ESAIM: M2AN
Vol. 40, N° 3, 2006, pp. 431–468
DOI: 10.1051/m2an:2006023

ESAIM: Mathematical Modelling and Numerical Analysis
www.edpsciences.org/m2an

NUMERICAL ANALYSIS OF EULERIAN MULTI-FLUID MODELS IN THE CONTEXT OF KINETIC FORMULATIONS FOR DILUTE EVAPORATING SPRAYS

Frédérique Laurent[1]

Abstract. The purpose of this article is the analysis and the development of Eulerian multi-fluid models to describe the evolution of the mass density of evaporating liquid sprays. First, the classical multi-fluid model developed in [Laurent and Massot, *Combust. Theor. Model.* **5** (2001) 537–572] is analyzed in the framework of an unsteady configuration without dynamical nor heating effects, where the evaporation process is isolated, since it is a key issue. The classical multi-fluid method consists then in a discretization of the droplet size variable into cells called sections. This analysis provides a justification of the "right" choice for this discretization to obtain a first order accurate and monotone scheme, with no restrictive CFL condition. This result leads to the development of a class of methods of arbitrary high order accuracy through the use of moments on the droplet surface in each section and a Godunov type method. Moreover, an extension of the two moments method is proposed which preserves the positivity and limits the total variation. Numerical results of the multi-fluid methods are compared to examine their capability to accurately describe the mass density in the spray with a small number of variables. This is shown to be a key point for the use of such methods in realistic flow configurations.

Mathematics Subject Classification. 35L05, 65L20, 76T10.

Received: April 6, 2005. Revised: October 4, 2005.

1. Introduction

In many industrial combustion applications such as Diesel engines, fuel is stored in condensed form and injected as a spray (spray is understood as a dispersed phase of liquid fuel droplets, *i.e.* where the liquid volume fraction is much smaller than one) carried by a gaseous flow. Two phase effects and more specifically the polydisperse character of the droplet size distribution can significantly influence flame structures, even in the case of relatively thin sprays [21]. The key point is then to precisely determine the amount of mass transferred from the liquid phase into the gaseous stream through the evaporation process.

Spray models have a common basis at what can be called the "kinetic level" under the form of a probability density function [4, 7, 15, 27–29]. So, assuming that the spray is constituted of spherical droplets characterized only by one geometry parameter ϕ, one velocity u_l and one temperature T_l, at low Weber number, the spray is characterized by a distribution function $f^\phi(t, x, \phi, u_l, T_l)$, so that $f^\phi \, \mathrm{d}x \, \mathrm{d}\phi \, \mathrm{d}u_l \, \mathrm{d}T_l$ is the probable number of droplets at time t, in the phase space elementary volume $\mathrm{d}x \, \mathrm{d}\phi \, \mathrm{d}u_l \, \mathrm{d}T_l$ around the point (x, ϕ, u_l, T_l).

Keywords and phrases. Spray, evaporation, multi-fluid method, kinetic schemes.

[1] EM2C, CNRS, École Centrale de Paris, Châtenay-Malabry, France. frederique.laurent@em2c.ecp.fr

The parameter ϕ denotes for example the volume V as in [9], the surface S as in [3] or the radius R ($f^R\,\mathrm{d}R = f^S\,\mathrm{d}S = f^V\,\mathrm{d}V$). The evolution of the spray is then described by the Williams transport equation [28]:

$$\partial_t f^\phi + \partial_x \cdot (u f^\phi) + \partial_\phi (R_\phi f^\phi) + \partial_u \cdot (F f^\phi) + \partial_T (E f^\phi) = \Gamma + Q \tag{1.1}$$

where R_ϕ denotes the rate of change of the droplet geometry due to evaporation, F, is the drag force due to the velocity difference with the gaseous phase, E, the rate of heat exchange between the two media, Γ, the coalescence and dissociation collision integral operator, and Q, the elastic collision integral operator. These quantities depend on the droplet size, velocity and temperature and on the local gas composition, velocity and temperature (and on f^ϕ for Γ and Q).

Several strategies may be used in order to solve the liquid phase. A first choice is to approximate the distribution function by a sample of discrete numerical parcels of particles of various sizes through a Lagrangian–Monte-Carlo approach [1, 7, 12, 22, 25]. This procedure has been widely used and has been shown to be efficient in a number of cases. Its main drawback, which has shown recently to be a major one with the development of new combustion chambers leading to combustion instabilities (lean premixed prevaporized combustor with spray injection), is the coupling of a Eulerian description for the gaseous phase to a Lagrangian description of the dispersed phase, thus offering limited possibilities of vectorization/parallelization and implicitation. Moreover for unsteady computations, a large number of parcels in each cell of the computational domain is generally needed, thus yielding large memory and CPU requirements.

This drawback makes the use of an Eulerian formulation for the description of the disperse phase attractive, at least as a complementary tool for Lagrangian solvers. Moments methods are thus used, since the high dimension of the phase space prevents a direct numerical integration of equation (1.1) with deterministic numerical methods like finite volumes. Two classes have been considered before, the first relies on "population balance" equations [23] derived for very small particles in the study of aerosols in chemical engineering; these equations may be integrated efficiently in order to follow the size distribution of particles, without inertia, experiencing some aggregation-breakage phenomena (quadrature of moment methods [19, 30]). However, the extension of such methods to sprays for which the inertia determines the dynamical behavior of the droplets has not yet received a satisfactory answer. The second type of method relies on moments in the velocity variable without dealing with the details of the size distribution or only considering one droplet size. The two-fluid type models for separated two-phase flows [2, 5, 10] belong to this class. The fact that no information is available on the droplet size distribution is generally too severe an assumption in most applications. One possibility is to use a semi-fluid representation based on velocity moment closure of the probability density function at sampled sizes [4]. However, one of the main drawback of most of the existing Eulerian models is the impossibility to treat droplet-droplet interactions because only a finite number of sizes are present in the problem. The use of moments methods leads to some information loss but the computational cost is usually much lower than that associated with Lagrangian calculations for two reasons, the first one is related to the fact that the polydisperse character of the sprays is not described by the model (the spray is mainly considered as mono-dispersed [24]) and the number of unknowns to be determined is very limited; the second one is related to the high level of optimization one can reach when the two phases are both described by an Eulerian model.

Recent Eulerian multi-fluid models [14–17, 20] are intermediate between a deterministic solution of the full spray equation (1.1) and the moment equation models, where some dimensions of the phase space are treated using moment equations and some others using a finite volume type discretization in such a way that enough precision is maintained at relatively low cost. More precisely, the kinetic equation (1.1) is averaged over velocity and temperature and yields a semi-kinetic model: at each droplet size one obtains a set of balance equations for the mean density, momentum and temperature. The continuous size phase space is then discretized using a finite volume formulation, the size intervals being called the sections in reference to Greenberg, Tambour *et al.* [8,9]. The conservation equations obtained in this view describe the evolution of mass density, momentum and temperature in each section. The derivation from the kinetic level was introduced in [15] for dilute sprays and extended to dense sprays where coalescence is taking place in [16]. The model has been validated for dilute sprays in the configuration of counterflow spray diffusion flames using comparisons with experimental measurements

[17, 21] and with the classical Lagrangian approach [15]. However, this method implies a projection step, associated with the choice of the distribution function profile in each section [3, 15] and for which no numerical analysis was available. Moreover, this choice, beyond its influence on the numerical analysis of the method, that is in the limit of well refined spatial meshes, has a strong impact if one wishes to reduce the number of sections and then, the computational cost [17]. As indicated in [15], numerical diffusion for the droplet size variable can reduce the precision of the method, especially for cases with non-smooth distribution. This lack of precision can be reinforced through a coupling with drag and coalescence effects [16].

The projection step and the numerical diffusion are completely stemming from the discretization of the droplet size variable in the evaporation process. For this reason, we focus in this study on a purely evaporating case, for which a rigorous numerical analysis is conducted. It is shown that the method is at most first order accurate, which sheds light on the observed numerical diffusion. The effect of the choice for the projection step on the convergence is specifically identified. But, as mentioned above, our purpose is to reduce the cost of the method in order to be able to tackle realistic configurations at a reasonable cost. We then introduce some higher order methods that improve the accuracy for regular distribution function, even if a small number of sections is used. A modification is proposed for the second order method in order to preserve the positivity of the variables and to have also a good accuracy for discontinuous function.

So, as in [3, 13, 15], the considered framework is an unsteady configuration of a motionless polydisperse spray evaporating in a hot gas at rest, without heating effects. For this configuration, the space phase can be reduced to ϕ: $f^\phi(t,\phi)$ is the droplet density at time t and as a function of the droplet size ϕ. Moreover, according to the d^2 law for the droplet evaporation, when the velocity and the composition of the gaseous phase are constant, the droplet surface area decreases linearly (*i.e.* R_S is constant). If ϕ is the droplet surface area S scaled by a maximum surface S_m and if the time t is scaled by S_m/R_S, the equation (1.1) reduces to:

$$\partial_t f^S - \partial_S f^S = 0, \qquad f^S(t=0,S) = f_0(S), \tag{1.2}$$

where $f_0(S)$ is the initial condition and S designates the non-dimensional droplet surface area. It can be noticed that the semi-kinetic system obtained by taking some moments in velocity and temperature of (1.1) reduces here to equation (1.2) [15]. The classical multi-fluid approach can be interpreted as an upwind finite volume discretization of this equation. This transport equation has been widely studied and numerical methods are available (see for example [18] and references therein). However, for the applications under consideration (*i.e.* for spray combustion), one is less interested in the function f^S than in its moment of order 3/2 which represents the mass density m of the spray, generally constituted of fuel. This is an important variable to predict the fuel gaseous mass fraction and thus, the flame structure. The non-dimensional mass density is defined by:

$$m(t) = \int_0^{+\infty} V(\phi) f^\phi(t,\phi)\,\mathrm{d}\phi, \tag{1.3}$$

where $V(\phi)$ is the non-dimensional volume corresponding to ϕ (in the same way, let $S(\phi)$ and $R(\phi)$ denote the non-dimensional surface and radius corresponding to ϕ).

It is also useful to introduce additional high order methods of moment type: a few surface moments (instead of one moment for the classical multi-fluid method) to describe the spray in each section. Compared to the classical multi-fluid method which only uses a single moment, a better approximation of the distribution in each section can be constructed by the use of polynomial functions which preserve the values of the moments and can be interpreted as the ones obtained with the least square method. The mass densities are then evaluated from this approximation. We first evaluate the order of accuracy of the different schemes in the case where the number of section tends to infinity. We next investigate the ability of the two or three moments, to describe the evolution of the mass density on a coarse discretization.

This article is organized as follows. First, the classical multi-fluid method is analyzed: the scheme is given for our particular case and for a general discretization type; results of convergence are proved and, in the case of an explicit Euler scheme for the time discretization, the corresponding CFL conditions for stability and positivity

of the droplet mass densities are provided. Higher order multi-fluid methods are introduced in a second step. Their consistency is proved and a condition for their stability is provided for the two and three moments cases. The particular case of the two moments method is examined, and an extension is proposed which preserves the positivity. Finally, these methods are numerically compared with a classical MUSCL method [26] for three types of initial distributions f_0: one regular, one singular and one regular but with a high Lipschitz constant. The computational efficiency is evaluated for a high level of refinement and shown to agree with the numerical analysis. Efficiency for really coarse discretizations is also examined, which will be a key point for the use of such methods in realistic configurations.

2. Numerical analysis of the classical Eulerian multi-fluid approach

The purpose of this section is the numerical analysis of the classical multi-fluid method, developed in [15], in the particular configuration described above. This method consists in a discretization of the droplet size variable in sections and provides a direct approximation of droplet mass density in each section.

The corresponding scheme is first detailed in the case where the initial distribution function has a compact support; the parameter characterizing the discretization are defined; a synthesis of convergence results as well as a discussion of the 'good way' to discretize the phase space are provided. The proof of these results are then set out: the global and local orders of the semi-discretized scheme and, in the case of an explicit Euler scheme for the time discretization, the conditions in the discretization step for stability and positivity of the scheme. Finally, an extension of the method is given for the case where the support of f_0 is not compact, as suggested in [17].

2.1. **A synthesis of the numerical analysis for the classical multi-fluid method**

2.1.1. *Presentation of the method*

The classical multi-fluid approach can be seen as an upwind finite volume discretization of the equation on the mass distribution $f_V^\phi = V(\phi)f^\phi$. This equation is obtained by multiplying (1.2) by the derivative of the droplet surface $S'(\phi) = \mathrm{d}S/\,\mathrm{d}\phi$ to obtain an equation on f^ϕ and by the droplet volume $V(\phi)$:

$$\partial_t(f_V^\phi) = \partial_\phi\left(\frac{f_V^\phi}{S'(\phi)}\right) - \frac{V'(\phi)}{V(\phi)S'(\phi)}f_V^\phi. \tag{2.1}$$

We can remark that, if ϕ is a power function of the non-dimensional droplet radius (ϕ is usually the droplet radius, surface or volume), then a singularity around $\phi = 0$ appears in the second term which can then be written: $\frac{3f_V^\phi}{2S(\phi)}$. For this reason, the classical finite volume methods cannot be used to integrate this equation and a specific method, the classical multi-fluid method, was developed.

Let us assume that $f_0(S)$, and then $f^\phi(0,\phi)$, has a compact support: $[0,1]$ for the non-dimensional variables. Let us then introduce the discretization $0 = \phi_0 < \phi_1 < \cdots < \phi_N = 1$ of this support. The classical multi-fluid approach gives an approximation $\tilde{m}_i(t)$ of the mass density $m_i(t) = \int_{\phi_{i-1}}^{\phi_i} V(\phi)f^\phi(t,\phi)\,\mathrm{d}\phi$ in the section i by the scheme [15]:

$$\partial_t\tilde{m}_i(t) = F_{i+1}\tilde{m}_{i+1}(t) - (F_i + M_i)\tilde{m}_i(t), \qquad \tilde{m}_i(0) = m_i(0). \tag{2.2}$$

The term $M_i\tilde{m}_i$ represents the mass flux between the section i and the gas; the term $F_i\tilde{m}_i$ represents the mass flux between the section i and the section $i-1$. These coefficients M_i and F_i are determined by a preliminary choice of the profile κ_i^ϕ of the distribution in each section: the function $f^\phi(t,\phi)$ is approximated by $f_a(t,\phi) = \kappa_i^\phi(\phi)\tilde{m}_i(t)$ for each $\phi \in [\phi_{i-1},\phi_i[$. We then have for $i \in \{1,\ldots,N\}$:

$$F_i = \frac{V(\phi_{i-1})}{S'(\phi_{i-1})}\kappa_i^\phi(\phi_{i-1}), \qquad M_i = \int_{\phi_{i-1}}^{\phi_i}\frac{V'(\phi)}{S'(\phi)}\kappa_i^\phi(\phi)\,\mathrm{d}\phi,$$

and $F_{N+1} = 0$. Due to the use of the $\tilde{m}_i(t)$ in the approximated function f_a, the κ_i^ϕ are such that:

$$\int_{\phi_{i-1}}^{\phi_i} V(\phi)\kappa_i^\phi(\phi)\,\mathrm{d}\phi = 1.$$

One then has to specify the discretization and the choice of the distribution profiles κ_i^ϕ in each section. It is natural to choose κ_i^ϕ as a constant function, but for a suitable variable ϕ. Since $\kappa_i^R dR = \kappa_i^S dS = \kappa_i^\phi d\phi$, the choice of ϕ is important, as emphasized in [3] and also in [15] and, thanks to simulations, an optimization of the choice of the profiles and the discretization in [17] in the case of a small number of sections and for experimental distribution functions. Here, the numerical analysis will be used to conclude on the asymptotic efficiency of the choice of profiles and discretization. It is convenient to focus on the general case $\phi = R^\beta$ with $\beta > 0$, so that the non-dimensional surface and volume of the droplets are $S(\phi) = \phi^{2/\beta}$ and $V(\phi) = \phi^{3/\beta}$. Moreover, a uniform discretization for a variable of the general form R^α, with $\alpha > 0$ is chosen:

$$\phi_i = (i\Delta\Psi)^{\beta/\alpha}, \qquad \Delta\Psi = 1/N.$$

The constants κ_i^ϕ and the coefficients F_i and M_i are then, for $i \in \{1, \ldots, N\}$:

$$\kappa_i^\phi = \frac{1}{\int_{\phi_{i-1}}^{\phi_i} \phi^{3/\beta}\,\mathrm{d}\phi}, \qquad F_i = \frac{\beta}{2}\frac{\phi_{i-1}^{1+1/\beta}}{\int_{\phi_{i-1}}^{\phi_i} \phi^{3/\beta}\,\mathrm{d}\phi}, \qquad M_i = \frac{3}{2}\frac{\int_{\phi_{i-1}}^{\phi_i} \phi^{1/\beta}\,\mathrm{d}\phi}{\int_{\phi_{i-1}}^{\phi_i} \phi^{3/\beta}\,\mathrm{d}\phi}. \tag{2.3}$$

The time discretization $t_k = k\Delta t$ is finally introduced, with $\Delta t > 0$ and $k \in \mathbb{N}$. The explicit Euler scheme corresponding to (2.2) provides an approximation $\hat{m}_i^k$ of $m_i^k = m_i(t_k)$ by:

$$\hat{m}_i^{k+1} = \hat{m}_i^k + \Delta t\, F_{i+1}\hat{m}_{i+1}^k - \Delta t\,(F_i + M_i)\hat{m}_i^k, \qquad \hat{m}_i^0 = m_i^0. \tag{2.4}$$

Results with the more general θ-schemes for the time discretization are given in [13]. However, the influence of the parameter θ on the CFL-like stability condition is a classical result which is not essential in this study and is then not detailed here.

2.1.2. *Synthesis of the results*

In the general case described above, some results will be proved depending on the value of α and β and then on the type of discretization. They first concern the semi-discretized scheme (2.2) for which consistency and convergence conditions on α and β are presented. Then, the time discretization leads to the introduction of some CFL conditions depending on α, so that the scheme is stable and the variables remain positive. These results allow to emphasize "optimal" choices for the discretization.

To evaluate the order of the semi-discretized scheme (2.2), the local truncation errors γ_i are introduced:

$$\gamma_i(t) = \partial_t m_i(t) + (F_i + M_i)m_i(t) - F_{i+1}m_{i+1}(t), \tag{2.5}$$

where m_i is the exact mass density of the section. As usual, the consistency of the scheme is proved in a "regular case", thanks to an expansion. It could seems natural that this regularity concerns the functions f^ϕ, as in [13]. However, a remark in this article has already emphasized that this regularity of f^ϕ cannot be ensured from initial conditions for every choice of ϕ. Indeed, the exact solution of equation (1.2) gives the following function f^ϕ:

$$f^\phi(t,\phi) = \frac{2}{\beta}\phi^{2/\beta-1} f_0(t + \phi^{2/\beta}). \tag{2.6}$$

In term of the initial distribution $f^\phi(0,.)$, it can be written: $\phi^{2/\beta-1}(t+\phi)^{\beta/2-1} f^\phi(0,(t+\phi^{2/\beta})^{\beta/2})$. This shows that for $\beta > 2$, a singularity appears, at $\phi = 0$ and for $t > 0$, even if the initial distribution is regular. Here, an

assumption independent of β is chosen: f_0 is supposed regular. Thus, results differ a little from those in [13]. The first one is the following local order theorem.

Theorem 2.1. *Let assume that $f_0(S)$ is a C^3 function for $S \in \mathbb{R}^+$. Then the semi-discretized scheme (2.2) is locally second order accurate in $\Delta\Psi$ if and only if $\beta \neq 2$ and $\alpha \leq 3/2$ or if $\beta = 2$ and $\alpha \leq 5/2$.*

Moreover, for these cases, the convergence of the variables $\tilde{m}_i$ can be proved, with a second order accuracy, which induces a first order accuracy for the total mass density $\sum \tilde{m}_i$.

