

# A momentum-conserving, consistent, Volume-of-Fluid method for incompressible flow on staggered grids

D. Fuster<sup>a</sup>, T. Arrufat<sup>a</sup>, M. Crialesi-Esposito<sup>b</sup>, Y. Ling<sup>f</sup>, L. Malan<sup>a,c</sup>, S. Pal<sup>a</sup>, R. Scardovelli<sup>d</sup>, G. Tryggvason<sup>e</sup>, S. Zaleski<sup>a</sup>

<sup>a</sup>*Sorbonne Université et CNRS,*

*Institut Jean Le Rond d'Alembert, UMR 7190, Paris, France*

<sup>b</sup>*CMT-Motores Térmicos, Universitat Politècnica de València, Camino de Vera, s/n, Edificio 6D, Valencia, Spain*

<sup>c</sup>*InCFD, Dept. of Mechanical Engineering, University of Cape Town, South Africa*

<sup>d</sup>*DIN - Lab. di Montecuccolino, Università di Bologna, I-40136 Bologna, Italy*

<sup>e</sup>*Mechanical Engineering, Johns Hopkins University, Baltimore, USA*

<sup>f</sup>*Dept. of Mechanical Engineering, Baylor University, Waco, TX, USA*

---

## Abstract

The computation of flows with large density contrasts is notoriously difficult. To alleviate the difficulty we consider a consistent mass and momentum-conserving discretization of the Navier-Stokes equation. Incompressible flow with capillary forces is modelled and the discretization is performed on a staggered grid of Marker and Cell type. The Volume-of-Fluid method is used to track the interface and a Height-Function method is used to compute surface tension. The advection of the volume fraction is performed using either the Lagrangian-Explicit / CIAM (Calcul d'Interface Affine par Morceaux) method or the Weymouth and Yue (WY) Eulerian-Implicit method. The WY method conserves fluid mass to machine accuracy provided incompressibility is satisfied which leads to a method that is both momentum and mass-conserving. To improve the stability of these methods momentum fluxes are advected in a manner “consistent” with the volume-fraction fluxes, that is a discontinuity of the momentum is advected at the same speed as a discontinuity of the density. To find the density on the staggered cells on which the velocity is centered, an auxiliary reconstruction of the density is performed. The method is tested for a droplet without surface tension in uniform flow, for a droplet suddenly accelerated in a carrying gas at rest at very large density ratio without viscosity or surface tension, for the Kelvin-Helmholtz instability, for a falling raindrop and for an atomizing flow in air-water conditions.

**Keywords:** Multiphase Flows, Navier-Stokes Equations, Volume of Fluid, Surface Tension, Large Density Contrast

---

## 1. Introduction

Multiphase flows abound in nature, but their stable and accurate computation remains elusive in many cases. As a case in point, many numerical methods used for two-phase incompressible flow are strongly unstable for large density contrasts and

large Reynolds numbers. Experience with such simulations shows that the presence of surface tension is an aggravating factor. The large density contrasts that are of interest are air/water or gas/liquid-metal, with  $\rho_l/\rho_g$  of the order of  $10^3$  or  $10^4$ . The large density contrasts are a difficulty whether one deals with any of the three major interface advection methods, Level-Set, Volume-of-Fluid (VOF) or Front-Tracking, or with combinations such as CLSVOF. (The term density contrast is preferable to density ratio since it encompasses ratios both much larger than one and much smaller than one.)

Several methods have been used to alleviate the high-density-contrast difficulties. It has been observed by several authors that making the momentum-advection method conservative improves the situation. For incompressible flow, momentum-conserving methods have been initially proposed by [? ], and by several other authors since [? ? ? ? ]. These methods have been shown to improve the stability of the numerical results in various situations. In particular, liquid-gas flows with very contrasted densities, as for example in the process of atomization, cause serious problems that are resolved by using momentum-conserving methods. In that case another oft-suggested solution is to increase the number of equations from the standard four equations to five, six or seven equations, by introducing new field variables in each phase. The addition of one more density  $\rho_i$ , momentum  $\rho_i \mathbf{u}_i$  or energy variable  $\rho_i e_i$  increases the number of equations. The authors of ref. [? ] used seven equations, those of [? ] used five equations and six equations were used in [? ]. The last three references also use a momentum-conserving formulation. Several authors, including some of those cited above, have argued that the difficulty may come from gas velocities of order  $u_g$  being mixed with liquid densities of order  $\rho_l$ . Both the  $\rho_l$  and  $u_g$  scales are large and the appearance of an unphysical  $\rho_l u_g^2$  dynamic pressure scale could create numerical pressure fluctuations of the same order and unphysical pressure spikes, as the one nicely illustrated in [? ] in the front part of a suddenly accelerated small droplet. One way of avoiding this unphysical mixture of liquid and gas quantities is to extrapolate liquid and gas pressure and velocity in the “other” phase, as in the ghost fluid method. This extrapolation strategy was used successfully in [? ].

It may be argued that a way to avoid this numerical diffusion of liquid and gas quantities is to advect the volume fraction and the conserved quantities that depend on it (density, momentum and energy) in a consistent manner. In incompressible flow in which we specialise in this paper, it means that the volume fraction and momentum or velocity must be advected in the same way. This is equivalent to request that the discontinuity of the Heaviside function  $H$ , marking the phase transition, should be advected at the same speed as the discontinuity in momentum. This can be expressed by the following consistency requirement: if momentum is initially exactly proportional to volume fraction, it should remain so after advection. We call such a method VOF-consistent.

To satisfy this requirement the idea is to solve the advection equation for momentum with the same numerical scheme that is used for the VOF color function. This consistency property minimizes the unphysical transfer of momentum from one phase to another due to the differences in the numerical schemes used. The consistency is especially important when dealing with fluids with a large density contrast where a small numerical momentum transfer from the dense phase to the light phase results in large numerical errors in the velocity field which in turn creates numerical instabilities.

In this work, we present a modification of the classical momentum-preserving scheme proposed by [?] for the case of a staggered grid and VOF method. We also focus more on the consistency requirement between VOF and momentum advection, and only partially conserve momentum in the scheme. This will illustrate mostly the benefit of using a consistent advection scheme.

The paper is organized as follow: the second section deals with the continuum mechanics formulation for incompressible flow and sharp interfaces. Section 3 describes our numerical method, starting with an overview of already-known methods for spatial discretization, time-stepping, and VOF advection. We continue with the new momentum advection-VOF-consistent method. Section 4 is devoted to tests of the method, followed by a conclusion.

Among the authors, Gretar Tryggvason, Ruben Scardovelli, Yue Stanley Ling and Stephane Zaleski have been involved in the construction of the base of the ParisSimulator VOF and Front-Tracking code that was used to implement and test the ideas in this paper. ParisSimulator is itself based on a Front-Tracking code developed by Gretar Tryggvason, Sadegh Dabiri and Jiacai Lu. Daniel Fuster was involved in the development of the Momentum Conserving method with consistent VOF advection, with help from Tomas Arrufat, Leon Malan and Yue Stanley Ling. The falling raindrop testing and the corresponding figures were done by Tomas Arrufat for the CIAM method and Sagar Pal for the WY method. The atomisation testing was done by Marco Cicali-Esposito.

## 2. Navier–Stokes equations with interfaces

We model flows with sharp interfaces defined implicitly by a characteristic function  $H(\mathbf{x}, t)$  defined such that fluid 1 corresponds to  $H = 1$  and fluid 2 to  $H = 0$ . The viscosities  $\mu$  and densities  $\rho$  are calculated as an average

$$\mu = \mu_1 H + \mu_2 (1 - H), \quad \rho = \rho_1 H + \rho_2 (1 - H). \quad (1)$$

There is no phase change so the interface, almost always a smooth differentiable surface  $S$ , advances at the speed of the flow, that is  $V_S = \mathbf{u} \cdot \mathbf{n}$  where  $\mathbf{u}$  is the local fluid velocity and  $\mathbf{n}$  a unit normal vector perpendicular to the interface. Equivalently the interface motion can be expressed in weak form

$$\partial_t H + \mathbf{u} \cdot \nabla H = 0, \quad (2)$$

which expresses the fact that the singularity of  $H$ , located on  $S$ , moves at velocity  $V_S = \mathbf{u} \cdot \mathbf{n}$ . For incompressible flows, which we will consider in what follows, we have

$$\nabla \cdot \mathbf{u} = 0. \quad (3)$$

The Navier–Stokes equations for incompressible, Newtonian flow with surface tension may conveniently be written in operator form

$$\partial_t (\rho \mathbf{u}) = \mathcal{L}(\rho, \mathbf{u}) - \nabla p \quad (4)$$

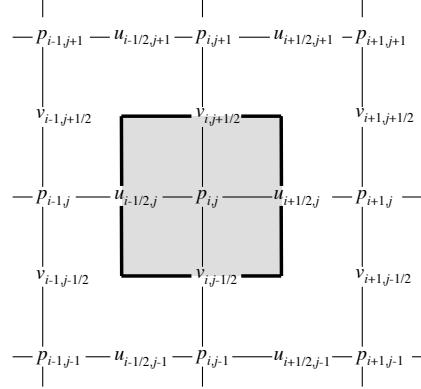


Figure 1: Representation of the staggered spatial discretisation. The pressure  $p$  is assumed to be known at the center of the control volume outlined by a thick solid line. The horizontal velocity component  $u_1 = u$  is stored in the middle of the left and right edges of this control volume and the vertical velocity component  $u_2 = v$  in the middle of the top and bottom edges.

where  $\mathcal{L} = \mathcal{L}_{\text{conv}} + \mathcal{L}_{\text{diff}} + \mathcal{L}_{\text{cap}} + \mathcal{L}_{\text{ext}}$  so that the operator  $\mathcal{L}$  is the sum of advective, diffusive, capillary force and external force terms. The first two terms are

$$\mathcal{L}_{\text{conv}} = -\nabla \cdot (\rho \mathbf{u} \mathbf{u}), \quad \mathcal{L}_{\text{diff}} = \nabla \cdot \mathbf{D}, \quad (5)$$

where  $\mathbf{D}$  is a stress tensor whose expression for incompressible flow is

$$\mathbf{D} = \mu \left[ \nabla \mathbf{u} + (\nabla \mathbf{u})^T \right], \quad (6)$$

where  $\mu$  is computed from  $H$  using (1). The capillary term is

$$\mathcal{L}_{\text{cap}} = \sigma \kappa \delta_S \mathbf{n}, \quad \kappa = 1/R_1 + 1/R_2, \quad (7)$$

where  $\sigma$  is the surface tension,  $\mathbf{n}$  is the unit normal perpendicular to the interface,  $\kappa$  is the sum of the principal curvatures and  $\delta_S$  is a Dirac distribution concentrated on the interface. We assume a constant surface tension value  $\sigma$ . Finally  $\mathcal{L}_{\text{ext}}$  represents external forces such as gravity.

### 3. Method

#### 3.1. Spatial discretization

We assume a regular cubic or square grid. This can be easily generalized to rectangular or cuboid grids, and with some efforts to quadtree and octree grids. We also use staggered velocity and pressure grids. This makes our method more complex than it would be on a collocated grid.

The staggered grid is represented in Figure 1. In a staggered grid discretization of the advection equation the control volumes of the velocity components  $u_1$  and  $u_2$  are shifted with respect to the control volume surrounding the pressure  $p$ . The use

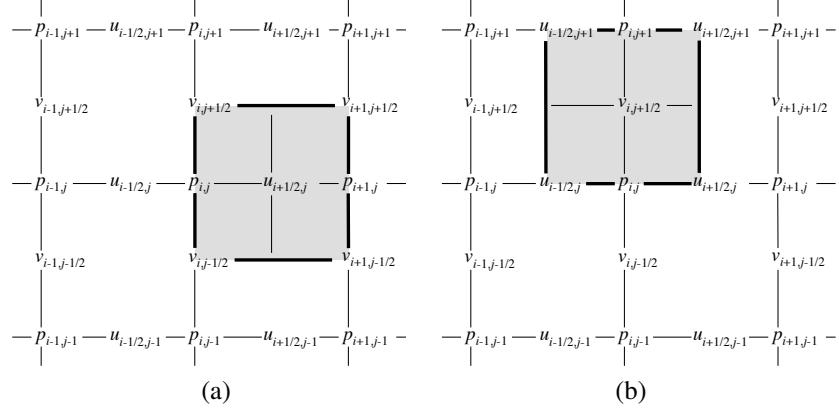


Figure 2: The control volumes for the  $u_1 = u$  and the  $u_2 = v$  velocity components are displaced half a grid cell to the right (horizontal velocities) and to the top (vertical velocities). Here the indices show the location of the stored quantities. Thus, half indices indicate those variables stored at the edges of the pressure control volumes.

of staggered control volumes has the advantage of suppressing neutral modes often observed in collocated methods but leads to more complex discretizations (see [?] for a more detailed discussion.) The control volumes for  $u_1$  and  $u_2$  are shown on Figure 2. This type of staggered representation is easily generalized to three dimensions. In what follows we shall use these control volumes for the velocity or momentum components.

Using the staggered grid leads to a compact expression for the continuity equation (3)

$$\frac{u_{1;i+1/2,j,k} - u_{1;i-1/2,j,k}}{\Delta x} + \frac{u_{2;i,j+1/2,k} - u_{2;i,j-1/2,k}}{\Delta y} + \frac{u_{3;i,j,k+1/2} - u_{3;i,j,k-1/2}}{\Delta z} = 0, \quad (8)$$

In what follows, we shall use the notation  $f = m\pm$ , with the integer index  $m = 1, 2, 3$ , to note the face of any control volume located in the positive or negative Cartesian direction  $m$ , and  $\mathbf{n}_f$  for the normal vector of face  $f$  pointing outwards of the control volume. On a cubic grid the spatial step is  $\Delta x = \Delta y = \Delta z = h$  so the continuity equation becomes

$$\nabla^h \cdot \mathbf{u} = \sum_{m=1}^3 (u_{m+} + u_{m-})/h = 0, \quad (9)$$

where  $u_f = u_{m\pm} = \mathbf{u} \cdot \mathbf{n}_f$  is the velocity normal to face  $f$ . The discretization of the interface location is performed using a VOF method. VOF methods typically attempt to solve approximately equation (2) which involves the Heaviside function  $H$ , whose integral in the cell  $\Omega$  indexed by  $i, j, k$  defines the volume fraction  $C_{i,j,k}$  from the relation

$$h^3 C_{i,j,k} = \int_{\Omega} H d\mathbf{x}. \quad (10)$$

$C_{i,j,k}$  represents the fraction of the cell labelled by  $i, j, k$  filled with fluid 1, taken to be the reference fluid. It is worth noting that in the staggered grid setup, the control volume for  $p$  is also the control volume for other scalar quantities, such as  $\rho$  and  $H$ .

