viscoelastic fluids. J. Non-Newton. Fluid Mech. 285, 104408.

Yuan, W., Zhang, M., Khoo, B.C., Phan-Thien, N., 2021. Hydrodynamic interaction and coalescence of two inline bubbles rising in a viscoelastic liquid. Phys. Fluids 33, 083102.

Zhou, X., Duan, L., Sun, J., Liu, D., Yuan, W., Chen, F., 2026. Role of non-newtonian rheology on gas-liquid mass transfer and bubble dynamics in T-junction microchannels. Chem. Eng. Sci. 326, 123525.