Corollary 2.2. *Let assume that $f_0(S)$ is a C^3 function for $S \in \mathbb{R}^+$. For $\beta \neq 2$ and $\alpha \leq 3/2$ or for $\beta = 2$ and $\alpha \leq 5/2$, the semi-discretized scheme (2.2) is convergent. Moreover, the variables $\tilde{m}_i$ are then second order accurate: there exists a constant C such that, for all $i \in \{1, \ldots, N\}$ and $t \geq 0$, $|\tilde{m}_i - m_i| \leq C(\Delta\Psi)^2$.*

Since we are interested in the total mass density of the spray, that is to say, the ℓ^1 norm of the m_i, we evaluate the global order of the method through the global truncation error: $\gamma(t) = \sum_i |\gamma_i(t)|$. In the proof of Theorem 2.1, a difference of the order on the local truncation errors appears between different sections. Compensations can happen so that there is not a second local order accuracy for all the sections, but a global first order accuracy is obtained. Cases of first order accuracy are thus given in the next theorem. Moreover, an asymptotically "optimal" estimate is given for the global truncation error in this theorem.

Theorem 2.3. *Let $f_0(S)$ be a C^3 function for $S \in \mathbb{R}^+$. The global truncation error $\gamma(t)$ is then such that:*

- *for $\beta \neq 2$ and $\alpha > 3$, $\gamma(t) = O\left(N^{-3/\alpha}\right)$, with $3/\alpha < 1$;*
- *for $\beta = 2$ and $\alpha > 5$, $\gamma(t) = O\left(N^{-5/\alpha}\right)$, with $5/\alpha < 1$;*
- *for $\beta \neq 2$ and $\alpha \leq 3$, $\gamma(t) = O(N^{-1})$;*
- *for $\beta = 2$ and $\alpha \leq 5$, $\gamma(t) = O(N^{-1})$.*

Thus, for the last two cases, the semi-discretized scheme (2.2) is first order accurate. More precisely, a bound of the global truncation error is then given by:

$$|\gamma(t)| \leq \Delta\Psi \left(\frac{|\beta - 2|K_0}{2\alpha} \|f_0\|_\infty + \frac{K_1}{\alpha} \|f_0'\|_\infty + \frac{2}{\alpha(7-\alpha)} \|f_0''\|_\infty \right) + o(\Delta\Psi) \tag{2.7}$$

where the constant K_0 is equal to 1 if $\alpha < 3$ and is positive otherwise. The constant K_1 is equal to 0 if $\alpha \leq 3$ and $\beta = 7 - \alpha$, is equal to 1 if $\alpha < 5$ and $\beta \neq 7 - \alpha$ and is bigger than 3 if $\alpha = 5$ and $\beta = 2$. Moreover, the coefficients before each norm are asymptotically "optimal" in this estimate.

The loss of accuracy for the large value of α has to be related with the singularity around $\phi = 0$ observed in equation (2.1). Indeed, when α increases whereas the number N of sections is constant, the discretization becomes coarser around this singularity.

Finally, the completely discretized scheme (2.4) has to be studied. Its consistency with a first order accuracy comes from previous estimates. We determine a type of CFL condition for its stability, and thus its convergence, and another one to ensure the positivity of the droplet mass densities $\hat{m}_i^k$:

Theorem 2.4. *A CFL condition for the stability of the scheme (2.4) is of the type $\frac{\Delta t}{\Delta\Psi^{2/\alpha}} \leq C$ if $\alpha < 2$ and $\frac{\Delta t}{\Delta\Psi} \leq C$ otherwise, where C is a constant. Moreover, the scheme (2.4) ensures the positivity of the $\hat{m}_i^k$, for any positive initial conditions m_i^0, if and only if a condition of the same type holds.*

The value of the constant is given by Theorems 2.7 and 2.9 coupled to the Lemma 2.6. It proves that CFL like conditions for the stability and the positivity of the scheme are less restrictive for $\alpha \geq 2$. For example, let us compare the standard case $\alpha = \beta = 1$ ($\Delta\Psi = \Delta R$) and the case $\alpha = \beta = 2$ ($\Delta\Psi = \Delta S$). It can be proved that the conditions of Theorem 2.7 are fulfilled in both cases (because the function λ of Lemma 2.6 is a decreasing function). The stability condition is, for both cases:

$$\frac{\Delta t}{(\Delta R)^2} \leq \frac{2}{3}, \qquad \frac{\Delta t}{\Delta S} \leq \frac{4}{5}. \tag{2.8}$$

TABLE 1. Synthesis of the convergence results for the classical multi-fluid method.

	α	0	3/2	2	5/2	3	5
$\beta \neq 2$	global order	$= 1$					< 1
	local order	$= 2$		< 2			
$\beta = 2$	global order	$= 1$					< 1
	local order	$= 2$			< 2		
CFL		restrictive			less restrictive condition		

The positivity conditions are of the same type as for the stability. So, the CFL conditions are more restrictive in the first case than in the second.

The CFL conditions for the positivity lead to a monotone scheme for the approximation of the distribution function. It is thus very robust but it is at most first order accurate [11].

The main results are gathered in Table 1 and in the "optimal" estimate (2.7). In order to have a first global order accuracy, but also a second local order accuracy (and convergence) and no restrictive CFL like condition, we have to impose $\beta = 2$ and $\alpha \in [2, 2.5]$. This value of β cancels the coefficient in front of $\|f_0\|_\infty$ in estimate (2.7) and then in the numerical diffusion. The other coefficients in this bound are decreasing for $\alpha \in [2, 2.5]$ but they lose less than 1/4 of their values. So, the "optimal" case (best compromise between precision, stability and positivity) is obtained for $\alpha \in [2, 2.5]$ and $\beta = 2$. This result, concerning β, is rather natural since the evolution equation for f^ϕ is a pure advection equation with constant coefficients if and only if $\beta = 2$. Since $\alpha = 2$ has more physical sense, we choose a discretization uniform in surface with, as approximated distribution, a constant function of the surface in each section. For this case, the CFL like condition for the stability and for the positivity are respectively: $\frac{\Delta t}{\Delta S} \leq \frac{4}{5}$ and $\frac{\Delta t}{\Delta S} \leq \frac{2}{5}$.

2.2. **Proof for the order and the convergence of the semi-discretized scheme**

As usual, to determine the order of the method, the truncation error has to be evaluated through an expansion of an assumed regular function: f_0. To this purpose, let us begin by expressing the local truncation errors γ_i defined by (2.5) with the function f_0. The term $\partial_t m_i$ is given by an integration of equation (2.1) on the section i:

$$\partial_t m_i(t) = \frac{\beta}{2}\left[\phi^{1+1/\beta} f^\phi\right]_{\phi_{i-1}}^{\phi_i} - \frac{3}{2}\int_{\phi_{i-1}}^{\phi_i} \phi^{1/\beta} f^\phi(t,\phi)\,\mathrm{d}\phi. \tag{2.9}$$

Let us introduce the difference $h_i(t,\phi)$ between the real distribution $f^\phi(t,\phi)$ and the approximated distribution $m_i(t)\kappa_i^\phi(\phi)$. Because of the expressions (2.3) of κ_i^ϕ and (2.6) of f^ϕ, the function h_i is written:

$$h_i(t,\phi) = \frac{2}{\beta}\frac{\int_{\phi_{i-1}}^{\phi_i} \varphi^{3/\beta}[\phi^{2/\beta-1} f_0(t+\phi^{2/\beta}) - \varphi^{2/\beta-1} f_0(t+\varphi^{2/\beta})]\,\mathrm{d}\varphi}{\int_{\phi_{i-1}}^{\phi_i} \varphi^{3/\beta}\,\mathrm{d}\varphi}. \tag{2.10}$$

Introducing the equation (2.9) in the expression (2.5) of γ_i, we finally obtain:

$$\gamma_i(t) = \frac{\beta}{2}\phi_i^{1+1/\beta} h_{i+1}(t,\phi_i) - \frac{\beta}{2}\phi_{i-1}^{1+1/\beta} h_i(t,\phi_{i-1}) - \frac{3}{2}\int_{\phi_{i-1}}^{\phi_i} \phi^{1/\beta} h_i(t,\phi)\,\mathrm{d}\phi, \tag{2.11}$$

where $h_{N+1} = 0$. Inverting the integrals, the last term of this expression can be written:

$$\int_{\phi_{i-1}}^{\phi_i} \phi^{1/\beta} h_i(t,\phi)\,\mathrm{d}\phi = \frac{2}{\beta}\int_{\phi_{i-1}}^{\phi_i} \phi^{3/\beta-1} f_0\left(t+\phi^{2/\beta}\right) d_i(\phi)\ \mathrm{d}\phi,$$

where the function d_i is defined by:

$$d_i(\phi) = 1 - \left(\frac{\phi}{\phi_{ci}}\right)^{2/\beta}, \qquad (\phi_{ci})^{2/\beta} = \frac{\int_{\phi_{i-1}}^{\phi_i} \varphi^{3/\beta}\,\mathrm{d}\varphi}{\int_{\phi_{i-1}}^{\phi_i} \varphi^{1/\beta}\,\mathrm{d}\varphi}\cdot$$

Performing a Taylor expansion of the C^3 function f_0 between each $t+\varphi^{\frac{2}{\beta}}$ or $t+\phi_{i-1}^{\frac{2}{\beta}}$ and $t+\phi_i^{\frac{2}{\beta}}$, the local truncation error becomes:

$$\gamma_i(t) = N^{-3/\alpha}\, v_i^0\, f_0(t+\phi_i^{2/\beta}) + N^{-5/\alpha}\, v_i^1\, f_0'(t+\phi_i^{2/\beta}) + N^{-7/\alpha}\,\frac{v_i^2}{2}\, f_0''(t+\phi_i^{2/\beta}) + r_i(t) \tag{2.12}$$

with

$$|r_i(t)| \le N^{-9/\alpha}\,\frac{v_i^3}{3}\,\|f_0^{(3)}\| \tag{2.13}$$

and where $v_i^j = i^{(3+2j)/\alpha}\, g^j\left(\frac{1}{i}\right)$ for $j=0,1,2,3$ and the $g^j(x)$ are defined for $x\in]0,1]$ by:

$$g^0(x) = 1-(1-x)^{\frac{3}{\alpha}} - \frac{\int_1^{(1+x)^{\frac{\beta}{\alpha}}} y^{\frac{5}{\beta}-1}\,\mathrm{d}y}{\int_1^{(1+x)^{\frac{\beta}{\alpha}}} y^{\frac{3}{\beta}}\,\mathrm{d}y} + (1-x)^{\frac{1+\beta}{\alpha}}\,\frac{\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{5}{\beta}-1}\,\mathrm{d}y}{\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{3}{\beta}}\,\mathrm{d}y} - \frac{3}{\beta}\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{3}{\beta}-1}\left[1-\frac{y^{\frac{2}{\beta}}}{c(x)}\right]\mathrm{d}y,$$

$$\begin{aligned} g^j(x) = {} & -\frac{\int_1^{(1+x)^{\frac{\beta}{\alpha}}} y^{\frac{5}{\beta}-1}\left(y^{\frac{2}{\beta}}-1\right)^j\,\mathrm{d}y}{\int_1^{(1+x)^{\frac{\beta}{\alpha}}} y^{\frac{3}{\beta}}\,\mathrm{d}y} + (1-x)^{\frac{1+\beta}{\alpha}}\,\frac{\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{5}{\beta}-1}\left(y^{\frac{2}{\beta}}-1\right)^j\,\mathrm{d}y}{\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{3}{\beta}}\,\mathrm{d}y} \\ & - (1-x)^{\frac{3}{\alpha}}\left[(1-x)^{\frac{2}{\alpha}}-1\right]^j - \frac{3}{\beta}\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{3}{\beta}-1}\left(y^{\frac{2}{\beta}}-1\right)^j\left[1-\frac{y^{\frac{2}{\beta}}}{c(x)}\right]\mathrm{d}y, \qquad j=1,2, \end{aligned}$$

$$\begin{aligned} g^3(x) = {} & \frac{\int_1^{(1+x)^{\frac{\beta}{\alpha}}} y^{\frac{5}{\beta}-1}\left(y^{\frac{2}{\beta}}-1\right)^3\,\mathrm{d}y}{\int_1^{(1+x)^{\frac{\beta}{\alpha}}} y^{\frac{3}{\beta}}\,\mathrm{d}y} + (1-x)^{\frac{1+\beta}{\alpha}}\,\frac{\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{5}{\beta}-1}\left(1-y^{\frac{2}{\beta}}\right)^3\,\mathrm{d}y}{\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{3}{\beta}}\,\mathrm{d}y} \\ & + (1-x)^{\frac{3}{\alpha}}\left[1-(1-x)^{\frac{2}{\alpha}}\right]^3 + \frac{3}{\beta}\int_{(1-x)^{\frac{\beta}{\alpha}}}^1 y^{\frac{3}{\beta}-1}\left(1-y^{\frac{2}{\beta}}\right)^3\left|1-\frac{y^{\frac{2}{\beta}}}{c(x)}\right|\mathrm{d}y, \end{aligned}$$

with

$$c(x) = \frac{\int_{(1-x)^{\beta/\alpha}}^1 y^{3/\beta}\,\mathrm{d}y}{\int_{(1-x)^{\beta/\alpha}}^1 y^{1/\beta}\,\mathrm{d}y}\cdot$$

Developments in the neighborhood of $x=0$ of these functions yield:

$$g^0(x) = \frac{(\beta-2)(3-\alpha)}{2\alpha^2}x^2 + \frac{(\beta-2)[(2\alpha-3)\beta+4(\alpha-3)]}{12\alpha^3}x^3 + O(x^4), \tag{2.14}$$

$$g^1(x) = \frac{\alpha+\beta-7}{\alpha^2}x^2 + \frac{40-12\alpha-\beta^2}{6\alpha^3}x^3 + O(x^4), \tag{2.15}$$

$$g^2(x) = -\frac{4}{\alpha^2}x^2 + O(x^3), \qquad g^3(x) = \frac{12}{\alpha^3}x^3 + O(x^4). \tag{2.16}$$

The expression (2.12), with (2.13), allows us to determine the global and local orders of the scheme.

2.2.1. *Proof of Theorem 2.1: local order*

The expression (2.12) of the local truncation error can be easily bounded by:

$$|\gamma_k(t)| \leq N^{-3/\alpha}|v_k^0|\,\|f_0\|_\infty + N^{-5/\alpha}|v_k^1|\,\|f_0'\|_\infty + N^{-7/\alpha}\frac{|v_k^2|}{2}\|f_0''\|_\infty + N^{-9/\alpha}\frac{v_k^3}{3}\|f_0^{(3)}\|$$

and where $v_k^j = k^{(3+2j)/\alpha} g^j\left(\frac{1}{k}\right)$ for $j = 0, 1, 2, 3$. It has also be shown, with the developments (2.14–2.16), that $g^0(x)/x^2$, $g^1(x)/x^2$, $g^2(x)/x^2$ and $g^3(x)/x^3$ have some finite limit when x tends to 0. Moreover, they are continuous for $x \in]0, 1]$. They are then bounded by some constants C_j for $j = 1, 2, 3$ or $|\beta - 2|C_0$ for $j = 0$. For γ_k, this leads to:

$$\begin{aligned}|\gamma_k(t)| \leq{}& N^{-3/\alpha}k^{3/\alpha-2}|\beta-2|C_0\,\|f_0\|_\infty + N^{-5/\alpha}k^{5/\alpha-2}C_1\,\|f_0'\|_\infty + N^{-7/\alpha}k^{7/\alpha-2}\frac{C_2}{2}\|f_0''\|_\infty \\ &+ N^{-9/\alpha}k^{9/\alpha-3}\frac{C_3}{3}\|f_0^{(3)}\|_\infty.\end{aligned}$$

Finally, $N^{-\lambda/\alpha}k^{\lambda/\alpha-i}$ has the following upper bound: $1/N^i$ if $\lambda/\alpha \geq i$ and $N^{-\lambda/\alpha}$ if $\lambda/\alpha < i$. This proves that the scheme is locally second order accurate in $\Delta\Psi$ if $\beta \neq 2$ and $\alpha \leq 3/2$ or if $\beta = 2$ and $\alpha \leq 5/2$.

Otherwise, the expression (2.12) shows that, if $\beta \neq 2$, then $\gamma_1(t)$ behaves as $N^{-3/\alpha}$ and if $\beta = 2$, then $\gamma_1(t)$ behaves as $N^{-5/\alpha}$. Then, this local truncation error cannot be second order accurate.

Remark 2.5. However, the local truncation error $\gamma_N(t)$ is second order accurate even if the preceding conditions on α and β are not satisfied. This shows that the local truncation errors have not the same order of accuracy.

The convergence of the semi-discretized scheme (2.2) can then be proved, with the same assumptions.

2.2.2. *Proof of Corollary 2.2: convergence*

There exists a unique solution C^1 on $\mathbb{R}^+$ of the equations (2.2) and it is easy to show that the obtained $\tilde{m}_i(t)$ are positive and tends to 0 when t tends to infinity. Theorem 2.1 proves that the semi-discretized scheme is locally second order accurate:

$$|\gamma_i(t)| \leq K\,(\Delta\Psi)^2.$$

Let us note $e_i(t) = m_i(t) - \tilde{m}_i(t)$ the local error on the mass density. This quantity satisfies the following equation:

$$e_i'(t) + (F_i + M_i)e_i - F_{i+1}e_{i+1} = \gamma_i, \qquad e_i(0) = 0.$$

Moreover, the limits of the C^1 function $e_i(t)$ when t tends to 0 or infinity are 0. So, $|e_i(t)|$ admits a maximal value, attained at a point t_0:

$$|e_i(t)| \leq |e_i(t_0)| = \left|\frac{\gamma_i(t_0) + F_{i+1}e_{i+1}(t_0)}{F_i + M_i}\right| \leq \frac{\|\gamma_i\|_\infty}{F_i + M_i} + \frac{F_{i+1}}{F_i + M_i}\|e_{i+1}\|_\infty.$$

Then, because of the inequality on γ_i:

$$\|e_i\|_\infty \leq K(\Delta\Psi)^2 u_i^N, \qquad u_i^N = \sum_{k=i}^{N}\frac{1}{F_k + M_k}\left(\prod_{p=i}^{k-1}\frac{F_{p+1}}{F_p + M_p}\right).$$

Let us now prove that u_i^N is bounded. This term can be written:

$$u_i^N = \sum_{k=i}^{N}\frac{2(1+\beta)}{(3+\beta)N^{\frac{2}{\alpha}}}\,\frac{k^{\frac{3+\beta}{\alpha}} - (k-1)^{\frac{3+\beta}{\alpha}}}{3k^{\frac{1+\beta}{\alpha}} + (\beta-2)(k-1)^{\frac{1+\beta}{\alpha}}}\prod_{p=i}^{k-1}\frac{(1+\beta)\left[1 - \left(1 - \frac{1}{p}\right)^{\frac{3+\beta}{\alpha}}\right]}{\left[3 + (\beta-2)\left(1 - \frac{1}{p}\right)^{\frac{1+\beta}{\alpha}}\right]\left[\left(1 + \frac{1}{p}\right)^{\frac{3+\beta}{\alpha}} - 1\right]}.$$

The factors are smaller than one: it is easy to prove in the case $\beta \leq 2$ since $(3+\beta)/\alpha \leq 1$ and, in the case $\beta > 2$, it is maximum for $p = +\infty$. Then:

$$u_i^N \leq \frac{1}{N^{\frac{2}{\alpha}}} \sum_{k=1}^{N} k^{\frac{2}{\alpha}} \lambda\left(\frac{1}{k}\right), \qquad \lambda(x) = \frac{2(1+\beta)}{3+\beta} \frac{1-(1-x)^{\frac{3+\beta}{\alpha}}}{3+(\beta-2)(1-x)^{\frac{1+\beta}{\alpha}}}.$$

The terms of the sum are equivalent to $\frac{2}{\alpha} k^{2/\alpha - 1}$ when k tends to infinity and form a diverging series. The upper bound of u_i^N then tends to 1 when N tends to infinity and u_i^N is then bounded for all $i \in \mathbb{N}$ and $N \in \mathbb{N}$. This completes the proof of the corollary.