### 3.2. Time Marching

The volume fraction field is updated as

$$C^{n+1} = C^n + \mathcal{L}_{VOF}(C^n, \mathbf{u}^n \tau/h), \quad (11)$$

where  $\mathcal{L}_{VOF}$  represents the operator that updates the Volume of Fluid data given the velocity field. Once volume fraction is updated, the velocity field is updated in a couple of steps. A projection method is first used, in which a provisional velocity field  $\mathbf{u}^*$  is computed

$$\rho^{n+1} \mathbf{u}^* = \rho^n \mathbf{u}^n + \tau \mathcal{L}_{conv}^h(\rho^n, \mathbf{u}^n) + \tau \left[ \mathcal{L}_{diff}^h(\mu^n, \mathbf{u}^n) + \mathcal{L}_{cap}^h(C^{n+1}) + \mathcal{L}_{ext}^h(C^{n+1}) \right] \quad (12)$$

It goes without saying that the above operators depend on the discretization steps  $\tau$  and  $h$  as well as the fluid parameters. The discussion of the  $\mathcal{L}_{conv}^h$  operator is the main point of this paper. In the second step, the projection step, the pressure gradient force is added to yield the velocity at the new time step

$$\mathbf{u}^{n+1} = \mathbf{u}^* - \frac{\tau}{\rho^{n+1}} \nabla^h p. \quad (13)$$

The pressure is determined by the requirement that the velocity at the end of the time step must be divergence free

$$\nabla^h \cdot \mathbf{u}^{n+1} = 0, \quad (14)$$

which leads to a Poisson-like equation for the pressure

$$\nabla^h \cdot \frac{\tau}{\rho^{n+1}} \nabla^h p = \nabla^h \cdot \mathbf{u}^*. \quad (15)$$

### 3.3. Volume-Of-Fluid

We follow the usual VOF method and will only detail the necessary steps to illustrate the momentum advection method based on the VOF method in the next section. Moreover we will use in this section a rescaling of the space and time variables so that the cell size is 1, and the time step is also 1. All velocities are then rescaled to  $u' = u\tau/h$ . Because of this space rescaling and in these new units,  $C_{i,j,k}$  is also the measure of the volume of reference fluid in cell  $i, j, k$ .

#### 3.3.1. Volume initialisation

Before any VOF interface tracking is performed, the field of  $C_{i,j,k}$  values must be initialized. We use the VoFI library described in [? ] and [? ]. This allows a high-accuracy numerical integration of the measure of the fluid volumes.

#### 3.3.2. Normal vector determination

The VOF method proceeds in two steps, reconstruction and advection. In the reconstruction step, one attempts to obtain the “best” possible representation of the interface using Volume of Fluid data  $C_{i,j,k}$ . The best representation is usually the one with the smallest reconstruction error for objects such as circles or ellipses, but issues of continuity and robustness at low resolution may also be considered. In the work reported here we use a standard method described in [? ]. In each cell cut by the interface one first determines the interface normal vector  $\mathbf{n}$ , then one solves the problem of finding a plane perpendicular to  $\mathbf{n}$  under which one finds exactly the volume  $C_{i,j,k}$ . Normal vector determination is performed using the MYC method described in [? ].

### 3.3.3. Plane constant determination

Once the interface normal vector  $\mathbf{n}$  is determined, a new, colinear normal vector noted  $\mathbf{m}$  and having unit  $L_1$  norm is deduced from  $\mathbf{n}$ , so that  $|m_x| + |m_y| + |m_z| = 1$ . Considering the volume  $V = C_{i,j,k}$  in cell  $i, j, k$  the plane constant  $\alpha$  is defined so that the plane

$$\mathbf{m} \cdot \mathbf{x} = \alpha \quad (16)$$

cuts exactly a volume  $V$  in the cell. Typically  $\alpha$  is determined by the resolution of a cubic equation as described in [? ].

### 3.3.4. General split-direction advection

Once the reconstruction has been performed at time  $t_n$ , it is used to obtain the approximate position of the interface, and the volumes  $C_{i,j,k}$  at time  $t_{n+1}$ . The following discussion of momentum advection is based on two VOF advection methods, (Figure 4) Lagrangian Explicit and Weymouth and Yue advection. We first describe the common features of these two methods.

After addition and subtraction of a term proportional to the velocity divergence, equation (2) leads to

$$\partial_t H + \nabla \cdot (\mathbf{u}H) = H \nabla \cdot \mathbf{u}. \quad (17)$$

A discrete representation of this equation is obtained by integrating the advection equation in time and space

$$C_{i,j,k}^{n+1} - C_{i,j,k}^n = - \sum_{\text{faces } f} F_f^{(c)} + \int_{t_n}^{t_{n+1}} dt \int_{\Omega} H \nabla \cdot \mathbf{u} dx, \quad (18)$$

where the first term on the right-hand side is the sum over faces  $f$  of cell  $i, j, k$  of the fluxes  $F_f^{(c)}$  of  $\mathbf{u}H$ . Obviously the “compression” term on the right-hand side disappears for incompressible flow, however it is useful to keep it in view of its utility in split-advection methods. In the previous equation  $F_f^{(c)}$  is defined by

$$F_f^{(c)} = \int_{t_n}^{t_{n+1}} dt \int_f u_f(\mathbf{x}, t) H(\mathbf{x}, t) dx, \quad (19)$$

where  $u_f = \mathbf{u} \cdot \mathbf{n}_f$  and  $\mathbf{n}_f$  is the unit normal vector of face  $f$ , pointing outwards of cell  $i, j, k$ .

Once an approximation for the evolution of  $\mathbf{u}(\mathbf{x}, t)H(\mathbf{x}, t)$  during the time step is chosen, a four-dimensional integral remains to be computed in equation (19). In what follows we describe the approximations used. However, all the methods we describe here are directionally split. The methods are also designed to preserve the property that  $0 \leq C_{i,j,k} \leq 1$  which we call  $C$ -bracketing. It is important to preserve  $C$ -bracketing despite the various approximations used, in order to avoid the arbitrary addition or removal of mass.

Directional splitting results in the breakdown of equation (18) into three equations

$$C_{i,j,k}^{n,m+1} - C_{i,j,k}^{n,m} = -F_{m-}^{(c)} - F_{m+}^{(c)} + c_m \delta_m^h u_m, \quad (20)$$

where  $m = 1, 2, 3$ ,  $C_{i,j,k}^{n,1} = C_{i,j,k}^n$ ,  $C_{i,j,k}^{n,4} = C_{i,j,k}^{n+1}$ , the face labelled “ $m-$ ” is the “left” face in direction  $m$  with  $F_{m-}^{(c)} \geq 0$  if the flow is locally from “right” to “left”, and finally the face “ $m+$ ” is the “right” face in direction  $m$  with  $F_{m+}^{(c)} \geq 0$  if the flow is from “left” to “right”. We have also approximated the compression term in (18) by

$$\int_{t_n}^{t_{n+1}} dt \int_{\Omega} H \partial_m u_m d\mathbf{x} \simeq c_m \partial_m^h u_m, \quad (21)$$

with no implicit summation rule. In the RHS of (20) and 21) the flux terms  $F_f^{(c)}$  and the partial derivative  $\partial_m u_m$  must be evaluated with the same discretized velocities. In particular,  $\partial_m^h u_m$  is a finite difference or finite volume approximation of the spatial derivative of the  $m$ th component of the velocity vector in direction  $m$ , and the “compression coefficient”  $c_m$  approximates the color fraction. Its exact expression is dependent on the version of the directionally split advection method used and will be described below. The value of  $c_m$  is chosen so as to preserve the  $C$ -bracketing condition. Note that because  $c_m$  may not be the same coefficient along the three Cartesian directions, even if the flow is incompressible the sum  $\sum_m c_m \partial_m^h u_m$  of the split compression terms is not necessarily vanishing.

Each application of equation (20) defines an advection substep. After each substep, a new reconstruction of the interface is performed with the updated volumes  $C_{i,j,k}^{n,m+1}$ , computing the normal  $\mathbf{m}$  and the  $\alpha$  constant. This reconstruction is used to estimate the fluxes  $F_f^{(c)}$  in the next split advection substep (20).

### 3.3.5. Lagrangian Explicit advection

The Lagrangian Explicit / Calcul d’Interface Affine par Morceaux CIAM advection method is described in what follows. “Calcul d’Interface Affine par Morceaux” is French for “Piecewise Linear Interface Calculation” (PLIC) but while PLIC refers to a generic VOF method with a piecewise linear reconstruction step, CIAM refers to a specific type of advection method first described in the archival literature in [? ] and classified as the “LE” method in [? ]. CIAM advection is most naturally explained as a Lagrangian transport of the Heaviside function. In each cell, the reconstruction step has defined a planar interface. Lagrangian transport consists in moving each point of this interface by a velocity field  $\mathbf{u}'$  defined as follows: 1) in each sub time step  $m$  of split advection the velocity field is in direction  $m$ , and depends only on the component  $x_m$  of the space variable; 2) the velocity field is interpolated between the face velocities  $u_{m-}$  on the left face and  $u_{m+}$  on the right face so that  $u'_m(x_m) = -u_{m-}(1-x_m) + u_{m+}x_m$  where the origin of the coordinate  $x_m$  was chosen on the left face. (As said before, space has been rescaled so the cell size is unity and recall  $u_f = \mathbf{u} \cdot \mathbf{n}_f$ .) The exact transport by this velocity field is  $dx_m/dt = u'_m(x_m)$  but we use a first-order, explicit approximation as  $x_m^{n+1} - x_m^n = u'_m(x_m^n)$ . (Time was rescaled so the time step is unity and disappears from the equations.) Using the approximation for the interpolated velocity we get

$$x_m^{n+1} = -u_{m-} + (1 + u_{m+} + u_{m-}) x_m^n. \quad (22)$$

This relation gives the advected points of the interface as a function of the original points. It is used to advect the faces and the planar interfaces as shown on Figure 3, for

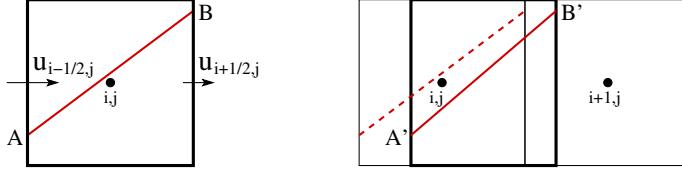


Figure 3: Advection of the AB interface segment. Face velocities at the left face,  $u_{i-1/2,j} = -u_{1-}$ , and the right face,  $u_{i+1/2,j} = u_{1+}$ , are used to create an interpolated linear velocity field, so that point A is advected to point  $A'$  at the velocity  $u_{i-1/2,j}$ ; point B to  $B'$  at the velocity  $u_{i+1/2,j}$ , and a point along AB at a velocity linearly interpolated between  $u_{i-1/2,j}$  and  $u_{i+1/2,j}$ .

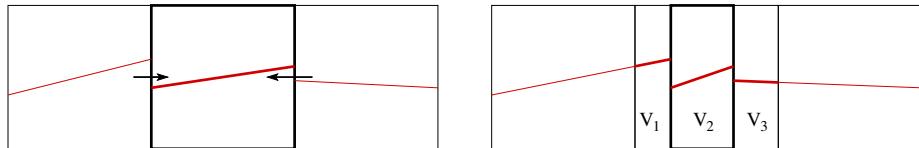


Figure 4: Formation of the volumes  $V_1$ ,  $V_2$  and  $V_3$  by Lagrangian advection in the  $x$  direction. Left: initial reconstruction with the horizontal velocities on the faces of the central cell. Right: segments and volumes  $V_i$  (see text) after Lagrangian advection (for simplicity it has been assumed  $u_{i-3/2,j} = u_{i+3/2,j} = 0$ )

a two-dimensional case. After advection, at most three interface fragments are found in each cell. The resulting volumes (areas in 2D) are shown in Figure 4 and are labelled  $V_1$ ,  $V_2$  and  $V_3$ . Eventually the color function is given by the sum of three contributions

$$C_{i,j,k}^{n,m+1} = V_1 + V_2 + V_3. \quad (23)$$

It is seen that the transform (22) compresses distances by a factor  $1 + \partial_m^h u_m = 1 + u_{m+} + u_{m-}$ . Thus the measure of any region of space is scaled by this amount under advection. The interface is then transformed in three steps: 1) compress all cells by the factor above; 2) move the cells by  $u_{m-}$ ; 3) the moved cells now overlap the original cells. The volume corresponding to the overlap of the cell with itself is  $V_2$ , while  $V_1$  and  $V_3$  correspond to the overlap with the moved cells from the left and right, respectively. The geometrical interpretation of the Lagrangian Explicit advection of Figs. 3 and 4 and the definition (19) of  $F_f^{(c)}$  leads to the following correspondence between the fluxes and the volume contributions  $V_i$ . To start with an example, for the central cell of Fig. 4 the flux on the left face is from left to right, since  $u_{1-} = -u_{i-1/2,j} < 0$ . This means that  $V_1 = -F_{1-,i}^{(c)} = F_{1+,i-1}^{(c)} > 0$  while  $V_3$  of the left cell, that is  $V_{3,i-1}$ , is zero corresponding to an absence in that cell. In all cases the fluxes across a face are equal to the  $V_1$  or  $V_3$  of one cell or the other adjacent to a face. The final expression of the split step is then

$$C_{i,j,k}^{n,m+1} = C_{i,j,k}^{n,m} (1 + u_{m+} + u_{m-}) - F_{m-}^{(c)} - F_{m+}^{(c)}, \quad (24)$$

which shows that the constant  $c_m$  in the compression term is  $C_{i,j,k}^{n,m}$  while the approximation of the derivative is  $\partial_m^h u_m = u_{m+} + u_{m-}$ .