2.2.3. *Proof of Theorem 2.3: global order*

Remark 2.5 suggests that some compensations can happen between the local errors so that the scheme can be first order accurate without being locally second order accurate for all the sections. Moreover, we want here to give an "optimal" estimate of the global error.

Thanks to (2.12), with (2.13), the global error is given by:

$$\gamma(t) = \sum_{k=1}^{N} \left| N^{-3/\alpha}\, v_k^0\, f_0(t+\phi_k^{2/\beta}) + N^{-5/\alpha}\, v_k^1\, f_0'(t+\phi_k^{2/\beta}) + N^{-7/\alpha}\, \frac{v_k^2}{2}\, f_0''(t+\phi_k^{2/\beta}) + r_k(t) \right|$$

with

$$\sum_{k=1}^{N} |r_k(t)| \leq N^{-9/\alpha} \sum_{k=1}^{N} \frac{\left|v_k^3\right|}{3} \|f_0^{(3)}\|.$$

We then have to evaluate each sum $N^{-(3+2j)/\alpha} \sum_{k=1}^{N} |v_k^j|$ and find an equivalent when N tends to infinity.

Thanks to the development (2.14) of g^0, $|v_k^0|$ is equivalent to $\frac{|\beta-2|(\alpha-3)}{2\alpha^2} k^{\frac{3}{\alpha}-2}$ if $\alpha \neq 3$ and to $\frac{\beta|\beta-2|}{108} k^{-2}$ if $\alpha = 3$. Consequently, when N tends to infinity,

- for $\alpha > 3$, $\sum |v_k^0|$ is a converging series and $N^{-3/\alpha} \sum_{k=1}^{N} |v_k^0| = O(N^{-3/\alpha})$, with $\frac{3}{\alpha} < 1$;
- for $\alpha = 3$, $\sum |v_k^0|$ is a converging series, its sum is denoted $\frac{|\beta-2|K_0}{2\alpha}$ and $N^{-3/\alpha} \sum_{k=1}^{N} |v_k^0|$ is then equivalent to $\frac{|\beta-2|K_0}{2\alpha N}$;
- for $\alpha < 3$, $\sum |v_k^0|$ is a diverging series and $N^{-3/\alpha} \sum_{k=1}^{N} |v_k^0|$ is equivalent to $\frac{|\beta-2|}{2\alpha N}$.

Thanks to the development (2.15) of g^0, $|v_k^1|$ is equivalent to $\frac{|\alpha+\beta-7|}{\alpha^2} k^{\frac{5}{\alpha}-2}$ if $\beta \neq 7-\alpha$ and to $\frac{|9-2\alpha+\alpha^2|}{6\alpha^3} k^{\frac{5}{\alpha}-3}$ if $\beta = 7-\alpha$. Then, if $\beta \neq 7-\alpha$ and when N tends to infinity:

- for $\alpha > 5$, $\sum |v_k^1|$ is a converging series and $N^{-5/\alpha} \sum_{k=1}^{N} |v_k^1| = O(N^{-5/\alpha})$, with $\frac{5}{\alpha} < 1$;
- for $\alpha = 5$, $N^{-5/\alpha} \sum_{k=1}^{N} |v_k^1| = O(\log(N)/N)$;
- for $\alpha < 5$, $\sum |v_k^1|$ is a diverging series and $N^{-5/\alpha} \sum_{k=1}^{N} |v_k^1|$ is equivalent to $\frac{|\alpha+\beta-7|}{\alpha(5-\alpha)N}$,

and if $\beta = 7-\alpha$, then:

- for $\alpha > 5/2$, $\sum |v_k^1|$ is a converging series and $N^{-5/\alpha} \sum_{k=1}^{N} |v_k^1| = O(N^{-5/\alpha})$;
- for $\alpha = 5/2$, $N^{-5/\alpha} \sum_{k=1}^{N} |v_k^1| = O(\log(N)/N^2)$;
- for $\alpha < 5/2$, $\sum |v_k^1|$ is a converging series and $N^{-5/\alpha} \sum_{k=1}^{N} |v_k^1| = O(N^{-2})$.

Moreover, in the case $\alpha = 5$ and $\beta = 2$, $\sum |v_k^1|$ is a converging series and $N^{-5/\alpha} \sum_{k=1}^{N} |v_k^1| = O(1/N)$.

The development (2.16) of the function $g^2(x)$ close to $x = 0$ shows that $|v_k^2|$ is equivalent to $\frac{4}{\alpha^2} k^{\frac{7}{\alpha}-2}$. Then,

- for $\alpha > 7$, $\sum |v_k^2|$ is a converging series and $N^{-7/\alpha} \sum_{k=1}^{N} |v_k^2| = O(N^{-7/\alpha})$, with $\frac{7}{\alpha} < 1$;
- for $\alpha = 7$, $\sum |v_k^2|$ is a diverging series and $N^{-7/\alpha} \sum_{k=1}^{N} |v_k^0| = O(\log(N)/N)$;
- for $\alpha < 7$, $\sum |v_k^2|$ is a diverging series and $N^{-7/\alpha} \sum_{k=1}^{N} |v_k^2|$ is equivalent to $\frac{4}{\alpha(7-\alpha)N}$.

The development (2.16) of the function $g^3(x)$ close to $x = 0$ shows that $|v_k^3|$ is equivalent to $\frac{12}{\alpha^3} k^{\frac{9}{\alpha}-3}$. Then,

- for $\alpha > 9/2$, $\sum |v_k^3|$ is a converging series and $N^{-9/\alpha} \sum_{k=1}^N |v_k^3| = O(N^{-9/\alpha})$;
- for $\alpha = 9/2$, $\sum |v_k^3|$ is a diverging series and $N^{-9/\alpha} \sum_{k=1}^N |v_k^3| = O(\log(N)/N^2)$;
- for $\alpha < 9/2$, $\sum |v_k^3|$ is a diverging series and $N^{-9/\alpha} \sum_{k=1}^N |v_k^3| = O(N^{-2})$.

Then, for $\alpha < 9$, $N^{-9/\alpha} \sum_{k=1}^N |v_k^3| = o(N^{-1})$.

Gathering the previous results, we obtain the order of the method given in the Theorem. Moreover, in the case of the first order accuracy (*i.e.* $\beta \neq 2$ and $\alpha \leq 3$ or $\beta = 2$ and $\alpha \leq 5$), a bound of the global truncation error can be written:

$$\gamma(t) \leq N^{-3/\alpha} \sum_{k=1}^N |v_k^0| \|f_0\|_\infty + N^{-5/\alpha} \sum_{k=1}^N |v_k^1| \|f_0'\|_\infty + N^{-7/\alpha} \sum_{k=1}^N |v_k^2| \|f_0''\|_\infty + o\left(\frac{1}{N}\right).$$

In this bound, the coefficients before each norm are asymptotically "optimal" in the sense that for each of them, one can find C^3 functions f_0 such that the equality is reached. With the previous results, this bound can be written as in (2.7).

2.3. **Full-discretized scheme: proof for stability and positivity**

We determine a type of CFL condition for the stability of the scheme (2.4) and another one to ensure the positivity of the droplet mass densities $\hat{m}_i^k$. Before that, we need a property on the coefficients $M_i + F_i$ and more precisely on $\mathcal{M} = \max\{M_i + F_i,\, i \in \{1, \ldots, N\}\}$.

Lemma 2.6. Coefficients – *The coefficients $F_i + M_i$ are positive and their maximum value is such that:*

- *if $\alpha < 2$ then $\mathcal{M} = O((\Delta\Psi)^{-2/\alpha})$;*
- *if $\alpha = 2$ then $\mathcal{M} = F_1 + M_1 = \frac{3}{2}\frac{3+\beta}{1+\beta}(\Delta\Psi)^{-1}$;*
- *if $\alpha > 2$ then, for a "big enough" N, $\mathcal{M} = F_N + M_N$ and $\mathcal{M} = O((\Delta\Psi)^{-1})$.*

Proof. It is first convenient to give another expression of the coefficients:

$$M_i + F_i = \frac{1}{2}\frac{3+\beta}{1+\beta}(\Delta\Psi)^{-2/\alpha}\lambda(i-1), \qquad \lambda(t) = \frac{3(t+1)^{\frac{1+\beta}{\alpha}} + (\beta-2)t^{\frac{1+\beta}{\alpha}}}{(t+1)^{\frac{3+\beta}{\alpha}} - t^{\frac{3+\beta}{\alpha}}}. \tag{2.17}$$

The function $\lambda(t)$ is equivalent to $\alpha\frac{\beta+1}{\beta+3}t^{1-2/\alpha}$ when t tends to $+\infty$.

If $\alpha > 2$, $\lambda(t)$ tends then to $+\infty$ when t tends to $+\infty$. So, if N is big enough, then $\lambda(i-1)$ is maximal for $i = N$ and $\mathcal{M} = F_N + M_N$ is equivalent to N when N tends to $+\infty$. It shows the third point of the lemma because $\Delta\Psi = 1/N$.

If $\alpha < 2$ then $\lambda(t)$ is a continuous function for $t \in \mathbb{R}^+$ and tends to 0 when t tends to $+\infty$. It is then bounded and $\mathcal{M}$ varies like $(\Delta\Psi)^{-2/\alpha}$.

If $\alpha = 2$, let us proceed with the following change of variables:

$$\forall y \in [0,1[\qquad \mu(y) = \lambda\left(\frac{y}{1-y}\right) = \frac{1-y}{1-y^{\frac{3+\beta}{2}}}\left[3 + (\beta-2)y^{\frac{1+\beta}{2}}\right] \tag{2.18}$$

noting that the function $y \to \frac{y}{1-y}$ is an increasing function from $y \in [0,1[$ in $\mathbb{R}^+$. Let us show that $\mu(y)$ admits a maximum value in $y = 0$.

For all $y \in [0,1[$, we have $y^{(3+\beta)/2} \leq y$. So, for $\beta \leq 2$ we have $\mu(y) \leq 3 + (\beta-2)y^{(1+\beta)/2} \leq 3 = \mu(0)$ and $\mu(y)$ admits a maximum value at $y = 0$.

For $\beta > 2$:

$$\mu(0) - \mu(y) = \frac{y}{1 - y^{\frac{3+\beta}{2}}}[3 - \varphi(y)], \qquad \varphi(y) = (\beta - 2)y^{\frac{\beta-1}{2}} + (5 - \beta)y^{\frac{1+\beta}{2}} \tag{2.19}$$

and the derivative of φ is positive. So $\varphi(y)$ is smaller than $\varphi(1) = 3$ for $y \in [0, 1[$ and $\mu(y)$ admits a maximum value in $y = 0$ because of (2.19). It proves the second point of the lemma. □

One may now give a stability result for the scheme (2.4):

Theorem 2.7. Stability – *Assume that the $M_i + F_i$ are pairwise distinct. Then, the scheme (2.4) is stable if and only if $\mathcal{M}\,\Delta t \leq 2$.*

Proof. To show this result, we introduce the matrix C of the scheme (2.4): $\hat{m}^{k+1} = C\hat{m}^k$, where $\hat{m}^k = (\hat{m}_1^k, \hat{m}_2^k, \ldots, \hat{m}_N^k)^t$. This matrix is upper triangular and its diagonal terms are: $c_{ii} = 1 - \Delta t\,(F_i + M_i)$. As the $M_i + F_i$, the c_{ii} are pairwise distinct. Then C is diagonalizable, its eigenvalues being the c_{ii}. So, the scheme is stable if and only if, for all $i \in \{1, \ldots, N\}$, we have $|c_{ii}| \leq 1$. This proves the theorem. □

Remark 2.8. The assumption that the $M_i + F_i$ are pairwise distinct can be avoided but, in this case, the inequality induced for $\mathcal{M}$ has to be strict. Indeed, the matrix C is then no more diagonalizable but a sufficient condition for stability is that its spectral radius is smaller than one, as explained in Section 3.1.5.

The shape of the scheme (2.4) leads straight to the following theorem:

Theorem 2.9. Positivity – *The scheme (2.4) ensures the positivity of the $\hat{m}_i^k$, for any positive initial conditions m_i^0, if and only if $\mathcal{M}\,\Delta t \leq 1$.*

Lemma 2.6 and Theorems 2.7 and 2.9 allows us to determine CFL conditions for stability and positivity for all values of α and β. In particular, in the case $\alpha < 2$: $\mathcal{M} = O(\Delta\Psi^{-2/\alpha})$. So the CFL conditions are of the following type: $\Delta t/\Delta\Psi^{2/\alpha}$ smaller than a constant. These results are gathered in Theorem 2.4. It shows that the CFL condition is then more restrictive for $\alpha < 2$ than in the case $\alpha \geq 2$.

2.4. **Exponentially decreasing distribution function at infinity**

In previous sections, we assumed that the distribution function had a compact support in $\mathbb{R}^+$. The practical cases satisfy at least approximately this property. It is not interesting to use a refined discretization in the tail of the distribution, where the mass density is small. Moreover, in some experiments [17], the tail of the distribution was shown to decrease exponentially as a function of the droplet surface. This property was used with the classical multi-fluid method: the profile of the last section was chosen to be also exponentially decreasing, allowing to reduce the number of sections [17] and also to completely cover $\mathbb{R}^+$ with sections, an important aspect to be able to account for droplet coalescence [16]. Indeed, this profile is compatible with the classical multi-fluid method because it is preserved in the purely evaporating case, as can be shown as follows. Assume that $f_0(S)$ is proportional to an exponential

$$\forall S \geq 1 \qquad f_0(S) = b\,\exp(-\lambda\,S), \tag{2.20}$$

where b and λ are two positive constants and where the value 1 is obtained through the choice of the non-dimensional surface S. Then, the exact solution of the equation (1.2) is $f^S(t, S) = f_0(S + t) = b\,\exp(-\lambda\,t)\exp(-\lambda\,S)$ for $S + t \geq 1$, indicating that f^S is still proportional to the same exponential. This type of profile is also used in [6] in a revisited multi-fluid approach, where the distribution is chosen to be exponential in each section, the parameters of this exponential being deduced from the mass density and also the number density of the droplets in each section.

The purpose of the following analysis is to show that, at least when the initial distribution tail is exponentially decreasing, the use of such a profile for the last section induces a convergent scheme. One assumes that the initial distribution function $f_0(S)$ is C^3 for $S \in \mathbb{R}^+$ and satisfy the assumption (2.20). Then, as in [17], only one section is used to describe this distribution tail, the profile of the approximated distribution in this section being exponentially decreasing as a function of the surface, with the same coefficient λ. Compared with previous formulations, we only have to add a $(N+1)$th section $[S_N, +\infty[$ (with $S_N = 1$) and the profile κ^S_{N+1} of the distribution function in this section is chosen such that:

$$\kappa^S_{N+1}(S) = \frac{\mathrm{e}^{-\lambda S}}{\int_1^{+\infty} \sigma^{3/2}\mathrm{e}^{-\lambda\sigma}\,\mathrm{d}\sigma}. \tag{2.21}$$

One may first remark that this last section has no influence on the order of the method. Indeed, the difference $h_{N+1}(t,S) = f^S(t,S) - m_{N+1}(t)\kappa^S_{N+1}(S)$ between the real distribution and the approximated distribution is equal to zero for $S \geq 1$. Furthermore, the parameter γ_{N+1} used to evaluate the local error for the last section and defined by (2.5) is still given by equation (2.11), for $i = N+1$. This parameter is then equal to zero: $\gamma_{N+1} = 0$. Moreover, the influence of the addition of the last section on the local order for the other sections can be seen as the addition in γ_N and then in γ of a term:

$$\frac{\beta}{2}\phi_N^{1+1/\beta}\frac{\int_{\phi_N}^{\phi_{N+1}} \varphi^{3/\beta}[f^\phi(t,\phi) - f^\phi(t,\varphi)]\mathrm{d}\varphi}{\int_{\phi_N}^{\phi_{N+1}} \varphi^{3/\beta}\,\mathrm{d}\varphi},$$

where ϕ_{N+1} is defined by $\phi_{N+1} = ((N+1)\Delta\Psi)^{\beta/\alpha}$. This term corresponds to the difference between the real distribution and an approximated distribution constant as a function of Ψ, in a section $[\phi_N, \phi_{N+1}]$. It is a first order term in $\Delta\Psi$ and then, the global order remains unchanged, like the local order of all sections except that of the Nth section.

The CFL conditions for the stability and the positivity are still given by Theorems 2.7 and 2.9. Let us then evaluate the coefficient $\mathcal{M} = \max\{M_i + F_i, i \in \{1,\ldots,N\}\}$. For the last section, the coefficient $M_{N+1} + F_{N+1}$ is written:

$$M_{N+1} + F_{N+1} = S_N^{3/2}\kappa^S_{N+1}(S_N) + \frac{3}{2}\int_{S_N}^{+\infty} S^{1/2}\kappa^S_{N+1}(S)\,\mathrm{d}S = \lambda.$$

Thus, for sufficiently small $\Delta\Psi$, the coefficient $\mathcal{M}$ is still controlled by the $M_i + F_i$ for $i \leq N$ because these coefficients are $O((\Delta\Psi)^{-1})$ or $O((\Delta\Psi)^{-2/\alpha})$ according to Lemma 2.6. The CFL conditions are therefore unchanged.

2.5. **Conclusion**

The classical multi-fluid method is then proved to be first order accurate under some compatibility assumption on the discretization. The study of the completely discretized scheme, using an explicit Euler scheme for time discretization, indicates that it is stable and also monotone and thus TVD (Total Variation Diminishing) in terms of the approximation of the distribution function. However, the main drawback of this method is its large numerical diffusion and its low convergence rate. This becomes critical when one wishes to reduce the computational cost, and hence, the number of sections and also when the initial distribution is not smooth. The necessity to develop higher order methods is then clearly identified.

3. Higher order methods

Several strategies can be developed in order to increase the order of a numerical method to predict the evolution of the mass density of a spray. However, this has to be done in the perspective of a generalization to more complex cases including particularly the dynamic of the droplets. Since the behavior of the droplets is very dependent of their size (for example, the drag force F in Eq. (1.1) varies like $1/S$), it seems important to discretize the phase space corresponding to the droplet size variable, and thus develop other multi-fluid methods.

Their accuracy may be increased by improving the distribution function approximation in each section. One way to do this is to use more than one moment (the mass density for the classical multi-fluid method) to describe each section. Several strategies can be followed. For example, in [6], the mass and the number density are used and the approximated distribution function is exponentially decreasing.

We here choose to use m moments in the droplet surface. This has the following advantages. First, the variable S is the most natural to use for some Godunov type scheme and leads to simpler numerical scheme because the transport equation is an advection equation with constant velocity when written with this variable. Second, the classical multi-fluid method is more accurate if the S variable is used for discretization. Finally, in the general case of equation (1.1), the drag force F is proportional to $1/S$, and use of moments in S can simplify the closure of the global system. With these moments, an approximate polynomial distribution function is constructed in each section. This polynomial, with a degree smaller than $m-1$, can be interpreted as the approximation of the distribution by the least square method.

In this section, a detailed presentation of the method is given in the general case where m moments are considered in each section. A preliminary result is derived. The order of accuracy of the mass density is analyzed in a second step (Thm. 3.4). The order of consistency of the method for the calculation of the moments (Thm. 3.5) are proved to be equal to the number m of moments used in each section. A stability condition is given and verified in the particular cases $m=2$ and $m=3$ (Thm. 3.6). But this method is neither positive nor TVD. In the particular case where 2 moments are used in each section, a modification of the method is proposed, so that it becomes positive. A condition to locate and deal with discontinuities is also proposed. This acts like a slope limiter condition and limits the increase of the total variation.

3.1. m moments method

Equation (1.2) is now discretized as follows. The phase space corresponding to the droplet surface is divided in N sections: for $i \in \{1,\dots,N\}$ let note $S_i = i\Delta S$ (with $\Delta S = 1/N$). It is convenient to use the mid points $S_{i-1/2} = (i-1/2)\Delta S$. Likewise, a time discretization is introduced: $t_k = k\Delta t$, where k is a positive integer. Here, we only used the distribution function f^S as a function of the surface and of the discrete time t_k. To simplify the notations, let $f^k(S) = f^S(t_k, S)$.

A finite volume kinetic scheme will be defined to yield an approximation of the moments:

$$\forall i \in \{1,\dots,N\} \quad \forall \alpha \in \{0,\dots,m-1\} \qquad \overline{nS_i^\alpha}^k = \int_{S_{i-1}}^{S_i} S^\alpha f^k(S)\,\mathrm{d}S. \tag{3.1}$$

It can be view like a three steps method. The first step is a reconstruction f_a^k of the distribution function from the approximated moments $\widetilde{nS_i^\alpha}^k$ at time t_k. This function is chosen such that, for each $i \in \{1,\dots,N\}$, $f_a^k|_{[S_{i-1},S_i[}$ is a polynomial function with a degree less or equal to $m-1$ and for which the moments are given by the $\widetilde{nS_i^\alpha}^k$:

$$\forall i \in \{1,\dots,N\} \quad \forall \alpha \in \{0,\dots,m-1\} \qquad \int_{S_{i-1}}^{S_i} S^\alpha f_a^k(S)\,\mathrm{d}S = \widetilde{nS_i^\alpha}^k. \tag{3.2}$$

The second step is the transport of this function, by solving the following problem:

$$\begin{cases} \partial_t \tilde{f} - \partial_S \tilde{f} = 0, & t \in [t_k, t_{k+1}], S \in \mathbb{R}^+, \\ \tilde{f}(t_k, S) = f_a^k(S), & S \in \mathbb{R}^+. \end{cases} \tag{3.3}$$

Finally, a projection step provides the $\widetilde{nS_i^\alpha}^{k+1}$ by taking the moments along the variable S of the function $\tilde{f}(t_{k+1}, S)$. Since this function is equal to $f_a^k(S+\Delta t)$, the scheme is described by the following equations,

for $i \in \{1, \dots, N\}$ and $\alpha \in \{0, \dots, m-1\}$:

$$\widetilde{nS_i^\alpha}^{k+1} = \int_{S_{i-1}+\Delta t}^{S_i+\Delta t} (S-\Delta t)^\alpha f_a^k(S)\,\mathrm{d}S = \sum_{p=0}^{\alpha} C_\alpha^p\,(-\Delta t)^{\alpha-p}\left(\widetilde{nS_i^p}^k + \Delta t\left(F_{i+1}^{p,k} - F_i^{p,k}\right)\right) \tag{3.4}$$

where the fluxes $F_i^{\alpha,k}$ are defined by:

$$F_i^{\alpha,k} = \frac{1}{\Delta t}\int_{S_{i-1}}^{S_{i-1}+\Delta t} S^\alpha f_a^k(S)\,\mathrm{d}S. \tag{3.5}$$

The droplet mass densities in each section are not directly evaluated by the scheme but the approximated distribution function is used for that:

$$\tilde{m}_i^k = \int_{S_{i-1}}^{S_i} S^{3/2} f_a^k(S)\,\mathrm{d}S. \tag{3.6}$$

Remark 3.1. Let introduce the scalar product $(f,g) \mapsto \int_{S_{i-1}}^{S_i} f(x)g(x)\mathrm{d}x$ of $\mathrm{L}^2([S_{i-1}, S_i])$. In the case where all the moments are exact ($\widetilde{nS_i^\alpha}^k = \overline{nS_i^\alpha}^k$), the conditions (3.2) are equivalent to $(P, f_a^k) = (P, f^k)$ for all $P \in \mathbb{R}_{m-1}[X]$. Thus, the restriction of f_a^k to the section $[S_{i-1}, S_i[$ is the polynomial function of degree $m-1$ obtained by the least square method to approximate f^k on this interval.

3.1.1. *Approximated distribution function*

We first check the existence and the uniqueness of such a function f_a^k. For that, we introduce basis of the set $\mathbb{R}_{m-1}[X]$ of polynomials with a degree of $m-1$ or less. The restriction of f_a^k to each section is expressed with the help of one of these basis. To build them, bijections are used, as defined in the following proposition, which is a classical result of linear algebra.

Proposition 3.2. *For all real numbers a and b such that $a < b$, the applications $\varphi_{a,b}(P)$ defined for each P in $\mathbb{R}_{m-1}[X]$ and taking their value in $\mathbb{R}^m$:*

$$\varphi_{a,b}(P) = \left(\int_a^b P(x)\,\mathrm{d}x, \int_a^b xP(x)\,\mathrm{d}x, \dots, \int_a^b x^{m-1}P(x)\,\mathrm{d}x\right)^t \tag{3.7}$$

where the exponent t denotes the transpose, are invertible linear applications.

This shows the existence and uniqueness of the function f_a^k since each function φ_{S_{i-1},S_i} is invertible. Two basis of $\mathbb{R}_{m-1}[X]$ are then introduced by defining:

$$P_j^{(i)} = \varphi_{S_{i-1},S_i}^{-1}(e_{j+1}), \qquad R_j = \varphi_{-1,1}^{-1}(e_{j+1}), \tag{3.8}$$

where $(e_1, \dots, e_m)$ is the natural basis vector of $\mathbb{R}^m$ so that the moments of these polynomial are:

$$\int_{S_{i-1}}^{S_i} S^l\,P_j^{(i)}(S)\;\mathrm{d}S = \delta_{j,l}, \qquad \int_{-1}^{1} y^l\,R_j(y)\;\mathrm{d}y = \delta_{j,l}, \qquad l \in \{0, \dots, m-1\} \tag{3.9}$$

where $\delta_{j,l}$ is equal to 1 if $j = l$ and 0 in the other case. The basis $(P_0^{(i)}, \dots, P_{m-1}^{(i)})$ are used to give an expression of f_a^k as a function of the moments $\widetilde{nS_i^\alpha}^k$ whereas the basis $(R_0, \dots, R_{m-1})$ is independent of the discretization and will be used to evaluate the accuracy of the scheme.

The function f_a^k is then given by:

$$\forall S \in [S_{i-1}, S_i[\quad f_a^k(S) = \sum_{j=0}^{m-1} \widetilde{nS_i^j}^k P_j^{(i)}(S). \tag{3.10}$$

The fluxes can be written as:

$$F_i^{\alpha,k} = \frac{1}{\Delta t} \sum_{j=0}^{m-1} \widetilde{nS_i^j}^k \int_{S_{i-1}}^{S_{i-1}+\Delta t} S^\alpha P_j^{(i)}(S)\,\mathrm{d}S, \tag{3.11}$$

and the scheme becomes:

$$\widetilde{nS_i^\alpha}^{k+1} = \sum_{j=0}^{m-1} \left(\widetilde{nS_i^j}^k \int_{S_{i-1}+\Delta t}^{S_i} (S-\Delta t)^\alpha P_j^{(i)}(S)\,\mathrm{d}S \; + \; \widetilde{nS_{i+1}^j}^k \int_{S_i}^{S_i+\Delta t} (S-\Delta t)^\alpha P_j^{(i+1)}(S)\,\mathrm{d}S \right). \tag{3.12}$$

3.1.2. *Preliminary result*

In order to estimate the order of the method, we need to evaluate the error made by the reconstruction of a regular function from its moments. So that the bounds are independent of the sections, the interval $[-1,1]$ and the basis $(R_0,\ldots,R_{m-1})$ of $\mathbb{R}_{m-1}[X]$ defined by (3.8) are used. The difference, at a point $y \in [-1,1]$, between a function h and its reconstruction by a polynomial function having the same m first moments can be written:

$$\mathfrak{D}(z \mapsto h(z), y) = h(y) - \sum_{j=0}^{m-1} R_j(y) \int_{-1}^{1} z^j\, h(z)\ \mathrm{d}z. \tag{3.13}$$

We then have the following estimate of this difference.

Lemma 3.3. *Let $h(y)$ be a C^m function for $y \in [-1,1]$. Then, for $y \in [-1,1]$:*

$$|\mathfrak{D}(z \mapsto h(z), y)| \leq \frac{H_m}{m!} \left[1 + \|R\| \sum_{j=1}^{m} \frac{2}{j} \right]$$

where H_m is the norm in $L^\infty([-1,1])$ of $h^{(m)}$ and $\|R\|$ is defined by

$$\|R\| = \sup_{j\in\{0,\ldots,m-1\}} \sup_{s\in[-1,1]} |R_j(s)|. \tag{3.14}$$

Proof. The Taylor-Lagrange formula, at the order m, between 0 and y, may be used to give the following expression of h: $h = Q + \epsilon$, where Q is a polynomial with a degree $m-1$ at most and where ϵ is such that, for $y \in [-1,1]$:

$$|\epsilon(y)| \leq \frac{H_m}{m!}. \tag{3.15}$$

We then express ψ in terms of ϵ. For that, let us expand the polynomial Q on the R_j basis:

$$Q(y) = \sum_{j=0}^{m-1} R_j(y) \int_{-1}^{1} u^j\, Q(u)\ \mathrm{d}u.$$

Inserting this in the expression of $\mathfrak{D}(z \mapsto h(z), y)$:

$$\mathfrak{D}(z \mapsto h(z), y) = h(y) - \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} R_i(y) \int_{-1}^{1} z^i \, R_j(z) \, \mathrm{d}z \int_{-1}^{1} u^j \, Q(u) \, \mathrm{d}u - \sum_{j=0}^{m-1} R_j(y) \int_{-1}^{1} z^j \, \epsilon(z) \, \mathrm{d}z.$$

Using the definition of the R_j given by (3.9), we can write:

$$\begin{aligned} \mathfrak{D}(z \mapsto h(z), y) &= h(y) - \sum_{j=0}^{m-1} R_j(y) \int_{-1}^{1} u^j \, Q(u) \, \mathrm{d}u - \sum_{j=0}^{m-1} R_j(y) \int_{-1}^{1} z^j \, \epsilon(z) \, \mathrm{d}z \\ &= \epsilon(y) - \sum_{j=0}^{m-1} R_j(y) \int_{-1}^{1} z^j \, \epsilon(z) \, \mathrm{d}z. \end{aligned}$$

We then obtain, with (3.15), the estimate of $\mathfrak{D}(z \mapsto h(z), y)$ given in the lemma. □

This result is essential to determine the order for the evaluation of the droplet mass density by using the reconstructed distribution function.

3.1.3. *Calculation of the mass*

An approximation $\tilde{m}_i$ of the droplet mass density m_i in the section i is given by:

$$\tilde{m}_i^k = \int_{S_{i-1}}^{S_i} S^{3/2} \, f_a^k(S) \, \mathrm{d}S = \sum_{j=0}^{m-1} \widetilde{nS_i^j}^k \int_{S_{i-1}}^{S_i} S^{3/2} \, P_j^{(i)}(S) \, \mathrm{d}S. \tag{3.16}$$

If the values of the moments are exact, then the error made with the preceding formula is given by:

$$\zeta_i^k = m_i^k - \sum_{j=0}^{m-1} \overline{nS_i^j}^k \int_{S_{i-1}}^{S_i} S^{3/2} \, P_j^{(i)}(S) \, \mathrm{d}S, \tag{3.17}$$

that is to say:

$$\zeta_i^k = \int_{S_{i-1}}^{S_i} S^{3/2} g_i^k(S) \, \mathrm{d}S, \qquad g_i^k(S) = f^k(S) - \sum_{j=0}^{m-1} \overline{nS_i^j}^k \, P_j^{(i)}(S). \tag{3.18}$$

One may now determine the order of ζ_i^k as a function of ΔS:

Theorem 3.4. *If $f_0(S)$ is a C^m function for $S \in \mathbb{R}^+$, then $\zeta_i^k = O\left(\Delta S^{m+1}\right)$.*

Proof. We first derive an expression of the g_i^k and apply Lemma 3.3 in a second step. It is easy to see that the m first moments of $g_i^k(S)$ on $[S_{i-1}, S_i]$ are equal to zero. So, the m first moments of the function $g_i^k(S_{i-1/2} + y \frac{\Delta S}{2})$ of $y \in [-1, 1]$ are equal to zero. Moreover, $(f^k - g_i^k)(S_{i-1/2} + y \frac{\Delta S}{2})$, as a polynomial function of y, can be expanded on the R_j basis:

$$(f^k - g_i^k)\left(S_{i-1/2} + y \frac{\Delta S}{2}\right) = \sum_{j=0}^{m-1} R_j(y) \int_{-1}^{1} z^j \, f^k\left(S_{i-1/2} + z \frac{\Delta S}{2}\right) \mathrm{d}z$$

and the g_i^k can then be written:

$$g_i^k(S) = \mathfrak{D}\left(z \mapsto f^k\left(S_{i-1/2} + z \frac{\Delta S}{2}\right), 2 \frac{S - S_{i-1/2}}{\Delta S}\right). \tag{3.19}$$

Lemma 3.3 is applied to g_i^k, with $h(y) = f^k(S_{i-1/2} + y\,\frac{\Delta S}{2})$. Having $f^k(S) = f_0(t_k + S)$ and with $M_m = \sup_{y\in\mathbb{R}^+} |f_0^{(m)}(y)|$, the $L^\infty([-1,1])$ norm of $h^{(m)}$ is smaller than $\Delta S^m\, M_m\, /2^m$. Then:

$$\left|g_i^k(S)\right| \leq \frac{M_m}{2^m\, m!}\left[1 + \|R\| \sum_{j=1}^{m} \frac{2}{j}\right](\Delta S)^m,$$

where $\|R\|$ is defined by (3.14). The ζ_i^k are given, as a function of the g_i^k by (3.18). The result of the theorem is then obtained:

$$|\zeta_i^k| \leq \int_{S_{i-1}}^{S_i} S^{3/2}|g_i^k(S)|\,\mathrm{d}S \leq \frac{M_m}{2^m\, m!}\left[1 + \|R\| \sum_{j=1}^{m} \frac{2}{j}\right]\Delta S^{m+1}. \tag{3.20}$$

□

3.1.4. *Order of the method*

The order of the method providing the moments $\overline{nS_i^\alpha}$ may now be determined. We introduce the local truncation errors for the scheme (3.12), for $\alpha \in \{0,\ldots,m-1\}$ and $i \in \{0,\ldots,N\}$:

$$\epsilon_i^{\alpha,k} = \frac{1}{\Delta t}\left[\overline{nS_i^\alpha}^{k+1} - \sum_{j=0}^{m-1}\left(\overline{nS_i^j}^{k}\int_{S_{i-1}+\Delta t}^{S_i}(S-\Delta t)^\alpha P_j^{(i)}(S)\,\mathrm{d}S\right.\right. \tag{3.21}$$
$$\left.\left. +\overline{nS_{i+1}^j}^{k}\int_{S_i}^{S_i+\Delta t}(S-\Delta t)^\alpha P_j^{(i+1)}(S)\,\mathrm{d}S\right)\right].$$

Using the definition (3.1) given for the $\overline{nS_i^\alpha}^k$ and the transport equation for f, we obtain:

$$\overline{nS_i^\alpha}^{k+1} = \int_{S_{i-1}}^{S_i} S^\alpha f^k(S+\Delta t)\,\mathrm{d}S = \int_{S_{i-1}+\Delta t}^{S_i+\Delta t}(S-\Delta t)^\alpha f^k(S)\,\mathrm{d}S$$

and then:

$$\epsilon_i^{\alpha,k} = \frac{1}{\Delta t}\left(\int_{S_i}^{S_i+\Delta t}(S-\Delta t)^\alpha g_{i+1}^k(S)\,\mathrm{d}S - \int_{S_{i-1}}^{S_{i-1}+\Delta t}(S-\Delta t)^\alpha g_i^k(S)\,\mathrm{d}S\right), \tag{3.22}$$

where g_i^k is the function introduced by the equation (3.18). Let us remark that the moments of order between 0 and $m-1$ of g_i^k over $[S_{i-1}, S_i[$ vanish, so that we can replace the integral over $[S_{i-1}+\Delta t, S_i]$ by an integral over $[S_{i-1}, S_{i-1}+\Delta t]$.

In the case where the distribution is regular, the following consistency property holds:

Theorem 3.5. *If $f_0(S)$ is a C^{m+1} function for $S \in \mathbb{R}^+$, then $\epsilon_i^{\alpha,k} = O\left(\Delta S^{m+1}\right)$.*

Proof. We want to introduce a C^1 function $h^{\alpha,k}(s)$, such that the local truncation error can be written

$$\epsilon_i^{\alpha,k} = h^{\alpha,k}(S_{i+1/2}) - h^{\alpha,k}(S_{i-1/2}).$$

In the proof of the preceding proposition, we gave the expression (3.19) of $g_i^k(S)$. Replacing it in the second term of the expression between brackets in equation (3.22) and using a change of variables, we obtain:

$$\int_{S_{i-1}}^{S_{i-1}+\Delta t}(S-\Delta t)^\alpha g_i^k(S)\,\mathrm{d}S = \frac{\Delta S}{2}\int_{-1}^{-1+2\Delta t/\Delta S}\left(S_{i-1/2} + \frac{\Delta S}{2}y - \Delta t\right)^\alpha$$
$$\times\left[f^k\left(S_{i-1/2} + \frac{\Delta S}{2}y - \Delta t\right) - \sum_{j=0}^{m-1} R_j(y)\int_{-1}^{1} z^j f^k\left(S_{i-1/2} + \frac{\Delta S}{2}z - \Delta t\right)\right]\mathrm{d}y.$$

Let $\nu = \Delta t/\Delta S$ represent the constant ratio between the two discretization steps. A function $h^{\alpha,k}$ satisfying the previous property can then be given by:

$$h^{\alpha,k}(S) = \frac{1}{2\nu}\int_{-1}^{-1+2\nu}\left(S+\frac{\Delta S}{2}y-\Delta t\right)^{\alpha}\mathfrak{D}\left(z\mapsto f^k\left(S+\frac{\Delta S}{2}z-\Delta t\right),y\right)\,\mathrm{d}y. \tag{3.23}$$

The $\epsilon_i^{\alpha,k}$, given above can be written: $\epsilon_i^{\alpha,k} = \Delta S\,(h^{\alpha,k})'(S_c)$ where $S_c \in]S_{i-1/2}, S_{i+1/2}[$, that is to say

$$\begin{aligned}\epsilon_i^{\alpha,k} = \frac{\Delta S}{2\nu}\Bigg\{&\int_{-1}^{-1+2\nu}\alpha\left(S_c+\frac{\Delta S}{2}y-\Delta t\right)^{\alpha-1}\mathfrak{D}\left(z\mapsto f^k\left(S_c+\frac{\Delta S}{2}z-\Delta t\right),y\right)\,\mathrm{d}y\\ &+\int_{-1}^{-1+2\nu}\left(S_c+\frac{\Delta S}{2}y-\Delta t\right)^{\alpha}\mathfrak{D}\left(z\mapsto (f^k)'\left(S_c+\frac{\Delta S}{2}z-\Delta t\right),y\right)\,\mathrm{d}y\Bigg\}.\end{aligned}$$

Applying Lemma 3.3 twice, we obtain upper bounds for terms between brackets. This yields an upper bound of $|\epsilon_i^{\alpha,k}|$, using the notations introduced in equation (3.14):

$$|\epsilon_i^{\alpha,k}| \le \left[1+\|R\|\sum_{j=1}^{m}\frac{2}{j}\right]\left(\alpha\sup_{y\in\mathbb{R}^+}|f_0^{(m)}(y)|+\sup_{y\in\mathbb{R}^+}|f_0^{(m+1)}(y)|\right)\frac{\Delta S^{m+1}}{2^m m!},$$

and the proposition is proved. □

So, in the case of a regular function f_0, the method is mth order consistent.