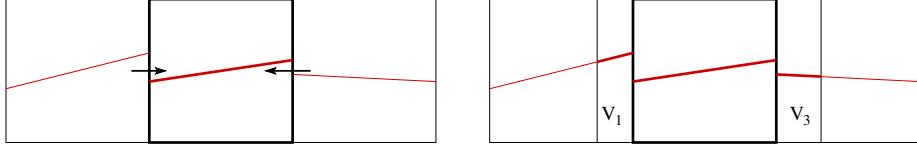


Figure 5: Eulerian flux representation for advection in the  $x$  direction. Left: same initial reconstruction of Fig. 4 with the horizontal velocities on the faces of the central cell. Right: fluxes, or volumes  $V_1$  and  $V_3$ , are calculated directly from the interface reconstruction in each cell

It is interesting to note that using three Lagrangian advects in a sequence does not result in volume conservation to machine accuracy. Indeed the summation of the three substeps (20) results in

$$C_{i,j,k}^{n+1} - C_{i,j,k}^n = - \sum_{\text{faces } f} F_f^{(c)} + \sum_{m=1}^3 C_{i,j,k}^{n,m} (u_{m+} + u_{m-}). \quad (25)$$

While the flux terms cancel upon integration over the domain, the sum of the compressive terms does not vanish since  $C_{i,j,k}^{n,m}$  changes during the three substeps. We recover eq. (20) with

$$c_m = C_{i,j,k}^{n,m} \quad (26)$$

### 3.3.6. Weymouth and Yue advection

Another method that we shall describe is the exactly mass-conserving method of Weymouth and Yue (WY) described in ref. [? ]. In that method the coefficient of the compression term (21) is independent of the direction  $m$  so that  $c_m = c$ . The coefficient  $c$  is defined as

$$c = \Theta(C_{i,j,k}^n - 1/2) \quad (27)$$

where  $\Theta$  is the one-dimensional Heaviside function (defined on the real axis). That is  $c = 0$  if  $C_{i,j,k}^n < 1/2$  and  $c = 1$  if  $C_{i,j,k}^n \geq 1/2$ . The fluxes  $F_f^{(c)}$  are also defined differently. The left fluxed volume in direction  $m = 1$  is equal to the volume fraction  $F_{1-}^{(c)}$  in a cuboid of width  $u_{i-1/2,j,k}$  (recall that  $\tau = 1$ ) adjacent to the left face  $f = 1-$ . This fluxed volume corresponds to “Eulerian Implicit” (EI) advection in the terminology of [? ] and is represented by the volume  $V_1$  on Figure 5. Using these definitions, Weymouth and Yue were able to show in ref. [? ] that the final result obeys  $C$ -bracketing (see Section 3.3.4).

Using this advection method results in volume conservation at machine accuracy. Indeed the summation of the three substeps (20) results in

$$C_{i,j,k}^{n+1} - C_{i,j,k}^n = - \sum_{\text{faces } f} F_f^{(c)} + c \sum_{m=1}^3 \partial_m^h u_m \quad (28)$$

Since  $\sum_{m=1}^3 \partial_m^h u_m = \sum_{m=1}^3 (u_{m+} + u_{m-})$  is the simple finite-volume expression for  $\nabla \cdot \mathbf{u}$ , it disappears and mass is conserved at the accuracy with which condition (3) is satisfied.

### 3.3.7. Clipping

The algorithm that has been coded involves a number of additional steps designed to avoid unwanted effects of arithmetic floating point round-off error. The most important one is clipping: at the end of each directional advection, the values of  $C_{i,j,k}$  are reset so that  $C_{ijk}$  is set to 0 if  $C_{ijk} < \epsilon_c$  and  $C_{ijk}$  is set to 1 if  $C_{ijk} > 1 - \epsilon_c$ . When there is no surface tension the choice  $\epsilon_c = 10^{-12}$  works well. Otherwise  $\epsilon_c = 10^{-8}$  gives more stable results with smoother interface shapes. This stronger clipping is a necessity for some simulations with WY, while the CIAM method can survive  $\epsilon_c = 10^{-12}$  with the diffusive interpolation schemes described below. As a matter of fact we observe that WY produces many more “wisps”, i.e. cells with tiny values of  $1 - C_{i,j,k}$  inside the liquid or  $C_{i,j,k}$  inside the gas. We have not yet been able to determine the origin of this need for a more forceful clipping with WY, but it could be related to the fact that the CIAM method has a geometrical interpretation, while WY is intrinsically algebraic in nature.

## 3.4. Momentum-advection methods

### 3.4.1. Advection of a generic conserved quantity

Consider the advection of a generic *conserved* quantity  $\phi$  by a continuous velocity field

$$\partial_t \phi + \nabla \cdot (\phi \mathbf{u}) = 0. \quad (29)$$

We assume that  $\phi$  is smoothly varying except on the interface where it may be discontinuous. Indeed finding a correct scheme for the advection of this discontinuity, at the same speed as the advection of the volume fraction, is the goal of the present study. The smoothness of the advected quantity away from the interface is verified for the density  $\rho$ , the momentum  $\rho \mathbf{u}$  or the internal energy  $\rho e$ . Integrating the advection equation in time one obtains

$$\phi_{i,j,k}^{n+1} - \phi_{i,j,k}^n = - \sum_{\text{faces } f} F_f^{(\phi)}, \quad (30)$$

where the sum of the right-hand side is the sum over faces  $f$  of cell  $i, j, k$  of the fluxes  $F_f^{(\phi)}$  of  $\phi$ . In equations,  $F_f^{(\phi)}$  is defined in a similar manner to equation (19) for the color function fluxes.

$$F_f^{(\phi)} = \int_{t_n}^{t_{n+1}} dt \int_f u_f(\mathbf{x}, t) \phi(\mathbf{x}, t) d\mathbf{x}, \quad (31)$$

where

$$u_f = \mathbf{u} \cdot \mathbf{n}_f, \quad (32)$$

and  $\mathbf{n}_f$  is the unit normal vector of face  $f$ , pointing outwards of cell  $i, j, k$ . In order to “extract” the discontinuity the flux can be expressed as

$$F_f^{(\phi)} = \int_{t_n}^{t_{n+1}} dt \int_f [u_f H \phi + u_f (1 - H) \phi] d\mathbf{x}, \quad (33)$$

which can be re-expressed in terms of the fluxes  $F^{(c)}$  in eq. (19) as

$$F_f^{(\phi)} = \bar{\phi}_1 \int_{t_n}^{t_{n+1}} dt \int_f u_f H d\mathbf{x} + \bar{\phi}_2 \int_{t_n}^{t_{n+1}} dt \int_f u_f (1 - H) d\mathbf{x} \quad (34)$$

where

$$\bar{\phi}_s = \frac{\int_{t_n}^{t_{n+1}} dt \int_f u_f H_s \phi \, d\mathbf{x}}{\int_{t_n}^{t_{n+1}} dt \int_f u_f H_s \, d\mathbf{x}}, \quad (35)$$

and  $H_1 = H$ ,  $H_2 = 1 - H$ . This can be expressed as

$$F_f^{(\phi)} = \bar{\phi}_1 F_f^{(c)} + \bar{\phi}_2 F_f^{(1-c)}, \quad (36)$$

where  $F_f^{(c)}$  is the flux of  $H$  defined in (19) and  $F_f^{(1-c)}$  is the flux obtained by replacing  $H$  by  $1 - H$  in (19).

### 3.4.2. Cloning the tracers

When a cell is cut by the interface, and the field  $\phi$  is non-smooth, it becomes difficult to estimate the integrals in (34). A possibility is to define two new fields  $\phi_{1,2}$  so that  $\phi_s = \phi$  inside phase  $s$ . Then

$$\phi = H\phi_1 + (1 - H)\phi_2. \quad (37)$$

In the discretized equation, two discrete fields  $\phi_{s,i,j,k}$  are defined over the entire domain  $\Omega$ . This is more costly in memory usage but simplifies considerably the computation of the averages in Equation (34). The two equations (29) and (2) are now replaced by three equations, the same volume fraction equation (2) and the two equations

$$\partial_t \phi_1 + \nabla \cdot (\phi_1 \mathbf{u}) = 0, \quad (38)$$

$$\partial_t \phi_2 + \nabla \cdot (\phi_2 \mathbf{u}) = 0. \quad (39)$$

The three equations (2,38-39) now imply (29). This addition of a pair of “cloned” variables to help deal with large variations of density is similar to the methods used for the resolution of the momentum and energy equations for compressible flow. For example Saurel and Abgrall used two density, momentum and energy variables in ref. [?], leading to their seven-equations model, while Allaire, Clerc and Kokh use two density variables in [?], leading to their five-equations model. The addition of a cloned tracer variable in incompressible isothermal flow was also implemented by Popinet in the “Basilisk” code [?].

### 3.4.3. Advection of the density field

Let us now apply the above considerations to the advection of the density field by a general continuous velocity field. Although in the incompressible case the density follows trivially the color function, it will allow us to introduce an important point about the equivalence of tracer and VOF advection. The density  $\rho(\mathbf{x}, t)$  obeys eq. (29) with  $\phi = \rho$ . The advection of density may be made consistent with the advection of the color function by using fluxes  $F_f^{(c)}$  and estimating the average occurring in equation (35),  $\bar{\phi}_1 = \bar{\rho}$  from averages over control volumes or approximations on faces.

The case relevant here is that of an incompressible velocity field, with uniform density in each phase. Then  $\rho$  is exactly proportional to  $H$  (Eq. 1). The constancy of  $\rho$

allows to extract it trivially from the integrals in eq. (34) and to obtain *exactly*  $\bar{\rho}_s = \rho_s$ . Then the flux of  $\rho$  is

$$F_f^{(\rho)} = \rho_1 F_f^{(c)} + \rho_2 F_f^{(1-c)}. \quad (40)$$

Using this flux definition for  $\rho$ , and any VOF method for the fluxes of the color function, one obtains a conservative method for  $\rho$  since eq. (30) evolves  $\rho$  as a difference of fluxes. That is, one conserves the total mass and  $\sum_{i,j,k} \rho_{i,j,k}^{(n)} = \sum_{i,j,k} \rho_{i,j,k}^{(n+1)}$ . However this result is not consistent with the advection of the color function in the CIAM case, since as we have seen above, CIAM does not conserve volumes exactly, that is  $\sum_{i,j,k} C_{i,j,k}^{(n)} = \sum_{i,j,k} C_{i,j,k}^{(n+1)}$  is not in general true. As a result the advection of  $\rho$  is not consistent with the advection of  $C$ .

The paradox may be resolved if one notices that the compression term is missing in the equation for the advection of  $\rho$ . One should keep the compression term, which is not changing the equation in the case of incompressible velocity fields, so that the equation for the conserved quantity  $\phi$  or  $\rho$  becomes

$$\partial_t \phi + \nabla \cdot (\phi \mathbf{u}) = \phi \nabla \cdot \mathbf{u} \quad (41)$$

It is then possible to define the evolution of  $\rho = \phi$  through a sequence of directionally split operations exactly equivalent to the operations performed on the color function.

$$\phi_{i,j,k}^{n,m+1} - \phi_{i,j,k}^{n,m} = -F_{m-}^{(\phi)} - F_{m+}^{(\phi)} + (\tilde{\phi}_1^m c_m^{(1)} + \tilde{\phi}_2^m c_m^{(2)}) \partial_m^h u_m \quad (42)$$

where  $F_{m\pm}^{(\phi)}$  is defined as above, and

$$\tilde{\phi}_s^m = \frac{\int_{t_n}^{t_{n+1}} dt \int_{\Omega} \phi H_s \partial_m^h u_m d\mathbf{x}}{\int_{t_n}^{t_{n+1}} dt \int_{\Omega} H_s \partial_m^h u_m d\mathbf{x}} \quad (43)$$

and where  $c_m^{(1)} = c_m$  is the compression coefficient computed as in the VOF advection from  $C_{i,j,k}^{n,m}$  and  $c_m^{(2)} = 1 - c_m$  corresponds to the symmetric color fraction  $1 - C_{i,j,k}^{n,m}$ . Specifically for  $\rho$  this gives

$$\rho_{i,j,k}^{n,m+1} - \rho_{i,j,k}^{n,m} = -F_{m-}^{(\rho)} - F_{m+}^{(\rho)} + C_m^{(\rho)}, \quad (44)$$

where the fluxes are given by (40) and the central compression part is

$$C_m^{(\rho)} = (\rho_1 c_m^{(1)} + \rho_2 c_m^{(2)}) \partial_m^h u_m \quad (45)$$

with no implicit summation on  $m$ . The  $c_m^{(s)}$  values are given by (26) or (27). It is interesting to note that for the WY method, the compression terms eventually cancel and mass is conserved at the same accuracy as the discrete incompressibility condition  $\partial_m^h u_m = 0$  is verified.

#### 3.4.4. Momentum advection: basic expressions

Momentum advection can be performed following the general outline above. Each momentum component can be treated by considering the transport of the conserved quantity  $\phi = \rho u_q$  where  $q$  is the component index. Reworking equations (33,34) and

(35), with this definition, we obtain for the weighted averages  $\bar{\phi}_s = \overline{\rho u_{q,s}}$  in phase  $s$  the expression

$$\overline{\rho u_{q,s}} = \rho_s \bar{u}_{q,s} \quad (46)$$

where

$$\bar{u}_{q,s} = \frac{\int_{t_n}^{t_{n+1}} dt \int_f u_q u_f H_s d\mathbf{x}}{\int_{t_n}^{t_{n+1}} dt \int_f u_f H_s d\mathbf{x}} \quad (47)$$

We term  $\bar{u}_{q,s}$  the “adverted interpolated velocity” and we will explain below how these “adverted” velocities are computed. Thus the evolution of the momentum is given by

$$\begin{aligned} (\rho_{i,j,k} u_{q;i,j,k})^{n,m+1} - (\rho_{i,j,k} u_{q;i,j,k})^{n,m} &= \\ -F_{m-}^{(\rho u)} - F_{m+}^{(\rho u)} + (\rho_1 \tilde{u}_{q,1}^m C_m^{(1)} &+ \rho_2 \tilde{u}_{q,2}^m C_m^{(2)}) \partial_m^h u_m \end{aligned} \quad (48)$$

where

$$F_f^{(\rho u)} = \rho_1 \bar{u}_{q,1} F_f^{(c)} + \rho_2 \bar{u}_{q,2} F_f^{(1-c)}, \quad (49)$$

and the “central interpolated velocity” corresponding to the  $\bar{\phi}_s$  are

$$\tilde{u}_{q,s}^m = \frac{\int_{t_n}^{t_{n+1}} dt \int_\Omega u_q^{n,p} H_s \partial_m^h u_m d\mathbf{x}}{\int_{t_n}^{t_{n+1}} dt \int_\Omega H_s \partial_m^h u_m d\mathbf{x}} \quad (50)$$

where for CIAM  $p = m$  and for WY  $p = 1$ . Thus  $\tilde{u}_{q,s}^m = \tilde{u}_{q,s}$  remains constant while the various steps  $m = 1, \dots, 3$  of directional splitting are performed. Note that we otherwise omit the superscript  $m$  for  $\bar{u}_q$  and  $\tilde{u}_q$  to avoid too complex notations. Notice that “cloning” the advected velocities  $\bar{u}_{q,1}$  and  $\bar{u}_{q,2}$  would make it easier to advect a velocity field with a jump on the interface. However in viscous flow without phase change the velocity is continuous on the interface, and to avoid an excessively complicated method we approximate the velocity field as continuous and we choose approximations of the “adverted interpolated velocity” such that  $\bar{u}_q = \bar{u}_{q,1} = \bar{u}_{q,2}$  and “central interpolated velocity” such that  $\tilde{u}_q = \tilde{u}_{q,1} = \tilde{u}_{q,2}$ . An important simplification is then

$$F_f^{(\rho u)} = \bar{u}_q F_f^{(\rho)} \quad (51)$$

(which is the central equation in this development) and thus

$$(\rho_{i,j,k} u_{q;i,j,k})^{n,m+1} - (\rho_{i,j,k} u_{q;i,j,k})^{n,m} = -\bar{u}_q F_{m-}^{(\rho)} - \bar{u}_q F_{m+}^{(\rho)} + \tilde{u}_q^m C_m^{(\rho)}, \quad (52)$$

where the density fluxes are defined in (40) and the central term  $C^{(\rho)}$  is defined in (45). In the above expression the weighted average velocities  $\bar{u}_q$  are computed using (47) on the corresponding left face  $m-$  or right face  $m+$ .