3.1.5. *Stability*

Stability is defined in this analysis in the sense of Lax and Richtmyer (see for example [18]). Then, if $H_{\Delta t}$ is the $Nm \times Nm$ matrix of the numerical method for a fixed ratio $\nu = \Delta t/\Delta S$ of the discretization steps, the method is stable if and only if for each time T, there is a constant C and a value $j > 0$ such that

$$\|H_{\Delta t}^k\| \le C \qquad \text{for all} \quad k\Delta t \le T, \quad \Delta t < j.$$

A sufficient condition to have this inequality is that the spectral radius of the matrix $H_{\Delta t}$ be smaller than one since, for a matrix M and a induced norm N, the spectral radius of M is such that: $\rho(M) = \lim_{j\to\infty}\left[N(M^j)\right]^{1/j}$.

Here, a variable change can be made to simplify the shape of the scheme matrix. Let $\widetilde{\delta^{\alpha} n_i}^{\,k}$ denote the quantity:

$$\widetilde{\delta^{\alpha} n_i}^{\,k} = \frac{2^\alpha}{(\Delta S)^\alpha}\sum_{j=0}^{\alpha}C_\alpha^j(-S_{i-1/2})^{\alpha-j}\widetilde{nS_i^j}^{\,k} = \int_{S_{i-1}}^{S_i}\left(2\frac{S-S_{i-1/2}}{\Delta S}\right)^{\alpha}f_a^k(S)\,\mathrm{d}S. \tag{3.24}$$

The approximated function f_a^k can be written, with these new variables:

$$\forall S\in[S_{i-1},S_i[\qquad f_a^k(S) = \frac{2}{\Delta S}\sum_{j=0}^{m-1}\widetilde{\delta^j n_i}^{\,k}R_j\left(2\frac{S-S_{i-1/2}}{\Delta S}\right). \tag{3.25}$$

The scheme for the new variables is written:

$$\widetilde{\delta^{\alpha} n_i}^{\,k+1} = \sum_{j=0}^{m-1}\widetilde{\delta^j n_i}^{\,k}\int_{-1+2\nu}^{1}(z-2\nu)^\alpha R_j(z)\,\mathrm{d}z + \sum_{j=0}^{m-1}\widetilde{\delta^j n_{i+1}}^{\,k}\int_{-1}^{-1+2\nu}(z+2(1-\nu))^\alpha R_j(z)\,\mathrm{d}z.$$

Let $X^k = (\widetilde{n_1}^k, \widetilde{\delta n_1}^k, \ldots, \widetilde{\delta^{m-1} n_1}^k, \ldots, \widetilde{n_N}^k, \widetilde{\delta n_N}^k, \ldots, \widetilde{\delta^{m-1} n_N}^k)^t$ the vector of unknown variables. This vector is such that $X^{k+1} = H_{\Delta t}\, X^k$ where:

$$H_{\Delta t} = \begin{pmatrix} A & B & 0 & \ldots & 0 \\ 0 & A & B & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & A & B \\ 0 & \ldots & \ldots & 0 & A \end{pmatrix}, \qquad \begin{aligned} A_{i,j} &= \int_{-1+2\nu}^{1} (z-2\nu)^{i-1} R_{j-1}(z)\, \mathrm{d}z \\ B_{i,j} &= \int_{-1}^{-1+2\nu} (z+2(1-\nu))^{i-1} R_{j-1}(z)\, \mathrm{d}z. \end{aligned} \tag{3.26}$$

Then, a sufficient stability condition is $\rho(A) < 1$. Here, stability will not be proved in the general case. Only the two particular cases $m = 2$ and $m = 3$ which will be used later are considered:

Theorem 3.6. *For $\nu \in [0,1]$, the scheme (3.12) corresponding to $m = 2$ or $m = 3$ is stable.*

Proof. As shown above, we just have to check that $\rho(A) < 1$ in each case. In the case $m = 2$, the polynomials R_0 and R_1 are:

$$R_0 = \frac{1}{2}, \qquad R_1 = \frac{3}{2} X$$

and the matrix A is:

$$A = (1-\nu) \begin{pmatrix} 1 & 3\nu \\ -\nu & 1 - 2\nu - 2\nu^2 \end{pmatrix}.$$

In the case $m = 3$, the polynomials R_0, R_1 and R_2 are:

$$R_0 = \frac{3}{8}(3 - 5X^2), \qquad R_1 = \frac{3}{2} X \qquad R_2 = \frac{15}{8}(-1 + 3X^2).$$

The matrix A is here:

$$A = \frac{1-\nu}{2} \begin{pmatrix} 2 + 5\nu - 10\nu^2 & 6\nu & 15\nu(-1+2\nu) \\ \nu(-7 + 5\nu + 5\nu^2) & 2(1 - 2\nu - 2\nu^2) & 15\nu(1 - \nu - \nu^2) \\ \nu(3 + 2\nu - 4\nu^2 - 4\nu^3) & 2\nu(-1 + 2\nu + 2\nu^2) & 2 - 13\nu + 2\nu^2 + 12\nu^3 + 12\nu^4 \end{pmatrix}.$$

The modulus of all the eigenvalues of the matrix A corresponding for both cases are plotted in Figure 1, as a function of the ratio ν. These quantities are all less unity thus concluding the proof. □

Remark 3.7. The Theorem is limited to the cases $m = 2$ and $m = 3$ used in this paper because for the larger values of m the analytical calculation of the spectral radius is not feasible. However, for all the values of m until $m = 10$, numerical calculation of the spectral radius as a function of ν has been done on a refined discretization of $[0, 1]$. It shows that the scheme should be stable at least for $m \leq 10$.

Remark 3.8. The variables $\widetilde{\delta^\alpha n_i}^k$ could be used to describe and compute the scheme. But, for a better generalization in the case where droplet velocities, and then their drag force, are taken into account, the natural variables are the $\widetilde{nS_i^\alpha}^k$ because the drag force is proportional to $1/S$.

This last result, combined with the consistency Theorem, proves that the method is mth order accurate, at least for $m = 2$ and $m = 3$. However, the polynomial functions used to reconstruct the distribution function are not necessarily positive as soon as $m \geq 2$. This could yield negative mass densities. To analyze this point, we examine the special case of the 2 moments method.

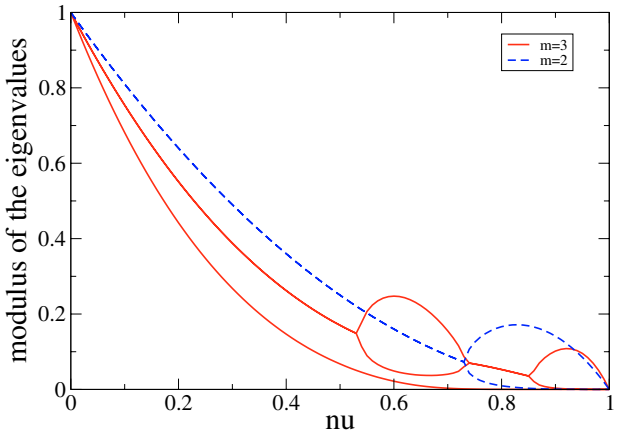

FIGURE 1. Modulus of the eigenvalues of the matrix A, as a function of the parameter ν, for $m = 3$ (solid lines) and for $m = 2$ (dashed lines).

3.2. **2 moments method**

The purpose of this section is to examine the 2 moments method, giving an expression of the approximated distribution function, the fluxes and the mass densities, and to determine positivity conditions.

3.2.1. *Numerical scheme for the 2 moments method*

The approximated distribution function defined in the preceding subsection and given by (3.25) is here of the form:

$$\forall S \in [S_{i-1}, S_i[\qquad f_a^k(S) = \frac{\widetilde{n_i}^k}{\Delta S} + \frac{3\,\tilde{\delta}_i^k \widetilde{n_i}^k}{\Delta S} \times \frac{2}{\Delta S}(S - S_{i-1/2}) \tag{3.27}$$

where the parameter $\tilde{\delta}_i^k$ characterizes the distance between the averaged surface $\widetilde{nS_i}^k / \widetilde{n_i}^k$ in the section $[S_{i-1}, S_i[$ and the mid point $S_{i-1/2}$ of this section:

$$\widetilde{nS_i}^k = S_{i-\frac{1}{2}} \widetilde{n_i}^k + \frac{\Delta S}{2} \tilde{\delta}_i^k \widetilde{n_i}^k. \tag{3.28}$$

This parameter can be related to the variable $\widetilde{\delta n_i}^k$ defined by (3.24): $\widetilde{\delta n_i}^k = \tilde{\delta}_i^k \widetilde{n_i}^k$. It is assumed that if $\widetilde{n_i}^k = 0$, then $\widetilde{\delta n_i}^k = 0$ and we then choose $\tilde{\delta}_i^k = 0$.

The numerical scheme can then be written:

$$\begin{cases} \widetilde{n_i}^{k+1} = \widetilde{n_i}^k + \Delta t \left(F_{i+1}^{0,k} - F_i^{0,k} \right), \\ \widetilde{nS_i}^{k+1} = \widetilde{nS_i}^k - \Delta t\, \widetilde{n_i}^k + \Delta t \left(F_{i+1}^{1,k} - F_i^{1,k} \right) - (\Delta t)^2 \left(F_{i+1}^{0,k} - F_i^{0,k} \right), \end{cases} \tag{3.29}$$

with

$$F_i^{0,k} = \frac{\widetilde{n_i}^k}{\Delta S} - 3(1-\nu)\frac{\tilde{\delta}_i^k \widetilde{n_i}^k}{\Delta S},$$
$$F_i^{1,k} = \left(S_{i-1} + \frac{\nu}{2}\Delta S \right) \frac{\widetilde{n_i}^k}{\Delta S} + 3\left[-(1-\nu) S_{i-1} + \frac{\nu}{6}(4\nu - 3)\Delta S \right] \frac{\tilde{\delta}_i^k \widetilde{n_i}^k}{\Delta S},$$

and with the ratio $\nu = \Delta t / \Delta S$.

The mass density is given by:

$$\tilde{m}_i^k = \frac{2}{5}\frac{\widetilde{n_i}^k}{\Delta S}\left(\frac{S_i^{5/2}-S_{i-1}^{5/2}}{\Delta S}\right) + \frac{6\,\tilde{\delta}_i^k \widetilde{n_i}^k}{35\Delta S}\left(3\frac{S_i^{7/2}-S_{i-1}^{7/2}}{\Delta S} - 7S_{i-1}S_i\frac{S_i^{3/2}-S_{i-1}^{3/2}}{\Delta S}\right).$$

3.2.2. *Positivity*

Assuming that the moments $\widetilde{n_i}^k$ and $\widetilde{nS_i}^k$ are non negative, the non negativity of the mass density $\tilde{m}_i^k$ and of the moments $\widetilde{n_i}^{k+1}$ and $\widetilde{nS_i}^{k+1}$ is ensured if f_a^k is a non negative function. The expression (3.27) of this function proves that the following condition is needed: $\tilde{\delta}_i^k \in [-1/3, 1/3]$, where the parameter $\tilde{\delta}_i^k$ is defined by (3.28). We determine a sufficient condition so that this property can be ensured at least for the exact solution. Let us then assume that $\widetilde{n_i}^k$ and $\widetilde{nS_i}^k$ are the moments of the exact solution $f^k(S) = f_0(S + t_k)$. In this case, $\tilde{\delta}_i^k$ is given by:

$$\widetilde{n_i}^k \tilde{\delta}_i^k = \frac{2}{\Delta S}\left(\widetilde{nS_i}^k - S_{i-1/2}\widetilde{n_i}^k\right) = \int_{S_{i-1}}^{S_i} \frac{2}{\Delta S}(S - S_{i-1/2}) f^k(S)\,\mathrm{d}S.$$

It is then easy to see that $\tilde{\delta}_i^k \in]-1, 1[$. A Taylor expansion of f^k around $S_{i-1/2}$ proves that $\tilde{\delta}_i^k$ behaves as $\frac{\Delta S\,(f^k)'(S_{i-1/2})}{6 f^k(S_{i-1/2})}$, which is small if $f^k(S_{i-1/2})$ is "not too large". The following proposition gives a more precise estimate and a sufficient condition to obtain $\tilde{\delta}_i^k \in [-1/3, 1/3]$.

Proposition 3.9. *If f_0 is a positive C^1 function over $\mathbb{R}^+$ and if there exists a positive constant K such that:*

$$\forall S \geq 0 \qquad |f_0'(S)| \leq K f_0(S), \tag{3.30}$$

then, for $3K\Delta S \leq 2$, the following inequality holds:

$$\forall i \in \{1, \ldots, N\} \quad \forall n \in \mathbb{N} \qquad \frac{2}{\Delta S}\left|\int_{S_{i-1}}^{S_i} (S - S_{i-1/2}) f^k(S)\,\mathrm{d}S\right| \leq \frac{1}{3}\left|\int_{S_{i-1}}^{S_i} f^k(S)\,\mathrm{d}S\right|. \tag{3.31}$$

Proof. For k fixed, let us denote $g(x) = f^k(S_{i-1/2} + x\Delta S/2) = f_0(t_k + S_{i-1/2} + x\Delta S/2)$, for $x \in [-1, 1]$. Condition (3.30) induces, for g:

$$\forall x \in [-1, 1] \qquad |g'(x)| \leq \frac{\Delta S}{2} K g(x). \tag{3.32}$$

If g is the null function over $[-1, 1]$, then the inequality (3.31) is immediate. In the other case, $\int_{S_{i-1}}^{S_i} f^k(S)\,\mathrm{d}S$ differs from zero and we can define:

$$R = \frac{2}{\Delta S}\left|\frac{\int_{S_{i-1}}^{S_i}(S - S_{i-1/2}) f(t, S)\,\mathrm{d}S}{\int_{S_{i-1}}^{S_i} f(t, S)\,\mathrm{d}S}\right| = \left|\frac{\int_{-1}^{1} x g(x)\,\mathrm{d}x}{\int_{-1}^{1} g(x)\,\mathrm{d}x}\right|.$$

An integration by parts and then, a use of the Rolle Theorem gives:

$$R = \left|\frac{g(1) - g(-1) - \int_{-1}^{1} x^2 g'(x)\,\mathrm{d}x}{2\int_{-1}^{1} g(x)\,\mathrm{d}x}\right| = \left|\frac{g'(c)}{2g(c)} - \frac{\int_{-1}^{1} x^2 g'(x)\,\mathrm{d}x}{2\int_{-1}^{1} g(x)\,\mathrm{d}x}\right|,$$

where c is a value in the interval $]-1, 1[$. The preceding inequality (3.32) is then applied:

$$R \leq \frac{K\Delta S}{2}$$

and the proposition is proved. □

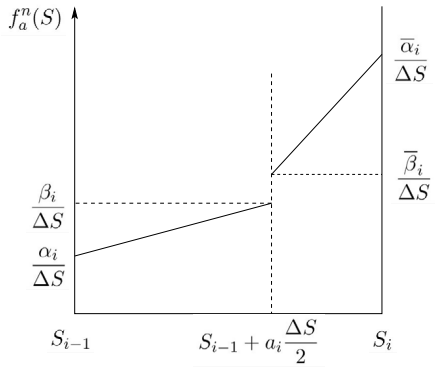

FIGURE 2. Graph of the function $f_a^n(S)$ in the section i.

Condition (3.30) induces that, for a decreasing function f_0, then $f_0(S) \geq \exp(-KS)$ for all $S \geq 0$. In particular, f_0 cannot vanish. For a function with a compact support, condition (3.31) cannot be satisfied, even for a refined discretization, so that the reconstructed function f_a^k can have negative values in sections where f^k is small.

Since the positivity condition cannot always be satisfied we introduce a modified scheme that has this property.

3.3. **The modified 2 moments method**

The solution proposed in this section is to use another profile for the distribution function: a function which is affine in two parts $[S_{i-1}\ ,\ S_{i-1}+a_i\Delta S/2[$ and $[S_{i-1}+a_i\Delta S/2\ ,\ S_i[$ composing the section. So, for $x \in [-1, 1[$, we define:

$$f_a^k\left(S_{i-1/2} + x\frac{\Delta S}{2}\right) = \frac{1}{\Delta S}\begin{cases} \alpha_i + (\beta_i - \alpha_i)\dfrac{1+x}{1+a_i} & \text{if } x \in [-1, a_i[, \\ \overline{\alpha}_i + (\overline{\beta}_i - \overline{\alpha}_i)\dfrac{1-x}{1-a_i} & \text{if } x \in [a_i, 1[, \end{cases} \tag{3.33}$$

with $-1 < a_i \leq 1$. This function is plotted in Figure 2. The total variation of f_a^k in the section is not artificially increased, we impose that this function is monotonic over each section $[S_{i-1}, S_i[$ (increasing if $\tilde{\delta}_i^k$ is positive and decreasing in the other case). For that, the values of $\beta_i - \alpha_i$, $\overline{\beta}_i - \beta_i$ and $\overline{\alpha}_i - \overline{\beta}_i$ are chosen proportional to $\tilde{\delta}_i^k\tilde{n}_i^k$, with a non-negative constant of proportionality. Four parameters are then introduced: ζ_i, $\lambda_i \geq -1$, $\overline{\lambda}_i \geq -1$ and $\mu_i \geq 0$ such that

$$\alpha_i = \zeta_i - \left[\frac{\mu_i}{2} + 3(1+a_i)(1+\lambda_i)\right]\tilde{\delta}_i^k\tilde{n}_i^k, \qquad \beta_i = \zeta_i - \frac{\mu_i}{2}\tilde{\delta}_i^k\tilde{n}_i^k,$$

$$\overline{\alpha}_i = \zeta_i + \left[\frac{\mu_i}{2} + 3(1-a_i)(1+\overline{\lambda}_i)\right]\tilde{\delta}_i^k\tilde{n}_i^k, \qquad \overline{\beta}_i = \zeta_i + \frac{\mu_i}{2}\tilde{\delta}_i^k\tilde{n}_i^k.$$

Then $\zeta_i/\Delta S$ represents the value of the midpoint of the discontinuity at $S = S_{i-1} + a_i\,\Delta S/2$, $\mu_i\tilde{\delta}_i^k\tilde{n}_i^k/\Delta S$ is the value of this discontinuity and the coefficients are chosen such that, for $\lambda_i = \overline{\lambda}_i = \mu_i = 0$, the affine case is recovered.