The scheme above must be combined with an advection scheme for the color fraction. When the CIAM scheme is used, the  $C_m^{(\rho)}$  term in (52) must be kept to ensure consistency with VOF advection. This term is computed using the above definition (26). This term does not cancel when the final momentum is computed from the three directionally-split advects and the result is not exactly conservative. On the other hand when the WY scheme is used the  $c_m^{(s)}$  and  $\tilde{u}_q^m$  factors are independent of  $m$  and

$$C_m^{(\rho)} = [\rho_1 c + \rho_2 (1 - c)] \partial_m^h u_m \quad (53)$$

so provided the velocity field is incompressible (that is  $\sum_{m=1}^3 \partial_m^h u_m = 0$ ) summing the three split advection expressions (52) one obtains a cancellation of the central terms and

$$(\rho_{i,j,k} u_{q;i,j,k})^{n,4} - (\rho_{i,j,k} u_{q;i,j,k})^n = - \sum_{\text{faces f}} \bar{u}_{q,f} F_f^{(\rho)}. \quad (54)$$

The version of our method coupled with WY advection is thus exactly conservative. Both CIAM and WY methods can be rewritten in a slightly different notation useful for future discussion

$$\hat{g}_q^{n+1} - g_q^n = - \sum_{\text{faces f}} \bar{u}_{q,f} F_f^{(\rho)}(C^n, u_f) + \sum_{m=1}^3 \tilde{u}_q^m C_m^{(\rho)}, \quad (55)$$

where  $\hat{g}_q^{n+1} = (\rho_{i,j,k} u_{q;i,j,k})^{n,4}$  is the  $q$ -component of the final momentum,  $g_q^n = (\rho_{i,j,k} u_{q;i,j,k})^n$  is the initial one, and we have written explicitly the dependence of  $F_f^{(\rho)}(C^n, u_f)$  on the face velocities and on the volume fraction field in the neighborhood of the interface.

#### 3.4.5. Momentum advection: interpolations and flux limiters

The evolution of momentum as expressed in eq. (55) can be approximated either 1) away from the interface, in the bulk of the phases or 2) in the neighborhood of the interface. In the first case the expression simplifies considerably since the density as well as the color fraction are constant thus  $\rho_{ijk} = \rho_0$  and  $C^n = C_0$ . The spurious central term is also removed so

$$u_q^* - u_q^n = - \sum_{\text{faces f}} \bar{u}_{q,f} u_f. \quad (56)$$

We can distinguish an “advecting” velocity  $u_f$  and an “advected velocity component”  $\bar{u}_{q,f}$  which must be estimated in an integral relating to face  $f$ . Both components are either centered on the face of the staggered set of cells / control volumes for the component  $u_q$  or estimated close to the face  $f$ . Typical hyperbolic equation methods use interpolants to perform these estimations. The interpolants we use here are one-dimensional and work from the node velocities on a segment aligned with the direction of advection, that is the direction perpendicular to face  $f$ . Thus the scheme in the bulk is

$$u_q^* - u_q^n = - \sum_{\text{faces f}} \bar{u}_{q,f}^{(\text{advected})} u_f^{(\text{advecting})}. \quad (57)$$

Near the interface we have instead

$$\hat{g}_q^{n+1} - g_q^n = - \sum_{\text{faces f}} \bar{u}_{q,f}^{(\text{advected})} F_f^{(\rho)}(C^n, u_f^{(\text{advecting})}) + \sum_{m=1}^3 \tilde{u}_q^m C_m^{(\rho)}, \quad (58)$$

To estimate the advecting velocities  $u_f^{(\text{advecting})}$  we use centered schemes. (see Figure 1 for a visual depiction of the arrangement of the staggered nodes.) Consider for example that the face is perpendicular to direction 1 so  $f = 1_{\pm}$ . There are two cases. In the first case the advected component is not aligned with the face normal so in the example  $q \neq 1$ . Consider specifically the case  $q = 2$  depicted in Figure 1(b). The staggered  $u_2$  control volumes are centered on  $i, j + 1/2, k$ . The face  $f = 1_{\pm}$  is then centered on

$i - 1/2, j + 1/2, k$ . Then  $u_f^{(\text{advection})} = u_{1_-}^{(\text{advection})} = u_{1;i-1/2,j+1/2,k}$  which is not given and has to be interpolated by

$$u_{1;i-1/2,j+1/2,k}^{(\text{advection})} = \frac{1}{2}(u_{1;i-1/2,j,k} + u_{1;i-1/2,j+1,k}). \quad (59)$$

In the second case the advected component is aligned with the face normal so in the example  $q = 1$  also depicted in Figure 1(a). The staggered  $u_1$  control volumes are centered on  $i + 1/2, j, k$ . The face  $f = 1_\pm$  is then centered on  $i - 1, j, k$  and the interpolation is

$$u_{1;i-1,j,k}^{(\text{advection})} = \frac{1}{2}(u_{1;i-1/2,j,k} + u_{1;i+1/2,j,k}). \quad (60)$$

Now we turn to the definition of the *adverted* velocity values  $u_{q,f}$  in expression (58). We still take the example where the face is perpendicular to direction 1 so  $f = 1_\pm$  (we take the case  $f = 1_-$  to illustrate, see Figure 6(a)) thus the *adverting* direction is  $x_1$ . We only need to consider the dependence of the advected velocity on  $x_1$  to define the interpolation. We consider the regularly spaced one-dimensional grid of advected velocities. In order to use lighter notations we will note  $\phi = u_q$  the *adverted* velocity component, with known node values at

$$\phi_p = u_q(x_{1,1_-} + ph) \quad (61)$$

where  $x_{1,1_-}$  is the coordinate of the face  $f = 1_-$ . The known values  $\phi_p$  in the staggered grid are at nodes that are a half-integer away from the face, which implies that the known values are

$$u_{1;i-3/2,j,k} = \phi_{-3/2}, \quad u_{1;i-1/2,j,k} = \phi_{-1/2}, \quad u_{1;i-1/2,j,k} = \phi_{1/2}, \dots \quad (62)$$

when  $q = 1$ ,  $x_{1,1_-} = ih$  and

$$u_{2;i-2,j+1/2,k} = \phi_{-3/2}, \quad u_{2;i-1,j+1/2,k} = \phi_{-1/2}, \quad u_{2;i,j+1/2,k} = \phi_{1/2}, \dots \quad (63)$$

when  $q = 2$ . For  $q = 2$  or 3, we have  $x_{1,1_-} = (i - 1/2)h$  so the  $q = 3$  case follows easily. We need to predict  $\phi_0$  to serve as an approximation of  $\bar{u}_q$  given in (47). An interpolation function is now defined that predicts this value as a function of the four nearest points, and in an upwind manner based on the sign of the *adverting* velocity  $u_f$ , given by centered interpolations. The value at the face is given by

$$\phi_0 = f(\phi_{-3/2}, \phi_{-1/2}, \phi_{1/2}, \phi_{3/2}, \text{sign}(u_f)) \quad (64)$$

In this study we have extensively tested two kinds of interpolations or schemes

1. A scheme that uses a QUICK third order interpolant in the bulk, away from the interface and a simple first order upwind flux near the interface. We call this scheme QUICK-UW.
2. A scheme that uses a Superbee slope limiter [?]. for the flux in the bulk and a more complex Superbee limiter tuned to a shifted interpolation point near the interface. We call naturally call this scheme “Superbee”.

We describe each scheme in turn, starting with the QUICK-UW. For positive advecting velocity and in the bulk we have

$$\phi_0 = \frac{3}{4}\phi_{-1/2} + \frac{3}{8}\phi_{1/2} - \frac{1}{8}\phi_{-3/2} \quad (65)$$

while near the interface  $\phi_0 = \phi_{-1/2}$ . For negative (to the left) advecting velocity and in the bulk we have

$$\phi_0 = \frac{3}{4}\phi_{1/2} + \frac{3}{8}\phi_{-1/2} - \frac{1}{8}\phi_{3/2} \quad (66)$$

while near the interface  $\phi_0 = \phi_{1/2}$  (see also [? ]) In the Superbee case, we start by defining a general family of interpolants that can be expressed as

$$f(\phi_{-3/2}, \phi_{-1/2}, \phi_{1/2}, \phi_{3/2}, \text{sign}(u_f)) = \phi_{-1/2} + Sh/2 \quad (67)$$

for a positive advecting velocity where the slope  $S$  is given by a slope-limiter function  $g$  such that  $S = g(\phi_{-3/2}, \phi_{-1/2}, \phi_{1/2})$  while for a negative advecting velocity

$$f(\phi_{-3/2}, \phi_{-1/2}, \phi_{1/2}, \phi_{3/2}, \text{sign}(u_f)) = \phi_{1/2} - Sh/2 \quad (68)$$

where  $s = g(\phi_{-1/2}, \phi_{1/2}, \phi_{3/2})$ .

To define the slope  $S$  used in Superbee consider the factor  $\alpha$  to be the smallest in absolute value of

$$\begin{aligned} \alpha^+ &= (\phi_{1/2,j} - \phi_{-1/2,j})/h, \\ \alpha^- &= 2(\phi_{-1/2,j} - \phi_{-3/2,j})/h. \end{aligned} \quad (69)$$

and the factor  $\beta$  to be the smallest in absolute value of

$$\begin{aligned} \beta^+ &= 2(\phi_{1/2,j} - \phi_{-1/2,j})/h, \\ \beta^- &= (\phi_{-1/2,j} - \phi_{-3/2,j})/h. \end{aligned} \quad (70)$$

Then one prescribes  $S$  as the largest, in absolute value, of  $\alpha$  and  $\beta$ .

A slightly different estimate of the Superbee advected velocity is used near the interface. First we extend the definition of the interpolants as we shall predict  $\bar{u}_q$  at a point  $\hat{x}$  slightly upwind from  $x_f$  using a new function  $\hat{f}$  so that

$$\hat{f}(\hat{x}, \phi_{-3/2}, \phi_{-1/2}, \phi_{1/2}, \phi_{3/2}, \text{sign}(u_f)). \quad (71)$$

We take  $\hat{x} = x_f - u_f \tau/2$  to be the midpoint of the fluxed region. (This is the region from which flow lines crossing the face originate see Figure 6(b).) The extended interpolant is defined for positive velocity as

$$\hat{f}(\hat{x}, \phi_{-3/2}, \phi_{-1/2}, \phi_{1/2}, \phi_{3/2}, \text{sign}(u_f)) = \phi_{-1/2} + S|\hat{x} - x_{-1/2}| \quad (72)$$

and for negative velocity as

$$\hat{f}(\hat{x}, \phi_{-3/2}, \phi_{-1/2}, \phi_{1/2}, \phi_{3/2}, \text{sign}(u_f)) = \phi_{1/2} - S|\hat{x} - x_{1/2}| \quad (73)$$

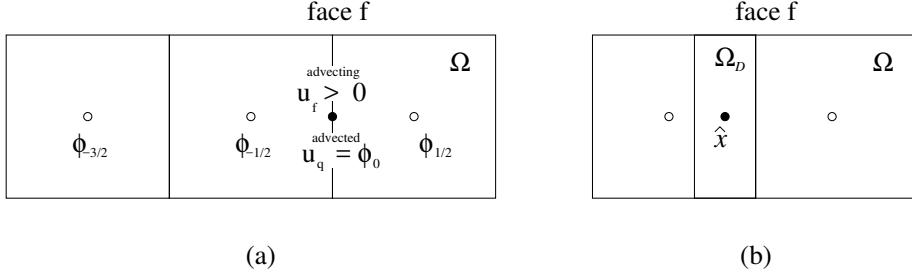


Figure 6: (a) The advected and advecting velocities in expression (58) are both located on or near face  $f$  (in this case  $f = 1-$ ). The reference control volume  $\Omega$  for the advected velocity component  $u_q$  is also shown. The arrangement is the same whatever the component  $q$  but here the example of a horizontal advection (case  $f = 1-$ ) is taken. The value of the advected velocity component  $\bar{u}_q = \phi_0$  on the center of face  $f$  (full disk) is interpolated (see text) from the values  $\phi_p$  of  $u_q$  on the nodes (open disks). (b) A more sophisticated interpolation predicts the value  $\phi(\hat{x})$  where  $\hat{x}$  is at the center of the “donating” region  $\Omega_D$  (see text).

The rationale behind this choice is as follows. For a time-independent advecting velocity field  $u(\mathbf{x}, t)$ , the integrals in expression (47) can be simplified

$$\phi_0 = \bar{u}_q = \frac{\int_{\Omega_D} u_f u_q \, d\mathbf{x}}{\int_{\Omega_D} u_f \, d\mathbf{x}}. \quad (74)$$

where  $\Omega_D$  is the “donating region” (see [? ]) from which flow lines crossing the face originate. Then since we approximate the advecting velocity  $u_f$  by its midpoint value the integral can be further simplified as

$$\bar{u}_q = \frac{1}{|\Omega_D|} \int_{\Omega_D} u_q \, d\mathbf{x} \quad (75)$$

Since the “midpoint” at  $\hat{x}$  is the center of mass of the donating region  $\Omega_D$  the interpolation expression (71) follows.

### 3.4.6. VOF-consistent Momentum advection: staggered grids.

In order to apply the above method on the staggered grid, the fluxes and control volumes of the momentum must be in correspondence with the fluxes and control volumes of the VOF color function. This is realized by recreating in each velocity control volume the necessary color fraction data.

At the start of the velocity advection operations summarized by the operator  $\mathcal{L}_{\text{conv}}^h$  each velocity control volume say  $\Omega_{i+1,j,k}$  overlaps two VOF control volumes  $\Omega_{i,j,k}$  and  $\Omega_{i+1,j,k}$  ( see figure 7) The following operations are performed:

1. Reconstruction of the volume fractions and density  $C^n$  and  $\rho^n$  in the staggered control volumes  $\Omega_{i+1/2,j,k}$  corresponding to the momentum component  $\rho u_1$ , thus obtaining the “half fractions”  $\tilde{C}^n$  and  $\tilde{\rho}^n$ .
2. Computation of the initial momentum component  $\tilde{g}_1^n = \tilde{\rho}^n u_1^n$ .