In order that the conditions (3.2) are fulfilled, the parameters ζ_i and μ_i have the following values:

$$\zeta_i = \tilde{n}_i^k + \tilde{\delta}_i^k \tilde{n}_i^k \left[3a_i + (3 - 4a_i - a_i^2) \frac{(1+a_i)}{4(1-a_i)} \lambda_i - (3 + 4a_i - a_i^2) \frac{(1-a_i)}{4(1+a_i)} \overline{\lambda}_i \right],$$
$$\mu_i = -\frac{1+a_i}{1-a_i}(2-a_i)\lambda_i - \frac{1-a_i}{1+a_i}(2+a_i)\overline{\lambda}_i.$$

α_i, β_i, $\overline{\beta}_i$ and $\overline{\alpha}_i$ can be written, in terms of the three parameters a_i, λ_i, $\overline{\lambda}_i$:

$$\alpha_i = \tilde{n}_i^k + \tilde{\delta}_i^k \tilde{n}_i^k \left[-3 - (1+a_i)(5-a_i)\frac{\lambda_i}{4} + \frac{(1-a_i)^3}{4(1+a_i)} \overline{\lambda}_i \right],$$
$$\beta_i = \tilde{n}_i^k + \tilde{\delta}_i^k \tilde{n}_i^k \left[3a_i + (1+a_i)(7+a_i)\frac{\lambda_i}{4} + \frac{(1-a_i)^3}{4(1+a_i)} \overline{\lambda}_i \right],$$
$$\overline{\beta}_i = \tilde{n}_i^k + \tilde{\delta}_i^k \tilde{n}_i^k \left[3a_i - \frac{(1+a_i)^3}{4(1-a_i)} \lambda_i - (1-a_i)(7-a_i)\frac{\overline{\lambda}_i}{4} \right],$$
$$\overline{\alpha}_i = \tilde{n}_i^k + \tilde{\delta}_i^k \tilde{n}_i^k \left[3 - \frac{(1+a_i)^3}{4(1-a_i)} \lambda_i + (1-a_i)(5+a_i)\frac{\overline{\lambda}_i}{4} \right].$$

In this section, the obtained scheme is first described as a function of the parameter a_i, λ_i and $\overline{\lambda}_i$. Second, we examine conditions on these parameters that ensure the second order consistency of the scheme (Prop. 3.10) and its stability (Prop. 3.11). A reconstruction algorithm is then given together with the choice for the remaining parameters and then for the approximated distribution function so that stability and positivity conditions are satisfied and the condition for the second order is verified as often as possible. Finally, a special reconstruction is devised for the discontinuous case: a regularity condition is given which can be interpreted as a slope limiter and the approximated distribution function is then a step function in the concerned section. This last improvement is done in order to limit the total variation of the function.

3.3.1. *Scheme for the modified 2 moments method*

The scheme is still given by (3.29) where for if $\nu \leq \frac{1+a_i}{2}$ (*i.e.* $S_{i-1} + \Delta t \leq S_{i-\frac{1}{2}} + a_i \frac{\Delta S}{2}$), the fluxes are such that:

$$\Delta t\, F_i^{0,k} = \alpha_i \nu + (\beta_i - \alpha_i) \frac{\nu^2}{1+a_i},$$
$$\Delta t\, F_i^{1,k} - (\Delta t)^2 F_i^{0,k} = \alpha_i \nu \left(S_{i-1} - \frac{\nu}{2} \Delta S \right) + (\beta_i - \alpha_i) \frac{\nu^2}{1+a_i} \left(S_{i-1} - \frac{\nu}{3} \Delta S \right),$$

and the mass densities are:

$$\tilde{m}_i^k = \frac{1+a_i}{2} \left[\beta_i\, C(\sqrt{S_{i-1}}, \sqrt{S_a}) + (\alpha_i - \beta_i)\, K(\sqrt{S_{i-1}}, \sqrt{S_a}) \right] + \frac{1-a_i}{2} \left[\overline{\beta}_i\, C(\sqrt{S_a}, \sqrt{S_i}) - (\overline{\alpha}_i - \overline{\beta}_i)\, K(\sqrt{S_a}, \sqrt{S_i}) \right],$$

where the functions C and K are defined by:

$$C(a,b) = \frac{1}{b^2 - a^2} \int_{a^2}^{b^2} S^{3/2}\, \mathrm{d}S = \frac{2}{5} \frac{b^4 + ab^3 + a^2b^2 + a^3b + a^4}{b+a},$$
$$K(a,b) = \frac{1}{(b^2 - a^2)^2} \int_{a^2}^{b^2} S^{3/2} (S - a^2)\, \mathrm{d}S = \frac{2}{35} \frac{5b^5 + 10ab^4 + 8a^2b^3 + 6a^3b^2 + 4a^4b + 2a^5}{(b+a)^2}.$$

3.3.2. *Order of the method*

Since the approximated distribution function f_a^k is not necessarily affine in each section, the scheme will not necessarily provide a second order accurate approximation of the distribution function for a L^∞ norm, assuming that the moments are exact. However, we are only interested in some moments of this function. And then, to ensure second order accuracy, some additional conditions are needed in a regular case.

Proposition 3.10. *Assume that f_0 is a positive C^3 function over $\mathbb{R}^+$. The local truncation errors of the previously defined scheme are some $O(\Delta S^3)$ if λ_i is a constant independent of i and if $\overline{\lambda}_i$ is such that:*

$$\overline{\lambda}_i = \frac{1+a_i}{(1-a_i)^3}[(1+a_i)(5-a_i) - 12\nu]\lambda_i. \tag{3.34}$$

In this case, the scheme is then second order consistent.

Proof. As f_0, f^k is a C^3 function over $\mathbb{R}^+$ and some Taylor expansions yield the following expressions for the local truncation errors $\epsilon_i^{0,k}$ and $\epsilon_i^{1,k}$:

$$\epsilon_i^{0,k} = (p_i - p_{i+1})\frac{\Delta S}{6}(f^k)'(S_i) - (p_i + p_{i+1})\frac{(\Delta S)^2}{12}(f^k)''(S_i) + O((\Delta S)^3),$$
$$\epsilon_i^{1,k} - S_i\epsilon_i^{0,k} = \left[-p_i + \frac{\nu}{2}(p_{i+1} - p_i) - \frac{\nu^2}{2}(\lambda_{i+1} - \lambda_i)\right]\frac{(\Delta S)^2}{6}(f^k)'(S_i) + O((\Delta S)^3),$$

where p_i is defined by:

$$p_i = \frac{\lambda_i}{4}[12\nu - (1+a_i)(5-a_i)] + \frac{(1-a_i)^3}{4(1+a_i)}\overline{\lambda}_i. \tag{3.35}$$

Then the global second order of the scheme is ensured if the parameters p_i are equal to zero and if all the λ_i are the same, concluding the proof. □

3.3.3. *Stability*

To prove the stability, a change of variables was introduced in Section 3.1.5. The scheme can then be written:

$$(\widetilde{n_i}^{k+1}, \widetilde{\delta n_i}^{k+1})^t = A_i(\widetilde{n_i}^k, \widetilde{\delta n_i}^k)^t + B_i(\widetilde{n_{i+1}}^k, \widetilde{\delta n_{i+1}}^k)^t \tag{3.36}$$

with, if p_i is defined by (3.35),

$$A_i = \begin{pmatrix} 1-\nu & 3\nu(1-\nu) - \nu p_i \\ -\nu(1-\nu) & (1-\nu)(1-2\nu-2\nu^2) - \nu^3\lambda_i + \nu(1-\nu)p_i \end{pmatrix}$$

$$B_i = \nu \begin{pmatrix} 1 & -3(1-\nu) + p_{i+1} \\ 1-\nu & -3+6\nu-2\nu^2+\nu^2\lambda_{i+1} + (1-\nu)p_{i+1} \end{pmatrix}.$$

We just have to evaluate the spectral radius of A_i to give a sufficient stability condition.

Proposition 3.11. *For $\nu \in [0, 1/2]$, the scheme defined by (3.36) and satisfying the condition (3.34) is stable if, for all $i \in \{1, \ldots, N\}$, $\lambda_i \leq 7$.*

Proof. The spectral radius of A_i is smaller than one if and only if the two following conditions are fulfilled: $\det(A_i) < 1$ and $|\operatorname{tr}(A_i)| < 1 + \det(A_i)$. This can be proved, considering that $\operatorname{tr}(A_i)$ and $\det(A_i)$ are two real number respectively representing the sum and the product of the two eigenvalues of A_i. Since

$$\operatorname{tr}(A_i) = 2(1-\nu)(1-\nu-\nu^2) - \nu^3\lambda_i + \nu(1-\nu)p_i$$

and

$$\det(A_i) = (1-\nu)^4 - \nu^3(1-\nu)\lambda_i + \nu(1-\nu)p_i,$$

it is easy to show that the previous conditions are satisfied for $p_i = 0$ and $\lambda_i \leq 7$. □

3.3.4. *Reconstruction algorithm*

As a consequence of Propositions 3.10 and 3.11, to ensure the third local order of convergence, the values of the parameters α_i, β_i, $\overline{\alpha}_i$, $\overline{\beta}_i$, μ_i and $\overline{\lambda}_i$ have to be:

$$\begin{aligned}
\alpha_i &= \tilde{n}_i^k - 3\tilde{\delta}_i^k \tilde{n}_i^k \left[1 + \nu\lambda_i\right] \\
\beta_i &= \tilde{n}_i^k + 3\tilde{\delta}_i^k \tilde{n}_i^k \left[a_i + (1 - \nu + a_i)\lambda_i\right] \\
\overline{\beta}_i &= \tilde{n}_i^k + 3\tilde{\delta}_i^k \tilde{n}_i^k \left\{a_i + \frac{1+a_i}{(1-a_i)^2}[\nu(7-a_i) - (3-a_i)(1+a_i)]\lambda_i\right\} \\
\overline{\alpha}_i &= \tilde{n}_i^k + 3\tilde{\delta}_i^k \tilde{n}_i^k \left\{1 + \frac{1+a_i}{(1-a_i)^2}[2(1+a_i) - \nu(5+a_i)]\lambda_i\right\} \\
\mu_i &= -\frac{12\lambda_i}{(1-a_i)^2}[1 + a_i - \nu(2+a_i)] \\
\overline{\lambda}_i &= \frac{1+a_i}{(1-a_i)^3}[(1+a_i)(5-a_i) - 12\nu]\lambda_i,
\end{aligned}$$

where λ_i have to be independent of i and smaller than 7. One wishes that the function f_a^k be non negative and then that the parameters α_i, $\overline{\alpha}_i$, and μ_i are non negative and that λ_i and $\overline{\lambda}_i$ are greater than -1. Finally, one also wishes to be as close as possible to the linear case in each section. That would impose that $|\lambda_i|$ would be as small as possible. A change of the value of λ_i between two neighboring sections reduces to 2 the local order of the method, but it has not a lot of influence on the global order if it only concerns a few sections or if the amplitude of its changes is small.

We now describe the reconstruction algorithm that allows to preserve the positivity and, as often as possible, the third local order:

- For $|\tilde{\delta}_i^k| \leq \frac{1}{3}$, the reconstruction is linear: $\lambda_i = 0$.
- For $\frac{1}{3} < \tilde{\delta}_i^k \leq \frac{1}{3(1-\nu)}$, the smallest possible value for $|\lambda_i|$ is chosen (it corresponds to $\alpha_i = 0$):

$$\lambda_i = -\frac{1}{\nu}\left(1 - \frac{1}{3\tilde{\delta}_i^k}\right),$$

and a_i is chosen to minimize the total variation in the section: $a_i = \frac{2-3\nu}{3(1-\nu)}$ if $\tilde{\delta}_i^k = 1/3(1-\nu)$ and otherwise, $a_i = 1 - 2(u_+ + u_-)$, with

$$u_\pm^3 = \frac{3\tilde{\delta}_i^k - 1}{2[\nu - (1-\nu)(3\tilde{\delta}_i^k - 1)]}\left(2 - 3\nu \pm \sqrt{\frac{\nu(2-3\nu) + \nu(1-\nu)(4-5\nu)(3\tilde{\delta}_i^k - 1)}{\nu - (1-\nu)(3\tilde{\delta}_i^k - 1)}}\right).$$

- For $\frac{1}{3} < -\tilde{\delta}_i^k \leq \frac{2-3\nu}{3(1-\nu)}$ or for $\frac{1}{3} < -\tilde{\delta}_i^k \leq \frac{1}{1+3\nu}$ if $\nu > 1/3$, the smallest value for $|\lambda_i|$ is obtained for $\overline{\alpha}_i = \overline{\beta}_i = 0$:

$$a_i = \frac{2 + 6(1-\nu)\tilde{\delta}_i^k + \sqrt{3[1 + (1+3\nu)\tilde{\delta}_i^k][3 - 4\nu + 3(1-\nu)\tilde{\delta}_i^k]}}{1 - 3\tilde{\delta}_i^k(1-\nu)}$$
$$\lambda_i = -\frac{(1-a_i)^2}{(1+a_i)[2(1+a_i) - \nu(5+a_i)]}.$$

In the other cases, the previous conditions for the third local order accuracy are not compatible with the positivity. We choose the following functions:

- For $\tilde{\delta}_i^k > \frac{1}{3(1-\nu)}$, we take $\alpha_i = \beta_i = 0$ and

$$\overline{\beta}_i = 0 \qquad \overline{\alpha}_i = \frac{4\tilde{n}_i^k}{3(1-\tilde{\delta}_i^k)} \qquad a_i = 3\tilde{\delta}_i^k - 2$$

or

$$\overline{\beta}_i = \overline{\alpha}_i = \frac{\tilde{n}_i^k}{1-\tilde{\delta}_i^k} \qquad a_i = 2\tilde{\delta}_i^k - 1$$

choosing the case which gives the value of $\overline{\alpha}_i$ which is the closest to the average $(\tilde{n}_{i+1}^k + \tilde{n}_i^k)/2$.

- For $\tilde{\delta}_i^k < -\frac{2-3\nu}{3(1-\nu)}$ or, if $\nu > 1/3$, for $\tilde{\delta}_i^k < -\frac{1}{1+3\nu}$, we take $\overline{\alpha}_i = \overline{\beta}_i = 0$ and

$$\beta_i = 0 \qquad \alpha_i = \frac{4\tilde{n}_i^k}{3(1+\tilde{\delta}_i^k)} \qquad a_i = 2 + 3\tilde{\delta}_i^k$$

or

$$\beta_i = \alpha_i = \frac{\tilde{n}_i^k}{1+\tilde{\delta}_i^k} \qquad a_i = 1 + 2\tilde{\delta}_i^k$$

choosing the case which gives the value of α_i which is the closest to the average $(\tilde{n}_{i-1}^k + \tilde{n}_i^k)/2$.

With Proposition 3.9, we show that for a regular distribution and a sufficiently refined discretization, the first case ($|\tilde{\delta}_i^k| \leq 1/3$) is the most common, except for the sections corresponding to the small values of the distribution. The discontinuous case requires a special treatment described below.

3.3.5. *The discontinuous case*

The numerical analysis is only valid for regular functions and it is well-known that, if no precaution is taken with high order method, some oscillations (called the Gibbs phenomena in spectral methods) happen around discontinuities of the distribution function. To avoid that, the MUSCL scheme, for example, uses a slope limiter, which imposes that the scheme is TVD, losing its second order accuracy around discontinuities and extrema. An similar strategy for the scheme developed here could be to replace the affine or bi-affine approximated distribution function by a constant function in some sections. But the information on the droplet surface moment would then be lost. Moreover, it seems not necessary that the scheme is TVD for our method, since we only want to have a good estimation of the mass density, that is to say of the moment of order 3/2 of the approximated distribution function. However, it is important to limit the total variation to avoid too high oscillations and thus a loss of accuracy.

Since we want to conserve the surface first moments, we choose an alternative shape for the reconstruction function f_a^k which is a particular case of the bi-affine function defined by (3.33). It is easy to see that, at least for $|a_i| \leq 1/2$, the bi-affine function which has the minimal variation over the section is such that $\lambda_i = \lambda_i = -1$.

It is the following bi-constant function:

$$f_a^k\left(S_{i-1/2}+x\frac{\Delta S}{2}\right)=\frac{\tilde{n}_i^k}{\Delta S}\begin{cases}1-\dfrac{2\delta_i^k}{1+a_i} & \text{if } x\in[-1,a_i[,\\[2ex] 1+\dfrac{2\delta_i^k}{1-a_i} & \text{if } x\in[a_i,1[.\end{cases} \tag{3.37}$$

The value of a_i is chosen, among the values which preserve the positivity of f_a^k, to minimize the following quantity which appears in the total variation of f_a^k and which represents the sum of the absolute values of the discontinuities at $S=S_{i-1}$, $S=S_{i-1}+a_i\Delta S/2$ and $S=S_i$:

$$F_i(a_i)=\left|\tilde{n}_i^k-\frac{2\tilde{n}_i^k\delta_i^k}{1+a_i}-\overline{\alpha}_{i-1}^k\right|+\left|\tilde{n}_i^k+\frac{2\tilde{n}_i^k\delta_i^k}{1-a_i}-\alpha_{i+1}^k\right|+\left|4\frac{\tilde{n}_i^k\delta_i^k}{1-(a_i)^2}\right|\cdot \tag{3.38}$$

It still remains to determine when the reconstruction (3.37) is chosen to replace the one described in Section 3.3.4, for which the total variation in the section i is $\left|\alpha_i^k-\overline{\alpha}_i^k\right|$. A condition is then introduced, based on a comparison of the total variation in the section and its neighbors between the two types of reconstruction. So that this condition is independent of the reconstruction in the neighboring sections, the values $\tilde{n}_{i\pm1}/\Delta S$ are used for these sections and the condition is written:

$$\left|\tilde{n}_i^k-\frac{2\tilde{n}_i^k\delta_i^k}{1+a_i}-n_{i-1}^k\right|+\left|\tilde{n}_i^k+\frac{2\tilde{n}_i^k\delta_i^k}{1-a_i}-n_{i+1}^k\right|+\left|4\frac{\tilde{n}_i^k\delta_i^k}{1-(a_i)^2}\right|\leq\left|\tilde{n}_{i-1}^k-\alpha_i^k\right|+\left|\alpha_i^k-\overline{\alpha}_i^k\right|+\left|\overline{\alpha}_i^k-\tilde{n}_{i+1}^k\right|.$$

The numerical tests detailed in Section 4 show that this condition is a good detector for the discontinuities in the distribution and is fulfilled only for isolated sections so that the quantity (3.38) is really the one appearing in the total variation of f_a^k.

3.4. **Conclusion**

A new class of multi-fluid methods is introduced in this article, theoretical results of convergence were demonstrated, when the number of sections tends to infinity. A second order integration method was proposed to preserve positivity and deal with discontinuous distributions.

This last method could seem complex but, in practical cases, since only a few sections are concerned by the bi-affine reconstruction, the increase of the cost is low compared to the 2 moments method. Moreover it conserves some great advantages of the classical multi-fluid methods: it preserves the positivity of the variables and limits the total variations of the reconstructed function.

The next step is to test these new methods on a number of representative situations.

4. Numerical comparisons

The purpose of this section is to evaluate the methods proposed in this article and compare them with a more standard MUSCL second order scheme [26]. This last method is not applied directly to the mass density because the conservation equation is singular as explained in Section 2.1.1. It is used to determine the number density of the droplets, that is to say through a discretization of equation (1.2). The mass density is then evaluated through an integration of the reconstructed function with a minmod slope limiter.

Problem (1.2) serves as a test for three types of initial distribution functions. The first one is a regular, but non monotonic (Fig. 3 left), typical of functions that can be measured experimentally in sprays [17]. The second is a discontinuous function (Fig. 3 middle), which is used to examine the behavior of the method on a Riemann type problem. Finally, the third initial profile is bimodal (Fig. 3 right) typical of aluminum particles in solid propellant rockets [12]. This last case is intermediate between the two other: the function is regular, but with a high Lipschitz constant.

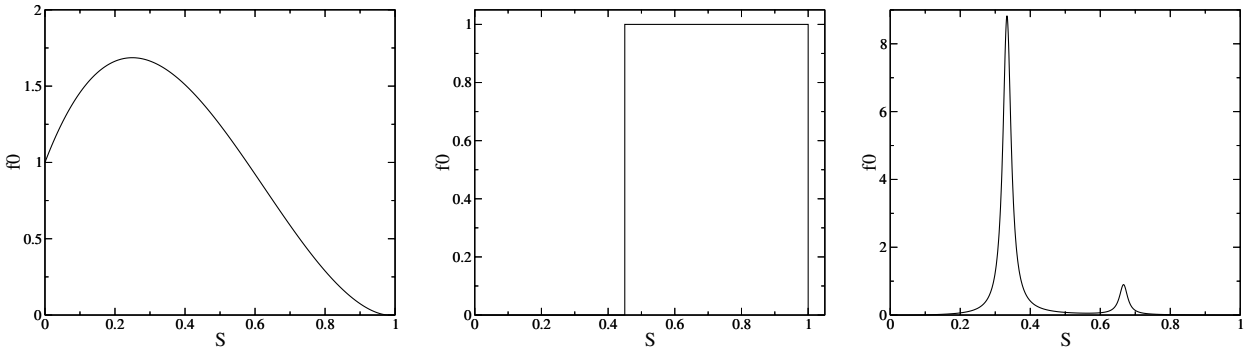

FIGURE 3. Initial distribution functions f_0 used for the computations.

TABLE 2. Maximal error (in percent) on the total mass density for the regular case done for each method with a refined discretization (100 sections) or a coarse discretization (3 sections).

	CMF	MUSCL	2M	M2M	3M
refined discretization	0.33	0.013	2×10^{-5}	1.7×10^{-5}	4.3×10^{-8}
coarse discretization	5.5	14	1	0.18	0.036

In each case, calculations are carried out with five different methods:

- the **2M** (2 moments) method introduced previously and using the moments $\tilde{n}_i^n$ and $\widetilde{nS_i}^n$;
- the **M2M** (modified 2 moments) method, with a bi-affine reconstruction of the distribution in each section;
- the **3M** (3 moments) method using the moments $\tilde{n}_i^n$, $\widetilde{nS_i}^n$ and $\widetilde{nS_i^2}^n$;
- the **MUSCL** method on the number density of the spray;
- the **CMF** (classical multi-fluid) method with a constant distribution function as a function of the surface in each section.

We want to observe the convergence of the methods and their capacity to produce a good estimate of the total mass density with coarse discretizations. It is also important to examine the Sauter mean diameter of the spray which is given, for a distribution f, by

$$\frac{\int S^{3/2} f(t,S)\,\mathrm{d}S}{\int S\, f(t,S)\,\mathrm{d}S}$$

and is often quoted in spray combustion applications. The Sauter mean diameter is used to characterize the global evolution of the droplet size.

4.1. **The regular case**

In the first test case, the initial distribution function f_0 is regular (see Fig. 3 left). Convergence of the methods is verified through calculations on a refined grid: 100 sections and 300 time steps. Figure 4 (left) gives the evolution of the error on the total mass density of the spray $\int S^{3/2} f(t,S)\,\mathrm{d}S$ and Figure 4 (right) gives the evolution of the Sauter mean radius. The values of the maximal error on the total mass density are given in Table 2. It shows that the different methods converge, with, as expected, the best accuracy on the mass density for the 3M method, then a very good accuracy for the M2M, the 2M and to a lesser extent the MUSCL method, the first order method (CMF) being the least accurate. Moreover, the evolution of the Sauter mean radius is well predicted by all the methods, with just a difference appearing for the MUSCL method where the

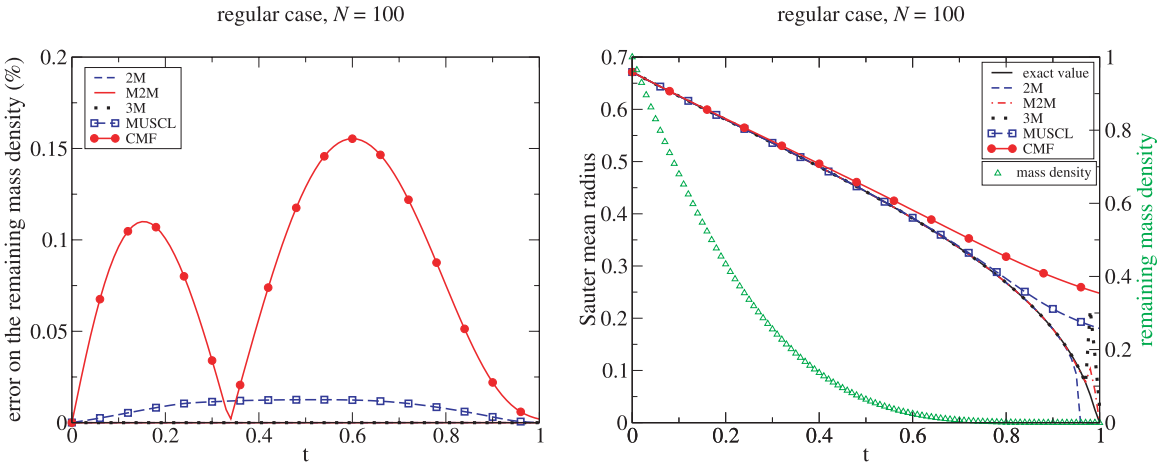

FIGURE 4. Evolution of the error on remaining mass density compared to the initial mass density (left) and of the Sauter mean radius (right) for calculations on a refined discretization (100 sections) with the 2 and 3 moments methods introduced in this article, with the MUSCL method and with the classical multi-fluid model.

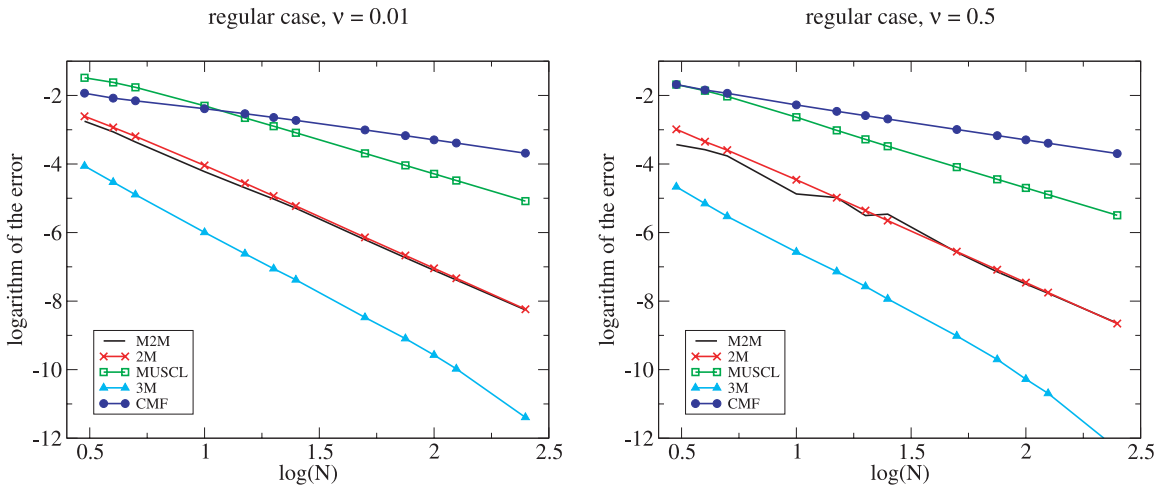

FIGURE 5. Decimal logarithm of the error on the mass density as a function of the decimal logarithm of the number of sections, for calculations with the 2 and 3 moments methods, with the MUSCL method and with the classical multi-fluid model, for $\nu = 0.01$ (left) or $\nu = 0.5$ (right).

remaining mass density is smaller than 0.7% of the initial mass and more early for the CMF methods, where the remaining mass density is smaller than 6%, showing their slow convergence.

This reduced rate of convergence of the CFM methods can also be seen in Figure 5 showing the evolution of this maximal error on the total mass density for an decreasing discretization step (with a constant value of $\nu = \Delta t/\Delta S$), on a log-log diagram. This figure also confirms the order of each method and allows to distinguish three groups of method depending their accuracy. First, the best is the 3M method. Then, the 2M and M2M have almost the same accuracy, the M2M being little better than the 2M. Finally, the MUSCL and the CFM methods are the worst methods, with a better accuracy for the MUSCL than for the CFM if the discretization

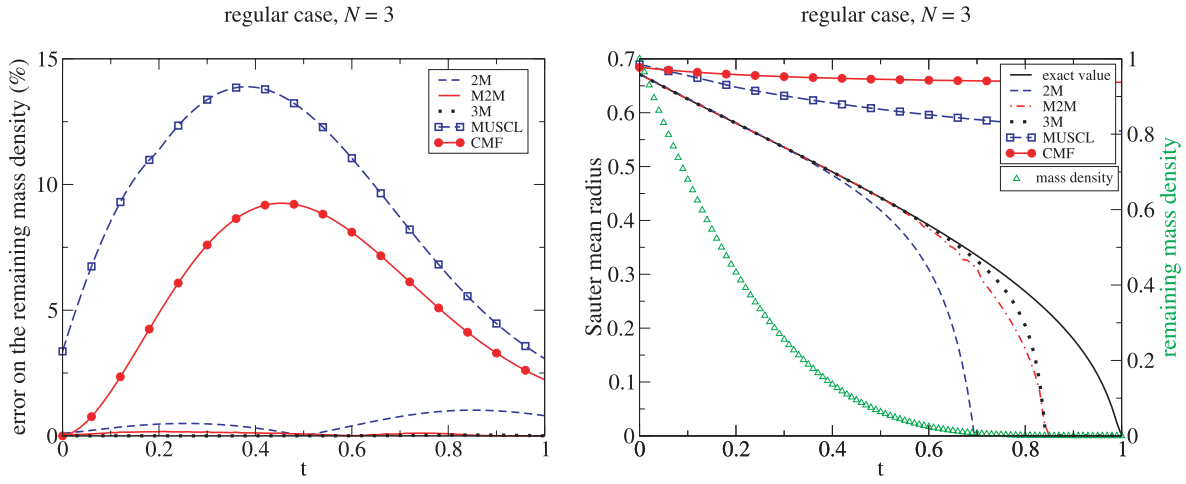

FIGURE 6. Evolution of the error on remaining mass density compared to the initial mass density (left) and of the Sauter mean radius (right) for calculations on a coarse discretization (3 sections) with the 2 and 3 moments methods introduced in this article, with the MUSCL method and with the classical multi-fluid model.

TABLE 3. Maximal error (in percent) on the total mass density for the singular case done for each method with a refined discretization (100 sections) or a coarse discretization (5 sections).

	CMF	MUSCL	2M	M2M	3M
refined discretization	0.33	0.13	6.1×10^{-4}	3.4×10^{-5}	4.4×10^{-5}
coarse discretization	3.5	7.8	0.27	0.17	0.039

is refined enough. The good accuracy of the 2M, M2M and 3M methods is apparent even for a really coarse discretization, with only 3 sections (it is the first point of the graphs in Fig. 5).

Results for this coarse discretization are plotted in Figure 6, with a small value of ν ($\nu = 0.03$) so that the only first order method for the time discretization of the classical multi-fluid method has little influence. For such a case, where the cost of the methods is very low, the 3M, M2M and 2M methods are still very good to predict the mass density (see Tab. 2) whereas the MUSCL method becomes much less accurate than the CMF which is already not very good. Moreover, the MUSCL and CMF methods give a poor evaluation of the Sauter mean radius for such a discretization whereas the other methods give good results till the point where the remaining mass density is less than 0.7%.

This last example shows the capability of the new methods to give good results with a very coarse discretization and then with a very low cost, in the case where the distribution is regular.

4.2. **The singular case**

For the second test case, the initial distribution function is a discontinuous step function (see Fig. 3 middle). As for the regular case, a first calculation is made with a refined mesh (100 sections and 300 time steps), to show the convergence of the methods. The conclusions concerning the prediction of the mass density and the Sauter mean radius of the spray are similar to the ones for the regular case with the refined discretization. The difference between the two cases appears in the error on the mass density (see Tab. 3): the MUSCL and also the 2M and 3M methods are less good than for the regular case whereas the CMF and the M2M give about the same error. The origins of such errors are different for MUSCL compared to 2M or 3M method, as shown by the reconstructed distribution function. It is given in Figure 7 for the time $t = 0.2$ s where almost 63% of

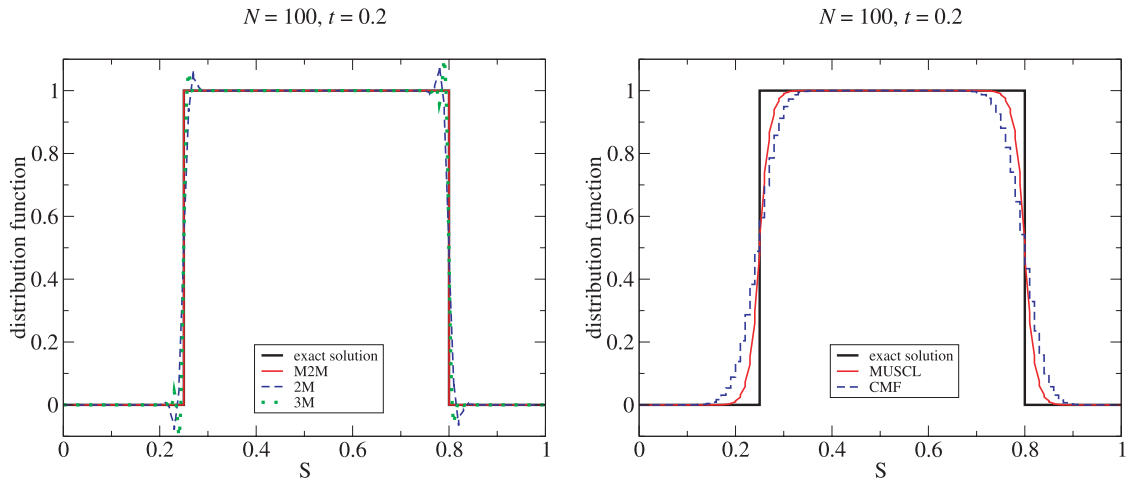

FIGURE 7. Distribution function at $t = 0.2$ for calculations on a refined discretization (100 sections) with the 2 and 3 moments methods (left) or with the MUSCL method and with the classical multi-fluid model (right).

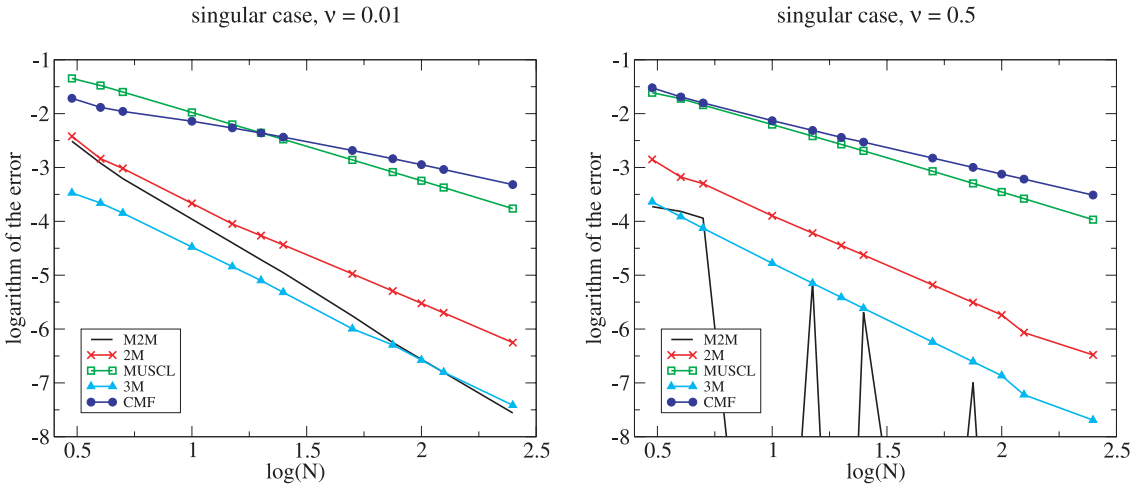

FIGURE 8. Decimal logarithm of the error on the mass density as a function of the decimal logarithm of the number of sections, for calculations with the 2 and 3 moments methods, with the MUSCL method and with the classical multi-fluid model, for $\nu = 0.01$ (left) or $\nu = 0.5$ (right).

the mass density remains. This figure shows that the error for the CMF method, and to a lesser extent for the MUSCL method is due to numerical diffusion, whereas for the 2M and 3M methods, it is essentially due to small oscillations near to the discontinuity but also to numerical diffusion. One may note that for the M2M method, no difference can be observed between the exact and the calculated distributions.

As for the regular case, the convergence of the methods is displayed on a log-log diagram (Fig. 8). This diagram shows the reduction of the order for all the methods in this discontinuous case, except for the M2M. Concerning this last method, for the small value of ν ($\nu = 0.01$), the error is more than second order accurate, with a value close to the one of the 2M method for coarse discretizations and smaller than the one of the

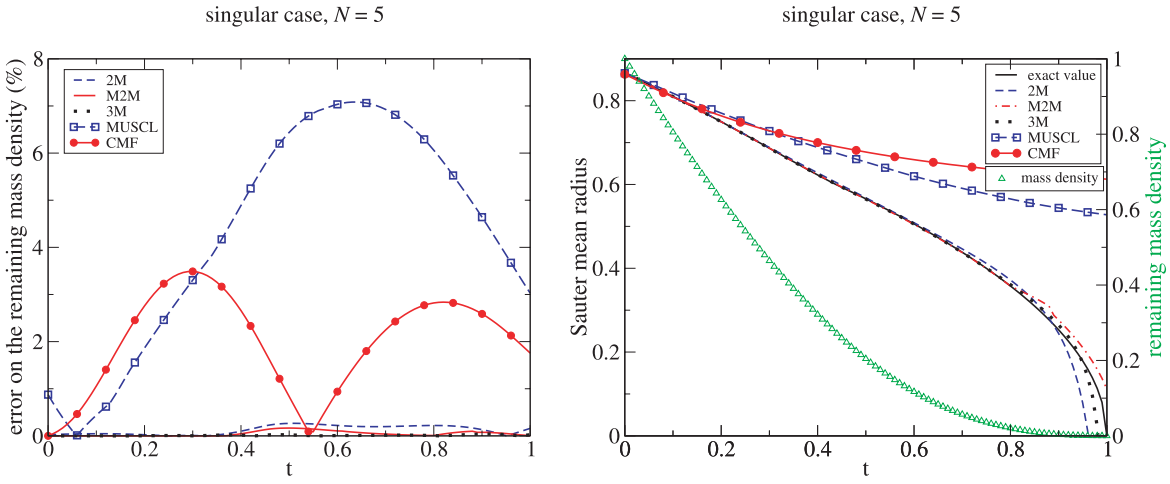

FIGURE 9. Evolution of the error on remaining mass density compared to the initial mass density (left) and of the Sauter mean radius (right) for calculations on a coarse discretization (5 sections) with the 2 and 3 moments methods introduced in this article, with the MUSCL method and with the classical multi-fluid model.

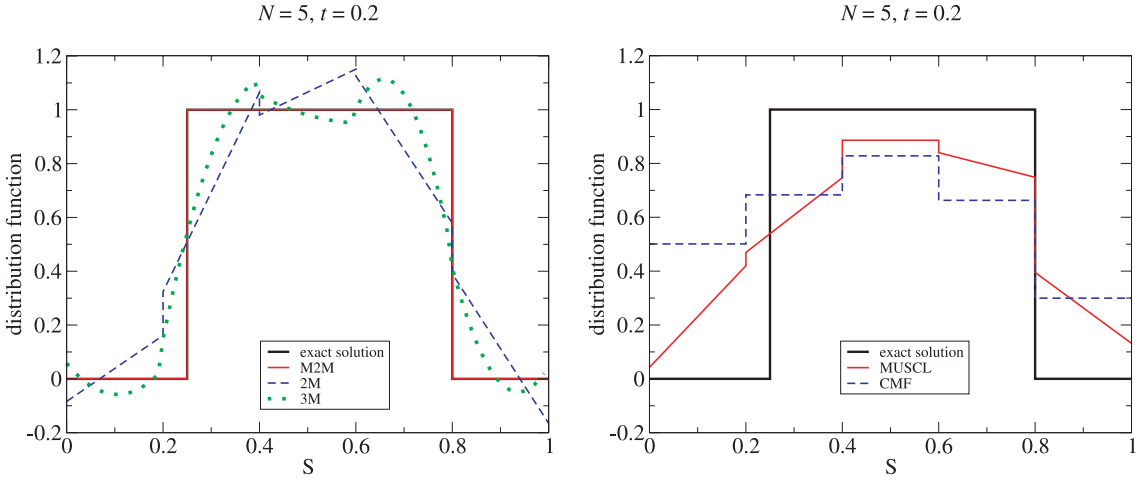

FIGURE 10. Distribution function at $t = 0.2$ for calculations on a coarse discretization (5 sections) with the 2 and 3 moments methods (left) or with the MUSCL method and with the classical multi-fluid model (right).