3. Split advection of the momentum in the  $x$  direction using the method in the previous sections to obtain the first component of the new momentum  $\hat{g}_1^{n+1}$ .
4. Split advection in the  $x$  direction of the volume fractions and density in the staggered control volumes using the volume of fluid method to obtain predicted densities  $\hat{\rho}^{n+1,1}$ .
5. Repeat the previous split advection operations for momentum and density in  $y$  and  $z$ . At each time step, the sequence  $x, y, z$  is permuted. Eventually one obtains  $\hat{g}_1^{n+1}$  and  $\hat{\rho}^{n+1}$ .
6. Repeat the previous operations for the two other momentum components  $g_2$  and  $g_3$ .
7. Extraction of the new velocities components  $\mathbf{u}^* = \hat{\mathbf{g}}^{n+1}/\hat{\rho}^{n+1}$ .
8. In parallel, computation of  $C^{n+1}$  from  $C^n$  on the centered cells using the VOF method.

We now cover the eight steps in detail. The reconstruction of the volume fractions in the staggered cells (step 1) implies considering each control volume for a velocity component, say  $\Omega_1 = \Omega_{i+1/2,j,k}$  for the  $u_1$  component, and finding the “half” volume fraction in each intersection  $\Omega_{i+1/2,j,k} \cap \Omega_{i,j,k}$  and  $\Omega_{i+1/2,j,k} \cap \Omega_{i+1,j,k}$  (see Figure 7). The “half-fractions” are found using the reconstruction parameters  $\alpha$ ,  $\mathbf{m}$  in equation (16) in the original centered control volume, and then computing the corresponding volume fractions in the half cells. Addition of the two half fractions leads to an estimate of the volume fraction  $\tilde{C}_{i+1/2,j,k}^n$  in the staggered cells, and thus of the densities  $\tilde{\rho}_{i+1/2,j,k}^n$ .

In step 2, now that the staggered cell densities at the beginning of the time step are known, obtain the momentum component  $g_{1,i+1/2,j,k}^n = \rho_{i+1/2,j,k}^n u_{1,i+1/2,j,k}^n$ . This momentum component can be advected (step 3) using the method described in the previous section, to obtain  $\hat{g}_{1,i+1/2,j,k}^{n+1}$  in a conservative manner following equation (55). Similarly the volume fractions in the staggered cells are also reconstructed and advected in direction  $x_1$  in step 4. Then in step 5 the other directionally split advects are performed. After each direction, the newly obtained densities and advected velocities are used for the next direction but the advecting velocities remain the same. The central velocity stepping is discussed below equation (50) above. Finally the advects are performed in each of the three systems of staggered cells (step 6) to obtain  $\hat{C}_{i+1/2,j,k}^{n+1} = \hat{C}_{i+1/2,j,k}^{n,4}$  where the notation  $C^{n,m}$  was introduced in equation (20) and  $m$  indicates which direction of the split time step was performed. Then  $\hat{\rho}_{i+1/2,j,k}^{n+1} = \rho_1 \hat{C}_{i+1/2,j,k} + \rho_2 (1 - \hat{C}_{i+1/2,j,k})$  obtains. Finally the  $\mathcal{L}_{\text{conv}}$  operator is obtained as (step 7)

$$\mathcal{L}_{\text{conv}} = \frac{1}{\tau} (\hat{\mathbf{g}}^{n+1} - \tilde{\rho}^n \mathbf{u}^n) \quad (76)$$

where as above we use the  $\tilde{\rho}$  notation to indicate that the density has been reconstructed in the staggered cells. In the absence of the other operators in  $\mathcal{L}_2$  this would result in

$$\mathbf{u}^* = \hat{\mathbf{g}}^{n+1}/\hat{\rho}^{n+1}. \quad (77)$$

However (leaving aside the pressure and other contributions in equation (13)) the momentum on the staggered cells at the beginning of the next time step is not equal to the momentum  $\hat{\rho}^{n+1} \mathbf{u}^{n+1}$  at the end of the previous time step but is rather defined as

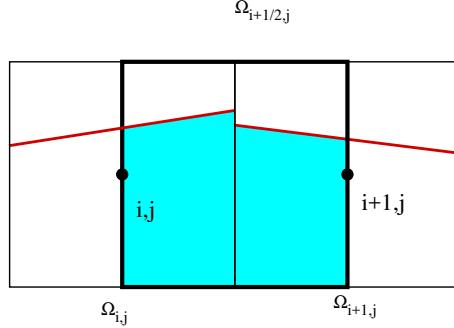


Figure 7: Computation of the staggered fractions from the half-fractions

$\tilde{\rho}^{n+1}\mathbf{u}^{n+1}$  where  $\tilde{\rho}^{n+1}$  is obtained at beginning of the time step from the half-fractions. That is,  $C^{n+1}$  is computed directly from  $C^n$  on the centered cells (step 8) and this allows to find  $\tilde{C}^{n+1}$  and from it,  $\tilde{\rho}^{n+1}$ , by running again step 1. These densities  $\tilde{\rho}^{n+1}$  are not the same as the  $\hat{\rho}^{n+1}$  densities of the previous time step. This implies non conservation of the momentum. We note that attempting to always use only the three sets  $C_{i+1/2,j,k}^n$ ,  $C_{i,j+1/2,k}^n$ ,  $C_{i,j,k+1/2}^n$  and evolve them by the VOF method on the staggered cells would maintain conservation but result in the three staggered grids evolving independently of each other and eventually diverging.

### 3.5. Description of the other time-split terms

The other time-split terms in equation (12) and in the projection step (13) are solved in a standard centered way. The density on the faces of the central cells  $\Omega_{i,j,k}$  is estimated using a centered average  $\rho_{i+1/2,j,k} = (\rho_{i+1/2,j,k} + \rho_{i+1/2,j,k})/2$ . Although this is less accurate and consistent than the usage of the half fractions  $\hat{\rho}_{i+1/2,j,k}$  described above the average is used in these terms both for simplicity reasons and because tests have shown the usage of  $\hat{\rho}_{i+1/2,j,k}$  to lead to less stable simulations than the centered average.

The velocities in the diffusion term are introduced in an explicit way. Although this requires small time steps of the order  $\rho h^2/\mu$  the capillary restriction on time steps is usually even smaller, being of order  $\tau = (\rho h^3/\sigma)^{1/2}$ . The two restrictions become of the same order when  $h \sim l_{\mu\sigma}$  where  $l_{\mu\sigma} = \mu^2/(\sigma\rho)$  is the length at which the viscous and capillary terms balance. For water, this length is of the order of 10 nanometers, and grids of that size are not used in the flows we consider. However, should the velocities be treated in an implicit manner, we do not believe this would change the conclusions of this paper.

Surface tension is computed using the Continuous Surface Force method proposed by [?], together with an estimate of the curvature through the computation of height functions, in a manner that closely follows the method of [?]. The external forces in equation (12) are only gravity and are computed in a trivial manner with  $\frac{1}{\rho^{h+1}} \mathcal{L}_{ext} = \mathbf{g}$ , where gravity  $\mathbf{g}$  is a constant.

## 4. Testing and Validation

### 4.1. Consistent cylinder advection

An elementary test of our method, that mostly verifies that the coding has been performed correctly, considers a uniform planar velocity field  $u_1 = u_2 = 1.6 \times 10^{-2}$  and a droplet of density  $\rho_l = 10^9$  in gas at density  $\rho_g = 1$  with a CFL number of  $0.0256\sqrt{2}$ . Viscosity and surface tension are set to zero in this test. The diameter of the droplet is  $D/h = 3.2$ . The unit domain is discretized on a  $16 \times 16$  grid. The droplet shapes that result are shown on the left of Figure 8. The irregularities seen in the advected droplet are due to the roughness of the VOF approximation at such low resolutions. We repeat the test with conditions close to air/water: viscosity  $\mu_l = 0.1$ ,  $\mu_g = 0.002$  and  $\rho_g = 1$ ,  $\rho_l = 10^3$  with identical results: a viscosity contrast will not generate numerical instabilities on a uniform velocity field, as shown on the right of Figure 8.

### 4.2. Kelvin Helmholtz Instability

The Kelvin-Helmholtz instability arises between unequal velocity fluid streams. It is closely related to the issues addressed in the current paper since it arises in many of the flows for which the current method is designed, such as atomisation. Moreover, the Kelvin-Helmholtz instability is particularly strong on a vortex sheet, since (as we show below) it has for an infinitely thin sheet a divergent growth rate as the wavenumber goes to infinity. Compounding the issue, the baroclinic term of the vorticity equation leads to the creation of a vortex sheet. In previous papers some of us have studied the Kelvin-Helmholtz instability in viscous flows with surface tension [? ? ?]

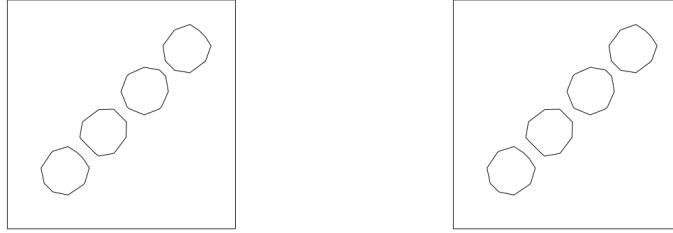


Figure 8: Large-density-ratio droplet in a uniform velocity field: a droplet with  $D/h = 3.2$  grid points per diameter is advected in 3D in the plane  $z = 0$  (see text); left: density ratio  $10^9$  without viscosity, right: density ratio  $10^3$  with viscosity

We focus in this paper on the inviscid, no surface tension case since the objective is to assess the manner in which the advection terms are treated numerically. Moreover the inviscid, no surface tension case is a kind of “worst-case scenario” without the stabilizing effects of viscosity and capillarity.

The simplest setup is that of vortex sheet, for which the growth rate is given by

$$\omega_i = \frac{2\sqrt{r}}{r+1} kU \quad (78)$$

#### 4.3. Sudden acceleration of a cylinder at large density contrast

A test that is often included in studies of momentum-conserving methods [? ? ? ? ?] and other methods designed to improve the stability of two-phase flow computations [?] is to initialize a droplet of very high density at velocity  $\mathbf{u}_l(\mathbf{x}) = U_0 \mathbf{e}_x$  with the other lighter fluid at rest, so that  $\mathbf{u}_g(\mathbf{x}) = 0$ . Surface tension and viscosity are not present as in the previous test, the only difference being the discontinuity of the initial velocity on the interface. The initial velocity condition amounts to a vortex sheet on the surface of the cylinder. After the first time step, the projection method (13) adds a dipole potential flow so that  $\mathbf{u}_g = U_0 \mathbf{e}_x + \tau \nabla p / \rho_g$  in the gas around the droplet, identical to the dipole flow around a solid object. However in addition to the dipole flow there could be small multipole components to the outer gas flow. There are three ways to estimate the perturbations created on the interface.

A vortex sheet is unstable with respect to the Kelvin-Helmholtz instability. The main results about the amplitude of the instability are as follows. Let  $A(t)$  be the maximum deviation of the interface which has a dimension of length. For flows with no viscosity and surface tension as is the case here, the Kelvin-Helmholtz instability amplitude  $A(t)$  should grow exponentially from a perturbation of wavenumber  $k$  as  $A(t) \sim A(0) \exp(s_k t)$  with  $s_k = |\Delta U|k / \sqrt{r}$  where  $\Delta U$  is the tangential velocity difference and  $r = \rho_l / \rho_g$ . In the limit of large  $r$  the growth rate becomes small. Since the maximum wavenumber on the grid is  $\pi/h$  an estimate of the growth rate of the small wavelength instabilities is  $\pi U_0 N / (D \sqrt{r})$  where  $N$  is the number of grid points per diameter. After advection by a droplet diameter, the elapsed time is  $\Delta t = D/U_0$ .

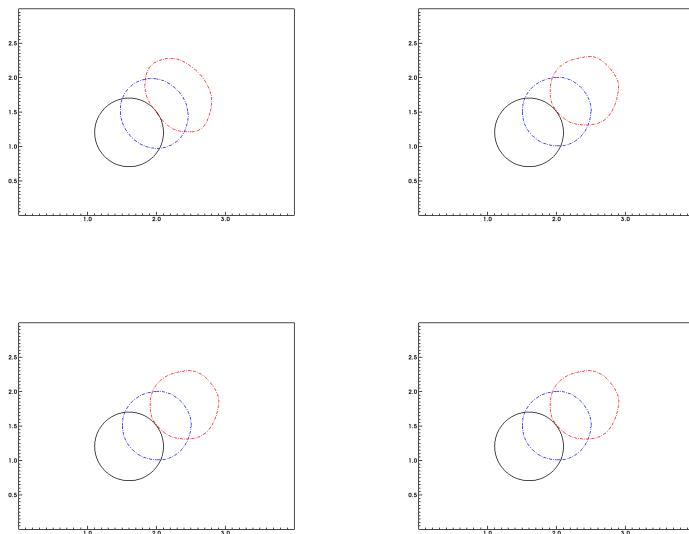


Figure 9: Deformation of a droplet moving through a light fluid. From left to right and top to bottom:  $r = 10, 10^3, 10^6, 10^9$ . Black shape at  $t = 0$ , blue at  $t = D/(2U)$  and red at  $t = D/U$ . The grid resolution is  $D/h = 20$ .

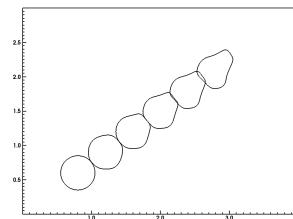


Figure 10: Deformation of a droplet moving through a light fluid at  $r = 10^9$ . The grid resolution is  $D/h = 20$ .

For typical values in the literature of  $r = 10^6$  and an arbitrary value of  $N = 32$  the amplitude growth would be

$$\exp(s_{\max}\Delta t) = \exp\left(\frac{\pi N}{\sqrt{r}}\right) = \exp(0.032\pi) = 1.1058\dots \quad (79)$$

which means the amplitude should grow by 10% after advection by one diameter and by  $e^{1/2}$  after the typical advection by 5 diameters.

Beyond the linear growth stage of the Kelvin-Helmholtz instability, there is a self-similar, non-linear growth stage for which dimensional analysis implies that  $A(t) \sim |\Delta U|t / \sqrt{r}$  [? ]. By this argument also the perturbation of the cylinder should remain of order  $A(\Delta t) \sim D / \sqrt{r}$  after advection by a droplet diameter. One should note that the self-similar growth is obtained for vanishing boundary layer thickness but we precisely expect such a velocity discontinuity in the present case with no viscosity.

A final, physical deformation is expected from the spatial pressure variation induced by the dipole gas flow. This variation involves a larger pressure at the aft and fore stagnation points and a lower pressure along the “equator” of the droplet [? ]. The resulting integrated stress is of order  $\rho_g U^2$  resulting in the growth of the droplet deformation as  $A(t) \sim U^2 t^2 / (Dr)$  and after advection by a droplet diameter as  $A(\Delta t) \sim D/r$ . This growth is observed experimentally [? ] and results in an elliptically shaped drop, albeit of much smaller amplitude than the two former Kelvin-Helmholtz-related growth mechanisms.

The results are shown on Figure 9 for several times and density ratios in a manner comparable to [? ]. For these numerical experiments, the exact manner of initializing the velocity fields has some importance. It is important not to allow in the first time step some gas velocity in the liquid to avoid large pressure gradients in the liquid. Thus the density is initialized to  $\rho_l$  to machine accuracy using the VoF library [? ? ] within a disk implicitly defined by  $x^2 + y^2 + z^2 < R^2$  and the velocity is initialized to 1 for all the velocity nodes inside a disk implicitly defined by  $x^2 + y^2 + z^2 < (R + nh)^2$ , where  $n$  is the size of the “halo” in number of grid points. The velocity in the other nodes is initialized to 0. The tests shown were performed with  $n = 1$ . Increasing the size of the halo from  $n = 0$  to  $n = 1$  improves the results in the first phase of the droplet motion, but not in the late phase.

The velocity of the motion has been oriented on the diagonal as in [? ]. The WY scheme is used with a QUICK interpolant. The droplet deforms little after advection by one droplet diameter (Figure 9). For longer advection the deformation is worse but, as explained above, this is to be expected at high resolution, except in the  $r = 10^9$  case for which we show advection by 5 diameters in Figure 10. If the non-momentum-conserving method is used, the droplet deforms rapidly and the simulation breaks down. As expected the results are better than those of [? ], that uses the skew-symmetric scheme but without momentum-conserving-like schemes, but worse than those of [? ? ? ? ]. It is also seen, as expected from the developments above, but perhaps contrary to intuition, that it is easier to have larger density contrasts. Increasing the number of grid points also helps.

A comment can be made on the nature of the instabilities seen. If the gas velocity numerically diffuses inside the liquid, then some vorticity may penetrate into the liquid

Method	VOF Scheme	velocity scheme	fraction completed
standard	WY	QUICK-UW	0.08
MC	WY	QUICK-UW	0.17
standard	CIAM	Superbee	0.50
MC	CIAM	Superbee	1

Table 1: Percentage of completion of the shear layer test with various methods. All simulations are performed on a  $128 \times 128$  grid with a CFL of 0.03 and a density ratio of 1000.

despite the fact that in inviscid flow vorticity should remain confined on the interface. If this happens, the growth rate of a single-phase Kelvin-Helmholtz instability inside the liquid is the much larger value  $\tilde{s}_k = |\Delta U|k/2$ , without the  $\sqrt{r}$  factor at denominator. (The derivation of this fact may be found for example in [? ].) The instability seen for the  $r = 10^9$  case for the long advection case in Figure 10 is a mode 3, which excludes the large linear growth discussed above for small wavelenghts rather points to the velocity diffusion mechanism. Minimizing numerical diffusion should then be an important quality for this test, while on the other hand numerical diffusion may stabilize other difficult test cases/situations.

#### 4.4. Sheared layer

In order to better analyze the behavior of the methods in flow under shear, such as vortex sheets and Kelvin Helmholtz instabilities, we setup directly a planar, parallel shear flow in a  $(-1/2, 1/2)^2$  domain with the following initial condition

$$\begin{aligned} u_1 &= 15 && \text{if } |x_2| > 1/10 \\ u_1 &= 1 && \text{if } |x_2| < 1/10 \\ u_2 &= 0.01 \sin(2\pi x_1) \exp(-20x_2^2) \end{aligned}$$

and  $\rho = 1$  if  $|x_2| > 1/10$  and  $\rho = 10^3$  otherwise. This flow is similar to a liquid sheet in high velocity gas. The flow is simulated until time  $t_f = 2$  unless the simulation blows up at an earlier time  $t_b < t_f$ . Table 1 shows the numerical schemes that have been used in a few simulations together with the fraction  $t_b/t_f$  of the final time that has been reached. All simulations have been performed on a  $128 \times 128$  grid with a CFL of 0.03 and a density ratio of 1000. The new method (MC) is systematically more stable than the standard method for the two combinations shown in Table 1: a) the WY VOF scheme and the QUICK-UW velocity interpolation schemes and b) the CIAM VOF with the Superbee slope limiters. The state of the two simulations with the first combination of schemes is shown in Figure 11, just before breakdown at time  $t_b = 0.34$  (or  $t_b/t_f = 0.17$ ) for the (MC) method, and at time  $t_b = 0.16$  (or  $t_b/t_f = 0.08$ ) for the standard method. We have also tested a number of other combinations, for example WY & Superbee that turns out to be very unstable. We have also performed simulations with smaller  $64 \times 64$  and  $32 \times 32$  grids that yield similar results. Finally, we note that this case is also unstable in a single-phase configuration when using the QUICK third-order velocity interpolation.

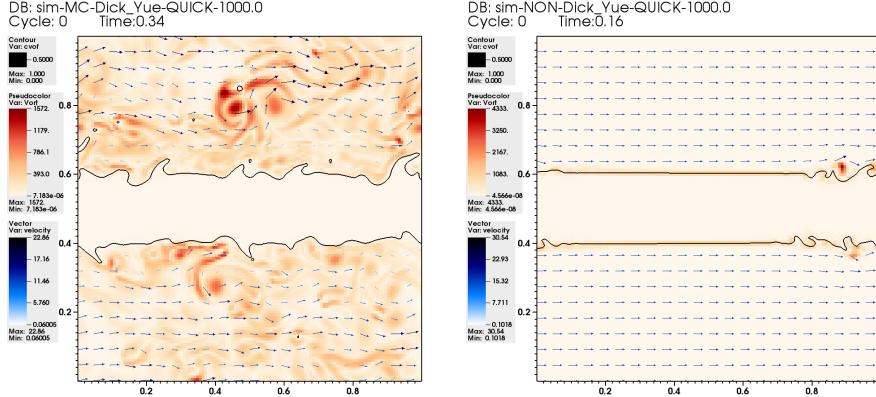


Figure 11: The state of the simulation of the sheared layer just before eventual breakdown of the simulation with the combination of WY VOF scheme and QUICK-UW interpolation. Left: (MC) method, right: standard method.

#### 4.5. Falling Raindrop

A flow configuration that combines the complexities of high density-ratios with the interaction between capillary, viscous and inertial stresses is that of a water droplet falling in air under the influence of gravitational acceleration. The problem is characterized by a combination of Reynolds, Weber and Bond numbers, the definitions of which are as follows :

$$We = \frac{\rho_{\text{air}} U^2 d}{\sigma} \quad , \quad Re = \frac{\rho_{\text{air}} U d}{\mu_{\text{air}}} \quad , \quad Bo = \frac{(\rho_{\text{water}} - \rho_{\text{air}}) g d^2}{\sigma} \quad (80)$$

In our particular numerical setup,  $We \approx 3.2$ ,  $Re \approx 1455$  and  $Bo \approx 1.2$ , thus corresponding to that of a 3mm diameter raindrop (a relatively large one) falling in air at an approximate terminal velocity of 8 m/s (interpolated from empirical data, refer to [?]). The parameters in the problem setup are given in Table 2, and the schematic diagram given by Fig. 12. The droplet is initially placed at the center of a cubic domain (3D), whose side is 4 times the diameter of the drop.

$\rho_{\text{air}}$ (kg/m <sup>3</sup> )	$\rho_{\text{water}}$ (kg/m <sup>3</sup> )	$\mu_{\text{air}}$ (Pa s)	$\mu_{\text{water}}$ (Pa s)	$\sigma$ (N/m)	$d$ (m)	$g$ (m/s <sup>2</sup> )
1.2	$0.9982 \times 10^3$	$1.98 \times 10^{-5}$	$8.9 \times 10^{-4}$	0.0728	$3 \times 10^{-3}$	9.81

Table 2: Parameter values used in the simulation of a falling water droplet in air.

In order to properly reproduce and analyse the dynamics of a relatively large drop (high Reynolds flow) such as in our case, the numerical method has to accurately resolve the thin boundary layers, the interaction of such layers with the capillary forces and finally the non-linear feedback of the complex 3D vortical structures present in the wake behind the droplet. Therefore, our objective behind the demonstration of this particular test case is not to develop a high fidelity model of a raindrop, but rather carry out a stringent evaluation of the robustness of our numerical method compared to methods which are not momentum consistent. For such a low Weber number the capillary forces dominate and the droplet should remain intact, and definitely not undergo any subsequent atomization.

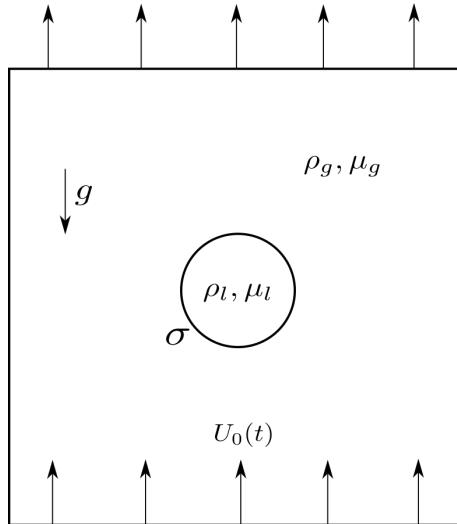


Figure 12: A 2D schematic of the numerical setup for the falling raindrop. We apply a uniform inflow velocity condition with  $U_0(t)$  and an outflow velocity condition at the top which corresponds to zero gradient of the velocity at the boundary. Boundary conditions on the side walls correspond to those of free slip (no shear stress).

Numerical simulations of this test case at moderate resolution (from  $D/h = 16$  to  $D/h = 64$ ) were tested *without* the consistent momentum-conserving scheme described in this paper, results in the catastrophic deformations of the droplet as shown in Fig. 13, which we describe as 'fictitious' or 'artificial' atomization as a result of rapidly growing numerical instabilities. This effect was previously uncovered by [?] in a similar case involving the sudden interaction of a droplet at rest with a uniform gas flow. The Weber number in that case was also approximately 3, therefore one would expect a nearly spherical droplet shape, but the authors of ref. [?] have reported a similar catastrophic deformation. We propose the following explanation in order to account for such numerical artifacts. To start with, we neglect gravity and viscous effects at this relatively large Reynolds number. Also, we are interested in steady-state



Figure 13: Rain drop test case: catastrophic breakup with non-conserving formulation,  $D/h = 30$ .

flow. On the axis and near the hyperbolic stagnation point at the front of the droplet one has  $u_2 = 0$  for the transverse (radial) velocity and for the axial momentum balance

$$u_1 \partial_1 u_1 = -\frac{1}{\rho} \partial_1 p. \quad (81)$$

Due to the large viscosity and density ratios, it is not possible for the air flow to immediately entrain the water, so the fluid velocity is significantly smaller inside the droplet. In the air the acceleration near the stagnation point is of the order  $U^2/D$ , whereas the pressure gradient is

$$\partial_1 p \sim \rho_a U^2 / D. \quad (82)$$

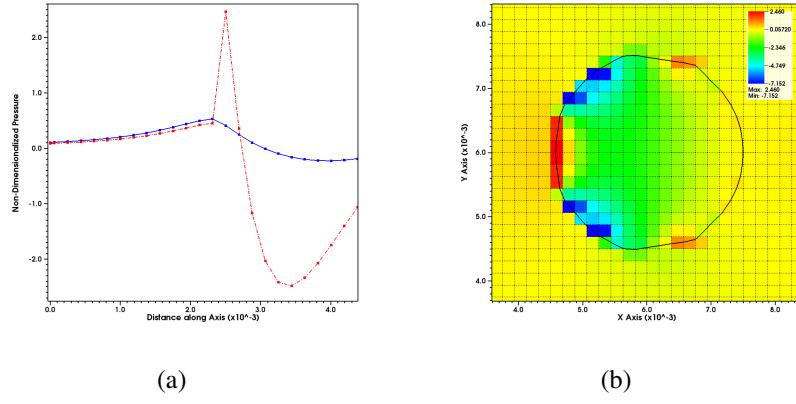
The pressure gradient in the water is much smaller, however in the case of a mixed cell the water density multiplies the air acceleration  $U^2/D$ , so that

$$\partial_1 p \sim \rho_w U^2 / D, \quad (83)$$

then a large pressure gradient results in the mixed cell or cells. This large pressure gradient results in a large pressure inside the droplet near the front stagnation point, as shown in Figure 14. This large pressure is balanced by surface tension only for a sufficiently large curvature near the droplet front. This explains the presence of a “dimple” often observed in low resolution simulations of the falling drop.

Visualization of the flow around the droplets (Figure 16) illustrates the challenging nature of the flow configuration, even for such a seemingly simple physical problem. As one can observe, the boundary layers are extremely thin. Therefore, even though the velocity field is continuous across the interface (in the discrete sense) in the absence of mass transfer, there is the appearance of strong velocity variations at the scale of the grid size for such coarse levels of grid refinement. We observe that applying the numerical method described in this paper brings a considerable and systematic improvement over a range of different flux limiters (WENO, ENO, Superbee, QUICK, Verstappen) and CFL numbers, as evidenced by comparing the figures 14 and 15.

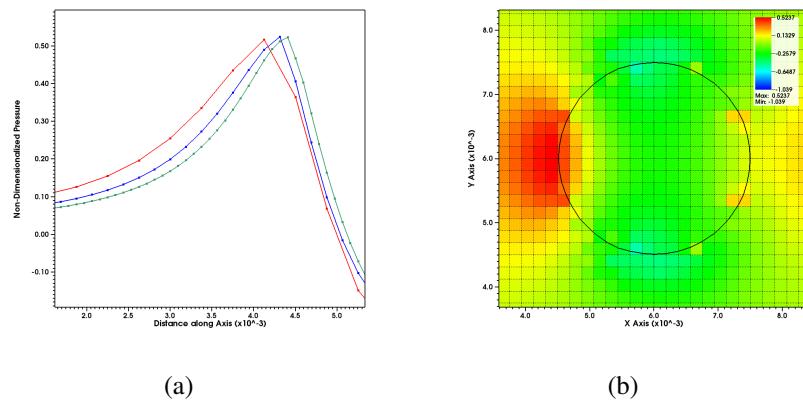
To summarize, the simulations broadly fall into three categories. The first two categories concern those that use the present momentum conserving method, and as a result maintain physical values of the kinetic energy and smooth interfacial shapes. Certain simulations amongst the ones carried out with the present method display some numerical instabilities (but not leading to catastrophic deformation) as a function of the



(a)

(b)

Figure 14: The origin of the pressure peak in the front of the droplet. (a) The profile of the pressure on the axis a few timesteps after initialisation with the standard, non-momentum conserving method (red) and the present method (blue). (b) The pressure distribution immediately after the start of the simulation using the standard, non-momentum-conserving method. The pressure peak does not result yet in the formation of a dimple. In all figures  $D/h = 16$ .



(a)

(b)

Figure 15: (a) The pressure profile on the axis a few timesteps after initialisation with the present method at various resolutions:  $D/h = 8$  (red), 16 (blue) and 32 (green). (b) The pressure distribution immediately after the start of the simulation using the present method and  $D/h = 16$ .

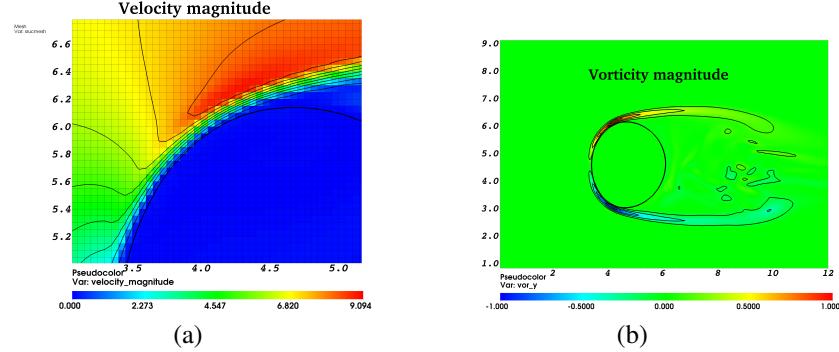


Figure 16: Flow field around the 3 mm droplet with 60 grid points per diameter. (a) The velocity magnitude. It is seen that even at this highest resolution there are only three points in the boundary layer. (b) The vorticity magnitude. The marked separation of the boundary layers is observed with a more complex vortical region in the wake.

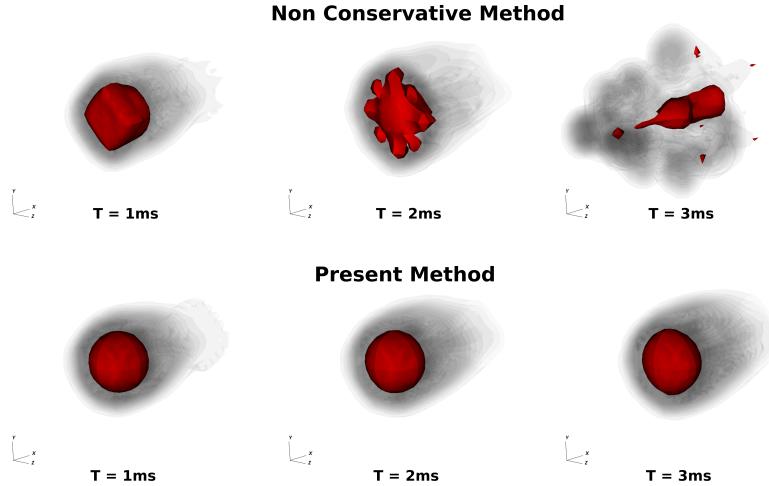


Figure 17: A comparison of the temporal evolution between the non conservative version and the present method, using a combination of the WY advection scheme with the QUICK flux limiter. The flow is along the positive X direction, with gravity along the opposite direction. The red contour indicates the isosurface of the volume fraction field corresponding to a value of 0.5, whereas the black contours surrounding the drop represent isosurfaces of the magnitude of vorticity. The raindrop in the non conservative version displays massive deformations leading to artificial breakup due to rapidly growing numerical instabilities. The droplet resolution for both methods is  $D/h = 8$ .

choice of advection scheme (CIAM/WY) and flux limiter (QUICK/WENO/ENO/Superbee/Verstappen) combination. Finally, there are simulations carried out without the present method that rapidly blow up, displaying marked peaks or spikes in kinetic energy as a function of time, associated with massively deformed interface shapes (see figure 17). Therefore, our studies suggest that certain combinations of the advection

scheme and the flux limiter are more numerically robust than others, in particular, the most stable combinations are that of CIAM advection with Superbee limiter and the WY advection with QUICK limiter. As a minor remark, the CIAM advection with the Superbee limiter appear to be quite diffusive.

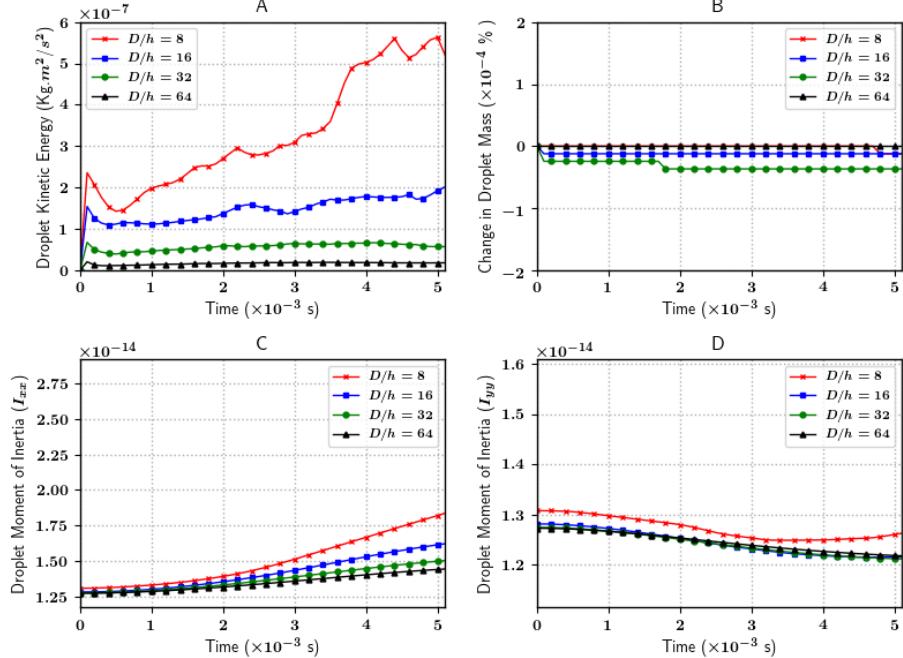


Figure 18: Temporal evolution of quantities of interest to evaluate the performance of our present method for different spatial resolutions. (A) Kinetic energy of the droplet. (B) Percentage change in the droplet mass from initialized mass. (C) Moment of inertia of the droplet along flow (X) direction. (D) Moment of inertia of the droplet along direction perpendicular to flow (Y,Z), evolution of  $I_{yy}$  is identical to  $I_{zz}$ , thus the latter is not shown.

We illustrate the performance of the present method through the results of our simulations in Figure 18. The spatial resolutions correspond to  $D/h = 8, 16, 32$  and  $64$  using the WY advection scheme in combination with the QUICK flux limiter, while keeping the same value for the inflow velocity boundary condition. The quantities of interest while examining the robustness of the method are the temporal evolution of the droplet kinetic energy (Fig. 18. A) and droplet mass (Fig. 18. B), as well as moment of inertia of the droplet along directions aligned to the inflow velocity (Fig. 18. C) and orthogonal (Fig. 18. D) to it. The moment of inertia is used as a descriptor of the ‘average’ droplet shape, with the three moments of inertia along the different axes  $I_m$  defined as -

$$I_m = \int_{\mathcal{D}} H x_m^2 d\mathbf{x}, \quad 1 \leq m \leq 3, \quad (84)$$

where  $\mathcal{D}$  is the computational domain and  $x_m$  is the distance of the interface relative to the center of mass of the droplet.

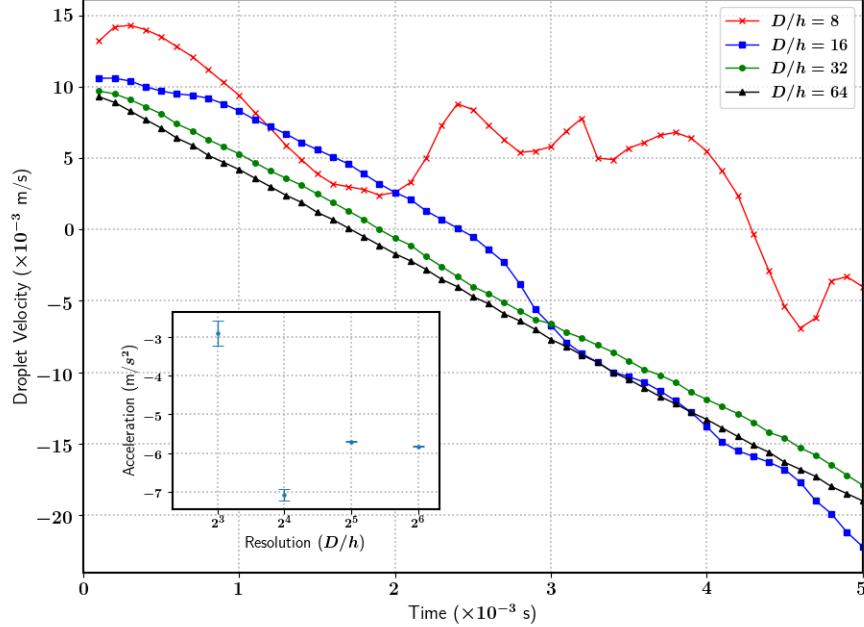


Figure 19: Comparison of droplet velocity as a function of time, for different droplet resolutions. The droplet velocities correspond to that of their respective center of masses. Inset : Convergence of the droplet acceleration as a function of resolution, computed using the best linear fit over the temporal variation of their respective velocities. The error bars signify the asymptotic standard error (least-squares) corresponding to the obtained linear fits.

Resolution ( $D/h$ )	8	16	32	64
Acceleration ( $m/s^2$ )	$-2.89 \pm 11.09\%$	$-7.08 \pm 1.97\%$	$-5.70 \pm 0.42\%$	$-5.82 \pm 0.12\%$

Table 3: Droplet acceleration in the frame of reference of the static box. The reduction in the error as we increase spatial resolution follows a rate of convergence roughly equal to 2.5.

The kinetic energy of the droplet evolves in a relatively smooth manner, without the presence of sudden spikes and falls which are emblematic of the non-conservative version of our method. Such abrupt changes in kinetic energy of the droplet have been found to be associated with instants when the droplet undergoes 'artificial' atomization or breakup, henceforth resulting in the catastrophic loss of stability for the numerical method. We observe a systematic drop in the amount of the droplet kinetic energy as we increase resolution, with the most probable explanation being that of the suppression of spurious interfacial oscillations which are rampant at low resolutions. There is also a component of the kinetic energy of the droplet associated with the internal coherent vortical structures generated due to the interaction of aerodynamic shear at

the interface, evidenced by the non-zero value of the kinetic energy even for the most highly resolved droplets. In terms of mass conservation, our method performs well, with the fractional loss of mass bounded within 0.0001% for all droplet resolutions tested. Finally, the moment of inertia of the droplet appears to evolve in a smooth manner for all droplet resolutions, with higher resolutions exhibiting lower amounts of inertia as a result of the more compact shapes obtained once the spurious interfacial deformation modes are subdued.

We have to keep in mind that the velocity inflow condition does not correspond to the 'exact' terminal velocity field an actual falling raindrop might experience, hence the droplet in our numerical setup does have some finite acceleration due to the imbalance between the aerodynamic forces and gravity. In Figure 19, we demonstrate the velocity of the center of mass of the droplet (in the frame of reference of the box enclosing the droplet) as a function of time, and its behavior as we increase the droplet resolution. As one can observe, due to the imbalance of forces acting on the droplet, it undergoes a net acceleration as evidenced by the approximately linear increase (absolute value) in the velocity as a function of time. The temporal variation in the droplet velocity is fitted to a linear polynomial in order to evaluate the droplet acceleration by means of a standard least-squares approach, for each droplet resolution. We illustrate (inset Figure 19) that we achieve more accurate fits as a consequence of higher droplet resolutions, as evidenced by decrease in the standard error on the fits ranging from  $\pm 11.1\%$  for  $D/h = 8$  to a value of  $\pm 0.12\%$  for the highest resolution of  $D/h = 64$ . As per the values given in Table 3, the reduction in the error as a function of the resolution corresponds to an order of convergence roughly equal to 2.5. The results obtained in this section indicate that especially when it comes to low resolutions, our numerical method can be used to get relatively good estimates of the underlying flow features of the droplet without observing an un-physical evolution as a consequence of discretization errors.

#### 4.6. Atomizing air and water planar jets

We also test the capability of the VOF-consistent momentum-conserving scheme to simulate complex air-water flows with an unstable shear flow. For that purpose we repeat the setup of reference [?]. Two jets of air and water are entering the computational domain from the left of Figure 20 at velocities comparable to those of experiments. However in order to save computational time the domain is smaller than in experiments. Physical properties of air and water are identical to those of the falling raindrop case given in Table 2. The flow and domain characteristics are given in Table 4 including a gas boundary layer and separator-plate size identical to those of reference [?]. The notations are as in reference [?]:  $H_p$  is the thickness of the jet of phase  $p$ , there is a separator plate of thickness  $l_y$  and a gas boundary layer of thickness  $\delta_g$ . The two streams have equal thickness  $H_l = H_g$ . The dimensionless parameters are given in Table 5. The CIAM advection method has been used. The number of grid points in the layer  $H_l/h = 16$  is relatively small (compared to the  $H_l/h = 32$  in the smallest (coarsest) simulation of reference [?]). It is thus all the more remarkable that the simulation is stable since using a smaller number of grid points usually increases the trend towards instability. It is interesting to note that the VOF calculations accounts for 31.5% of the

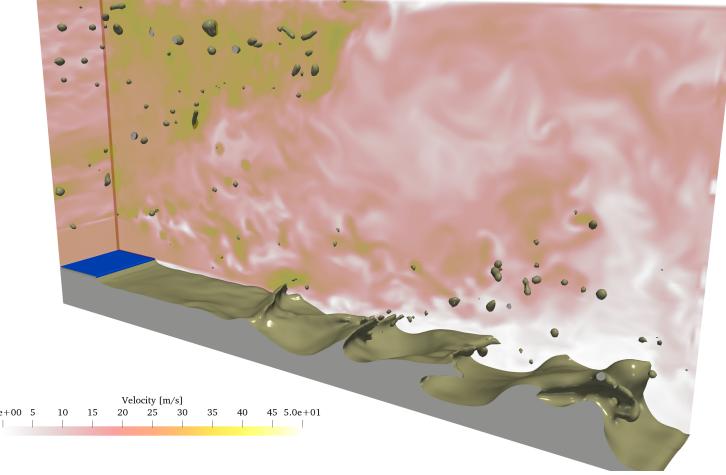


Figure 20: Atomizing layer with air/water properties.

$U_l$ (m/s)	$U_g$ (m/s)	$H_l$ (m)	$h$ (m)	$l_y$ (m)	$\delta_g/l_y$ (-)
1	25	$4 \cdot 10^{-3}$	$2.5 \cdot 10^{-4}$	$2.5 \cdot 10^{-4}$	2

Table 4: Physical parameters (defined in the text) for the atomizing layer setup. The fluid properties are the same as in Table 2.

total time, while the inversion of the Poisson operator for the pressure accounted for 51.5%. The whole simulations runs overnight on a present-day workstation.

## 5. Conclusion

We have presented and tested a simulation method for multiphase flow that shows increased stability at large density contrasts and large Reynolds numbers. The method is closely related to some variants of the VOF method including an implementation of the WY advection method. The increased stability is evidenced in the test case of a single 3-mm droplet of water falling in air, a typical raindrop. It is a reflection on the challenging nature of multiphase flow that such complex methods apparently need to

$M$	$r$	$m$	$\text{Re}_{g,\delta}$	$\text{We}_{g,\delta}$	$\text{Re}_g$
$\rho_g U_g^2 / (\rho_l U_l^2)$	$\rho_l/\rho_g$	$\mu_l/\mu_g$	$\rho_g U_g H_g / \mu_g$	$\rho_l U_l H_l / \mu_l$	$\rho_g U_g^2 H_g / \sigma$

Table 5: Dimensionless parameters for the atomizing layer setup.

be implemented to resolve such an everyday and simple phenomenon.

The method comes with a significant saving of computer time, since for similar problems with raindrops, our attempts with a non-momentum-conserving VOF approach led to catastrophic deformation of the drop or strong dimple formation. These problems have also been observed by us using other non-VOF-consistent and non-momentum-conserving methods such as the one of [? ]. In that case whenever less than 200 grid points per diameter are used numerically stable air-water drops accelerated at moderate Weber number cannot be found. However for higher resolutions they can be computed without difficulty as also found by the authors of ref. [? ]. Here, approximate solutions accurate within 15% are found with only 15 points per diameter, and non-divergent computations are found with as little as 2 points per diameter.

A particular advantage of the method is that it is conserving mass at the accuracy at which discrete incompressibility is enforced and opens a perspective for similar momentum conservation using WY advection. The method nevertheless is more complex and costly than a colocated method. This opens the perspective for systematic development of other methods with different grid arrangements. Another perspective is the potential of stable methods with large density contrasts, exact mass and momentum conservation and small droplets, that could be smoothly merged into models that represent the small droplets as Lagrangian Point Particles [? ].

## 6. Acknowledgements

This work has been supported by the ANR MODEMI project (ANR-11-MONU-0011) program and grant SU-17-R-PER-26-MULTIBRANCH from Sorbonne Université. Philipp Yecko acknowledges support from grants NSF-1362823 and NSF-1620158.

This work was granted access to the HPC resources of TGCC-CURIE, TGCC-IRENE and CINES-Occigen under the allocations t20152b7325, t20162b7760, 2017tgcc0080 and A0032B07760, made by GENCI and TGCC. The authors would also acknowledge the MESU computing facilities of Sorbonne Université.

We would thank Dr. W. Aniszewski, Dr S. Dabiri, Dr Jiacai Lu and Dr. P. Yecko for their contribution to the development of the code *PARIS-Simulator*, and we thank Dr. W. Aniszewski, Dr. V. Le Chenadec, Mr. C. Pairetti, Dr. S. Popinet and Dr. S. Vincent for useful conversations on the topics of this paper.

Finally, the simulation data are visualized by the software VisIt developed by the Lawrence Livermore National Laboratory.

## References

### Appendix A. Kelvin Helmholtz instability: numerical setup

We consider the 2D base flow shown on Figure A.22. Coordinates are noted  $(x, z)$  and vectors  $(u, w)$ . The height of the interface is  $h(x, t)$ . The flow has density  $\rho_1$  for  $z < h(x, t)$  and  $\rho_2$  for  $z > h(x, t)$  and is incompressible. The base flow is a parallel shear flow. The base flow is uniform with  $u = -U$  for  $z < 0$  and  $u = U$  for  $z > 2a$ , with a linear (Couette flow) boundary layer in between. When  $\rho_1 \neq \rho_2$  the heavier fluid is the

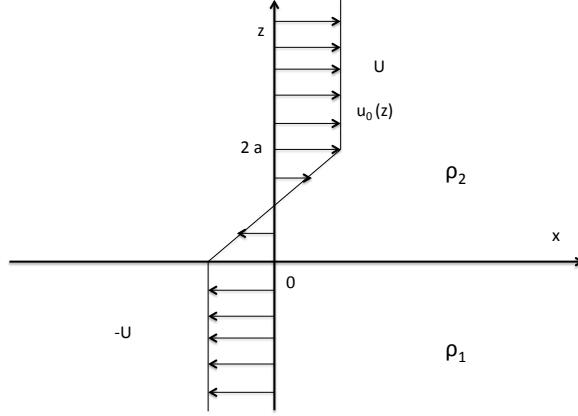


Figure A.21: Base velocity profile.

“liquid” and the lighter the “gas”. For  $\rho_1 > \rho_2$  the boundary layer is in the gas while otherwise the boundary layer is in the liquid. Similar flows have been studied in [? ?].

#### Appendix A.1. Dispersion relation

The Euler and incompressibility equations are as in Eqs. (4) with only  $\mathcal{L} = \mathcal{L}_{\text{conv}}$ , surface tension and viscosity are not included. The interface height  $h$  moves according to

$$\partial_t h + u \partial_x h = w \quad (\text{A.1})$$

We also use a stream function  $\psi$

$$w = -\partial_x \psi, \quad u = \partial_y \psi \quad (\text{A.2})$$

We consider a small perturbation of the base velocity in the form

$$\mathbf{u} = \mathbf{u}_0 + \epsilon \mathbf{u}_1 + O(\epsilon^2) \quad (\text{A.3})$$

the pressure expands as  $p = p_0 + \epsilon p_1(x, z, t) + O(\epsilon^2)$ , the height as  $h = \epsilon h_1(x, z, t) + O(\epsilon^2)$ , and similarly the stream function. We assume the following form for the perturbation

$$\begin{pmatrix} u_1 \\ w_1 \\ p_1 \\ \psi_1 \\ h_1 \end{pmatrix} = \begin{pmatrix} U_1(z) \\ W_1(z) \\ P_1(z) \\ \Psi_1(z) \\ A_h \end{pmatrix} e^{-ikx-i\omega t} \quad (\text{A.4})$$

where  $k$  is an arbitrary wavenumber and  $\omega$  a frequency to be determined. Although the expressions on the rhs in (A.4) are complex we understand the real part.

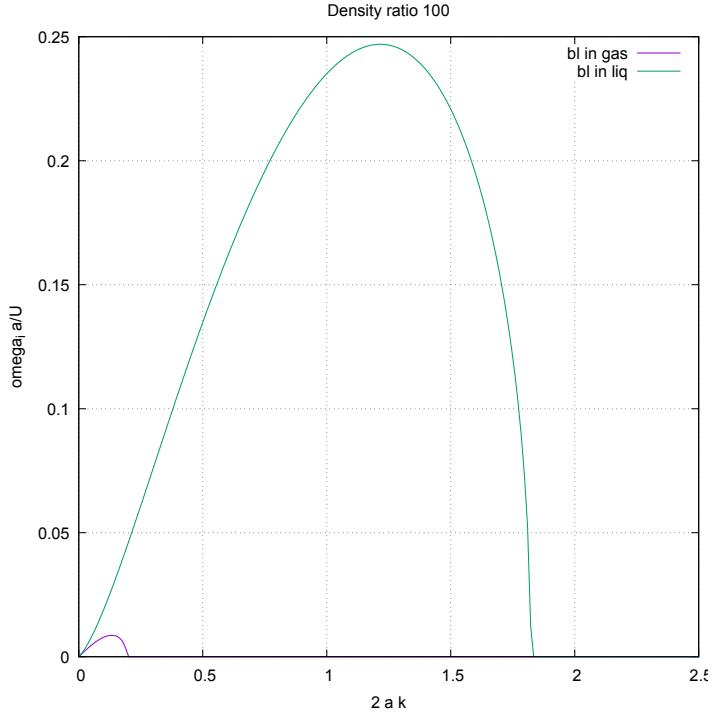


Figure A.22: Growth rate for  $r = 100$ .

It is convenient to define as Chandrasekhar [? ] the reduced wavenumber  $\kappa = 2ka$  and the reduced frequency  $\Omega = \omega a/U$ , which are related by

$$e^{-2\kappa} = (1 - 2\Omega - \kappa) \frac{2 + (r + 1)(2\Omega - \kappa)}{2 + (r - 1)(2\Omega - \kappa)}, \quad (\text{A.5})$$

see for example [? ], equation (135). The system is unstable whenever Eq. (A.5) has two complex conjugate non-real roots. Then the positive imaginary part  $\Omega_i$  is the growth rate, plotted on Figure ?? in the case  $r = 100$  where  $r = \max(\rho_1/\rho_2, \rho_2/\rho_1)$ . The case  $r = 100$  corresponds to the boundary layer in the gas and is much less unstable than the case with the boundary layer in the liquid. As the ratio  $r$  is increased the growth rate for the boundary layer in the gas decreases steadily while the growth rate for the boundary layer in the liquid converges to the one for a free surface, with the gas replaced by a void.

#### Appendix A.2. Special case

When  $a = 0$  the above is singular and a special computation is needed. One has  $\kappa = 0$  and the frequencies are

$$\omega = \left( \frac{r - 1}{r + 1} + \frac{2\sqrt{r}}{r + 1} i \right) kU \quad (\text{A.6})$$

### Appendix A.3. Construction of the unstable mode

We want to write the solution using the stream function  $\psi$  in order to have a divergence free initial condition in the computations. The construction of  $\psi$  is done by determining the relevant coefficients so that  $\psi$  is given by

$$z < 0 \quad \psi(x, z, t) = A_1 e^{kz} e^{-ikx-i\omega t} \quad (\text{A.7})$$

$$0 < z < 2a \quad \psi(x, z, t) = (A_0 e^{kz} + B_0 e^{-kz}) e^{-ikx-i\omega t} \quad (\text{A.8})$$

$$2a < z \quad \psi(x, z, t) = B_2 e^{-kz} e^{-ikx-i\omega t} \quad (\text{A.9})$$

From (A.1)

$$\partial_t h = -U_0(0)\partial_x h + w \quad (\text{A.10})$$

and from (A.2) and (A.8) we get

$$-i\omega A_h = -ikUA_h + ik(A_0 + B_0)$$

hence

$$(-i\omega + ikU)A_h = ik(A_0 + B_0)$$

and

$$A_h = -\frac{k}{\omega - kU}(A_0 + B_0)$$

Similar relations are obtained from the requirements of continuity of pressure and normal velocity at  $z = 0$  and  $z = 2a$ . We can then determine all the amplitudes of the constructed solution after an amplitude for the interface has been chosen. Typically the modulus  $|A_h|$  and the argument  $\phi$  are selected so that

$$A_h = |A_h|e^{i\phi}$$

then for  $\kappa > 0$  we have

$$C_0 = -\left(\frac{\omega}{k} - U\right)A_h(e^{-2\kappa} + 2\Omega + \kappa - 1)^{-1} \quad (\text{A.11})$$

$$A_0 = C_0 e^{-2\kappa} \quad (\text{A.12})$$

$$B_0 = C_0(2\Omega + \kappa - 1) \quad (\text{A.13})$$

$$A_1 = A_0 + B_0 \quad (\text{A.14})$$

$$B_2 = A_0 e^{2\kappa} + B_0 \quad (\text{A.15})$$

These expressions for the amplitudes together with (A.7-A.9) are used to initialize the stream function. The intermediate constant  $C_0$  is used to simplify the expressions.

#### Appendix A.4. Special cases: mode structure for $ka \rightarrow 0$

##### Appendix A.4.1. Mode structure for $a > 0$ in the limit $a \rightarrow 0$ .

For small or vanishing  $a$  and  $\kappa$  however the expression (A.11) is singular. Indeed in the limit  $\kappa \rightarrow 0$  we also have  $\Omega \rightarrow 0$  and

$$e^{-2\kappa} + 2\Omega + \kappa - 1 = -2\kappa + 2\kappa^2 + \kappa + O(\kappa^3) = -\kappa + O(\kappa^2)$$

and then  $A_0$  and  $B_0$  become spurious as the region  $(0, 2a)$  vanishes. From (A.11)

$$C_0 \simeq -\frac{\omega}{k} \frac{1}{\kappa} A_h$$

and from (A.13) and (A.15)

$$B_2 = C_0(2\Omega + \kappa) \simeq C_0\kappa$$

Thus for  $a = 0$

$$\begin{aligned} B_2 &= -\frac{\omega}{k} A_h \\ A_1 &= \frac{\omega}{k} A_h \end{aligned}$$

the latter being obtained directly from (A.14). Since  $B_2 \neq A_1$  there is a  $O(\epsilon)$  discontinuity of  $\psi$  which results in a jump of  $v(x, z, t)$  across the interface at  $z = 0$  and a related thin jet  $u_1(x, z, t) \simeq \epsilon f(x, t)\delta(z)$ . This is a consequence of placing the interface in the above calculations at  $z = 0$  instead of  $z = \epsilon h_1(x, t)$  and it conflicts with the solution obtained classically with the thin vortex sheet setup (that is, the setup in which  $a = 0$  is postulated at the beginning).

#### Appendix A.5. Mode structure in the $a = 0$ case

We now obtain the mode structure for the classic thin vortex sheet setup. In that case we keep only the terms in (??) and (??). The velocity continuity condition becomes

$$h_t = -uh_x + w = -u_0h_x + w_1 + O(\epsilon) \quad (\text{A.16})$$

which replaces (A.10). Since  $h_t$  must have the same expression above and below the interface

$$[w - h_x u_0] = 0 \quad (\text{A.17})$$

thus  $[w] = h_x[u_0] = 2h_xU$  and from (??) and (??)

$$z < 0 \quad W_1 = A'_1 e^{kz} \quad (\text{A.18})$$

$$z > 0 \quad W_1 = B'_2 e^{-kz} \quad (\text{A.19})$$

and from (A.16)

$$A'_1 = -i(\omega - kU)A_h \quad (\text{A.20})$$

$$B'_2 = -i(\omega + kU)A_h \quad (\text{A.21})$$

$$(\text{A.22})$$

The pressure equality at  $z = 0$  leads to

$$z < 0 \quad P_1 = \frac{i\rho_1}{k^2}(\omega - kU)A'_1ke^{kz} \quad (\text{A.23})$$

$$z > 0 \quad P_1 = \frac{i\rho_2}{k^2}(\omega + kU)B'_2(-ke^{-kz}) \quad (\text{A.24})$$

hence introducing  $\nu = \omega/(kU)$  and from (A.20)and (A.21)

$$(\nu + 1)A_1 - (\nu - 1)B_2 = 0 \quad (\text{A.25})$$

$$r(\nu - 1)A_1 + (\nu + 1)B_2 = 0 \quad (\text{A.26})$$

$$(\text{A.27})$$

from which one obtains

$$\nu = \frac{r - 1}{r + 1} + \frac{2\sqrt{r}}{r + 1}\text{i} \quad (\text{A.28})$$

identical to (A.6). Also from (A.20)and (A.21)

$$A_1 = (U - \omega/k)A_h \quad (\text{A.29})$$

$$B_2 = -(U + \omega/k)A_h \quad (\text{A.30})$$

These expressions should be used whenever  $a \ll A_h$  while the full expressions with the boundary layer would be valid for  $A_h \ll a$ . In both cases  $\Delta x \ll \min(A_h, a)$  may be required.