3M method for refined discretizations. For a bigger value of ν, the M2M method gives the exact solution for some discretizations and is as good as the 3M method for the others.

Let us consider a coarse discretization with only 5 sections. It corresponds to the third point in Figure 8 but with $\nu = 0.05$ (*i.e.* 100 time steps). Results given in Figure 9 and in Table 3 indicate that the CMF and MUSCL methods do not predict very well the evolution of the mass density and provide a poor description of the evolution of the mean Sauter radius. The other methods give good results for the mass density. Moreover, the description of the evolution of the mean Sauter radius is also good, especially before $t = 0.7$, where only 6% of the mass density remains. Finally, Figure 10 shows the reconstructed distribution function at time $t = 0.2$. One observes the difference of diffusion between the new methods and the MUSCL and CMF ones.

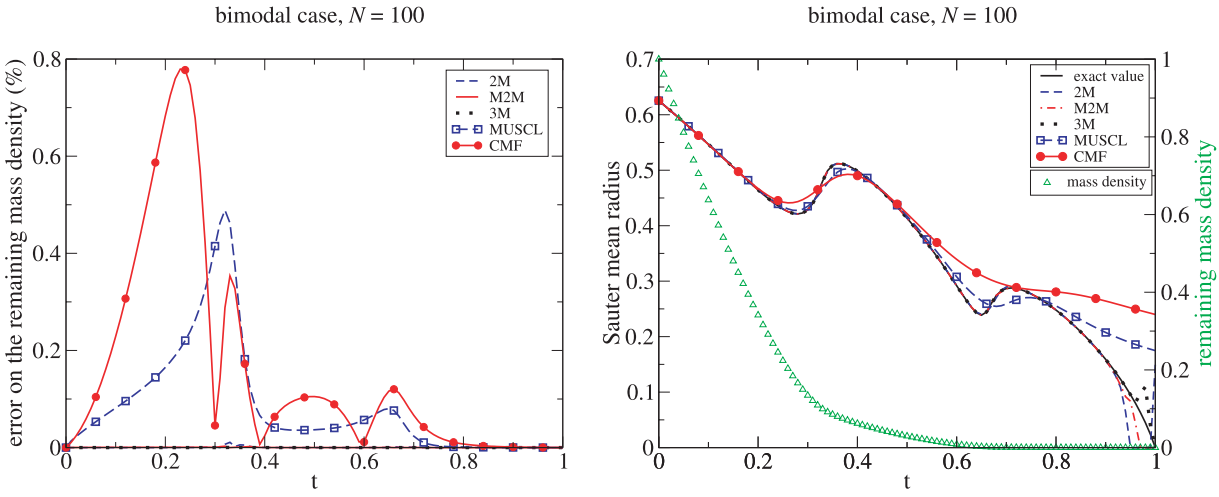

FIGURE 11. Evolution of the error on remaining mass density compared to the initial mass density (left) and of the Sauter mean radius (right) for calculations on a refined discretization (100 sections) with the 2 and 3 moments methods introduced in this article, with the MUSCL method and with the classical multi-fluid model.

TABLE 4. Maximal error (in percent) on the total mass density for the bimodal case done for each method with a refined discretization (100 sections) or a coarse discretization (10 sections).

	CMF	MUSCL	2M	M2M	3M
refined discretization	0.78	0.49	0.011	2.8×10^{-3}	2.9×10^{-4}
coarse discretization	3.2	11	0.89	0.61	0.22

This singular distribution constitutes an extreme case, interesting to test the methods but which is of less interest for application than the following intermediate distribution.

4.3. **The bimodal case**

For the third test case, the initial distribution function is regular but with a high Lipschitz constant, with essentially two groups of droplets of size around 0.33 and 0.66 (see Fig. 3 right). In this "bimodal" case, the Sauter mean radius evolves in a special way (see Fig. 11 right). For a refined case (100 sections and 300 time steps), the accuracy on the mass density is fairly good for the MUSCL and the CMF methods and is quite good for the other methods (see Tab. 4 and Fig. 11 left). But the CMF method and to a lesser extent the MUSCL method, have difficulties in describing the evolution of the Sauter mean radius, around $t = 0.3$ where 13% of the mass density remains and after $t = 0.5$ where less than 3% of the mass density remains.

These difficulties are due to the numerical diffusion of these schemes. This diffusion can be seen in Figure 12: at $t = 0.2$, where 34% of the mass remains, the approximation of the distribution is almost perfect for the M2M, 2M and 3M methods whereas the diffusion appear for the other methods, especially for the CMF.

As for the other cases, one observes the convergence of the methods on a log-log diagram (Fig. 13), with their theoretical order. These diagrams are comparable to the ones for the regular case except that here, the different between the M2M and 2M method is larger.

The behavior of the methods on a coarse discretization is specifically interesting. Results with 10 sections and 200 time steps are given in Figure 14 and in Table 4. As for the other case, the CMF and even more the MUSCL method are not very good and they are unable to follow the evolution of the Sauter mean radius.

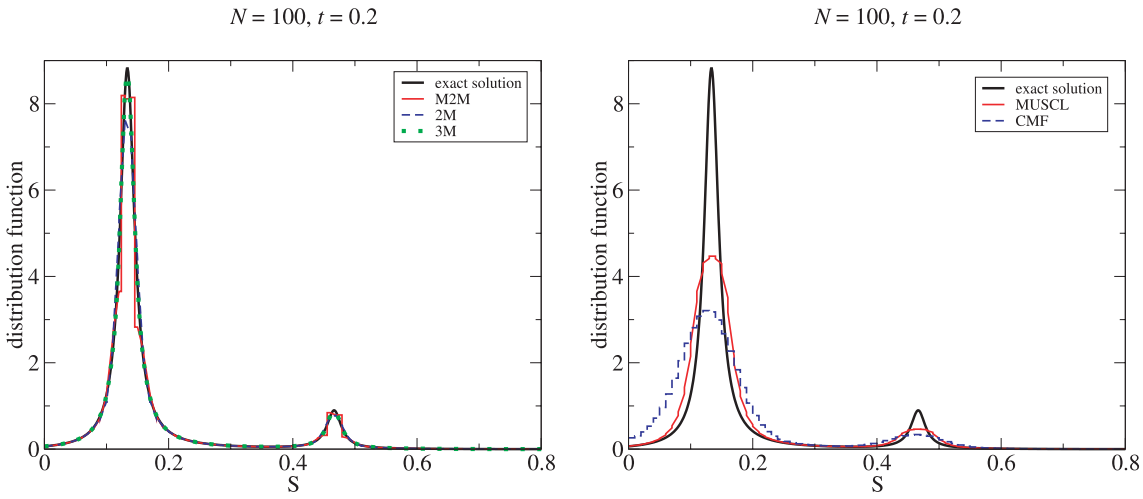

FIGURE 12. Distribution function at $t = 0.2$ for calculations on a refined discretization (100 sections) with the 2 and 3 moments methods (left) or with the MUSCL method and with the classical multi-fluid model (right).

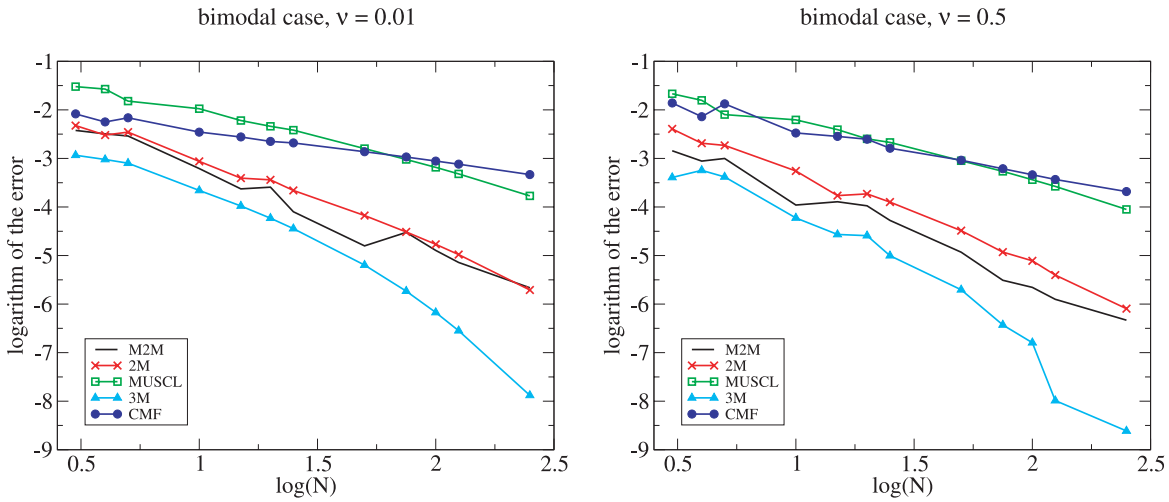

FIGURE 13. Decimal logarithm of the error on the mass density as a function of the decimal logarithm of the number of sections, for calculations with the 2 and 3 moments methods, with the MUSCL method and with the classical multi-fluid model, for $\nu = 0.01$ (left) or $\nu = 0.5$ (right).

The other methods have a good accuracy for the mass density (less than 1%) and they are able to capture the good qualitative behavior of the Sauter mean radius, least until $t = 0.5$, where only 3% of the mass remains.

5. Conclusion – discussion

It is shown in this article that the classical multi-fluid methods are first order accurate monotone and thus TVD with some CFL type condition. Their numerical analysis also concludes about the "best way" to discretize the equation. But this first order accuracy means also slow convergence and large numerical diffusion. Higher

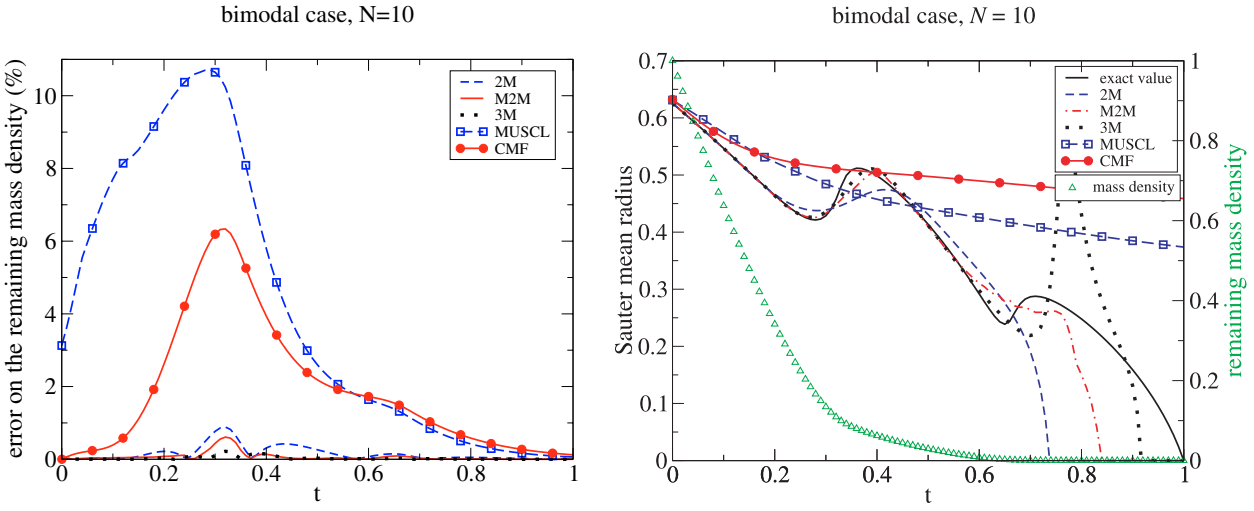

FIGURE 14. Evolution of the error on remaining mass density compared to the initial mass density (left) and of the Sauter mean radius (right) for calculations on a coarse discretization (10 sections) with the 2 and 3 moments methods introduced in this article, with the MUSCL method and with the classical multi-fluid model.

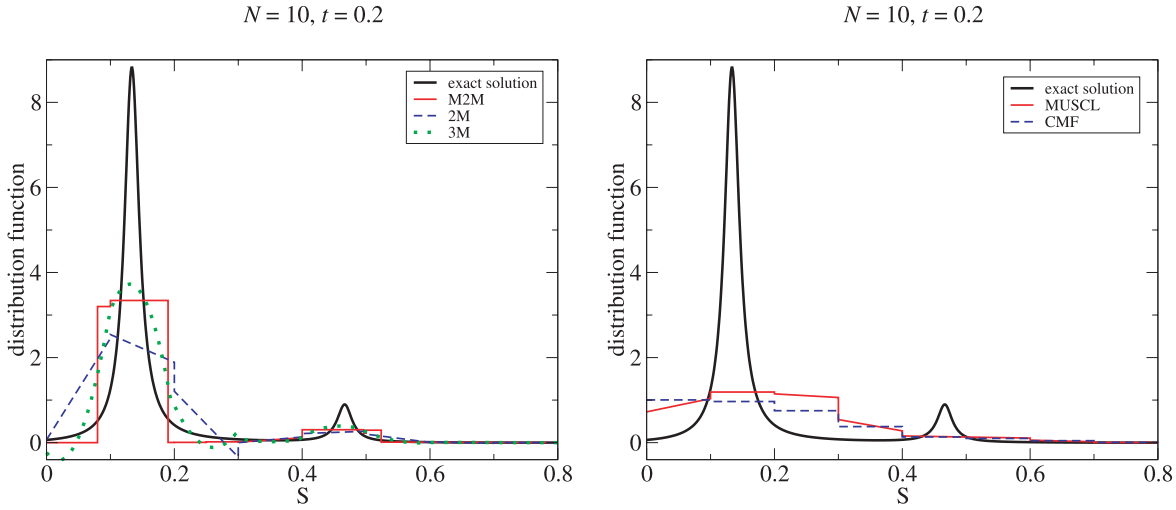

FIGURE 15. Distribution function at $t = 0.2$ for calculations on a coarse discretization (10 sections) with the 2 and 3 moments methods (left) or with the MUSCL method and with the classical multi-fluid model (right).

order methods providing number density, like the MUSCL method can be used. It is TVD, preserves the positivity of the mass density, but its accuracy is not good if the discretization is not sufficiently refined.

As an alternative, some moment methods have been introduced here: we consider some moments for the droplets surface variable in each section, the mass density being calculated through an approximation of the distribution. The consistency order of such a method is the same as the number of these moments. We also prove the stability and convergence of the 2 and 3 moments methods. These methods do not necessarily ensure the positivity of the mass density or of the moments. Moreover, they induce some artificial oscillations of

the approximated distribution function around discontinuities of the distribution. However, these oscillations have only a small impact on the variable of interest: the mass density of the spray and they yield a good approximation of the solution, even on a coarse mesh. An improvement of the 2 moments method is introduced with the modified 2 moments method. It preserves the positivity of the mass density and also reduces the oscillations of the solution around the discontinuity thanks to a slope limiter like condition.

These new methods, with their efficiency even for a small number of variables and a reduced CPU time, are good candidates to simulate evaporating sprays in more complex configurations.

Acknowledgements. The support of the French Ministry of Research through a grant in the framework of "Recherche aéronautique sur le supersonique" (Réseau de Recherche et d'Innovation Technologique) and through a grant in the framework of "Nouvelles Interfaces des Mathématiques" (Action Concertée Incitative, "Milieux Réactifs: vers de nouveaux outils de simulation pour les problèmes multi-échelles", 2003–2006) is gratefully acknowledged. The author wishes to thank Marc Massot and Jean-François Maître for helpful discussions and comments.

References

[1] A.A. Amsden, P.J. O'Rourke and T.D. Butler, *Kiva II, a computer program for chemically reactive flows with sprays.* Technical Report LA-11560-MS. Los Alamos National Laboratory, Los Alamos, New Mexico (1989).
[2] G. Chanteperdrix, P. Villedieu and J.P. Vila, A compressible model for separated two-phase flows computations, in *ASME Fluids Engineering Division Summer Meeting*, number 31141, Montreal (2002).
[3] K. Domelevo, The kinetic sectional approach for noncolliding evaporating sprays. *Atomization Spray.* **11** (2001) 291–303.
[4] K. Domelevo and L. Sainsaulieu, A numerical method for the computation of the dispersion of a cloud of particles by a turbulent gas flow field. *J. Comput. Phys.* **133** (1997) 256–278.
[5] D.A. Drew and S.L. Passman, *Theory of multicomponent fluids.* Applied Mathematical Sciences, Springer **135** (1999).
[6] G. Dufour and P. Villedieu, A second-order multi-fluid model for evaporating sprays. *ESAIM: M2AN* **39** (2005) 931–963.
[7] J.K. Dukowicz, A particle-fluid numerical model for liquid sprays. *J. Comput. Phys.* **35** (1980) 229–253.
[8] J.B. Greenberg, D. Albagli and Y. Tambour, An opposed jet quasi-monodisperse spray diffusion flame. *Combust. Sci. Technol.* **50** (1986) 255–270.
[9] J.B. Greenberg, I. Silverman and Y. Tambour, On the origin of spray sectional conservation equations. *Combust. Flame* **93** (1993) 90–96.
[10] H. Guillard and A. Murrone, *A five equation reduced model for compressible two phase flow problems.* Prepublication 4778, INRIA (2003).
[11] A. Harten, J.M. Hyman and P.D. Lax, On finite-difference approximations and entropy conditions for shocks. *Comm. Pure Appl. Math.* **29** (1976) 297–322. With an appendix by B. Keyfitz.
[12] J. Hylkema, *Modélisation cinétique et simulation numérique d'un brouillard dense de gouttelettes. Application aux propulseurs à poudre.* Ph.D. thesis, ENSAE (1999).
[13] F. Laurent, Analyse numérique d'une méthode multi-fluide Eulérienne pour la description de sprays qui s'évaporent. *C. R. Math. Acad. Sci. Paris* **334** (2002) 417–422.
[14] F. Laurent, *Modélisation mathématique et numérique de la combustion de brouillards de gouttes polydispersés.* Ph.D. thesis, Université Claude Bernard, Lyon 1 (2002).
[15] F. Laurent and M. Massot, Multi-fluid modeling of laminar poly-dispersed spray flames: origin, assumptions and comparison of the sectional and sampling methods. *Combust. Theor. Model.* **5** (2001) 537–572.
[16] F. Laurent, M. Massot and P. Villedieu, Eulerian multi-fluid modeling for the numerical simulation of polydisperse dense liquid spray. *J. Comput. Phys.* **194** (2004) 505–543.
[17] F. Laurent, V. Santoro, M. Noskov, A. Gomez, M.D. Smooke and M. Massot, Accurate treatment of size distribution effects in polydispersed spray diffusion flames: multi-fluid modeling, computations and experiments. *Combust. Theor. Model.* **8** (2004) 385–412.
[18] R.J. LeVeque, *Numerical methods for conservation laws.* Birkhäuser Verlag, Basel, second edition (1992).
[19] D.L. Marchisio, R.D. Vigil and R.O. Fox, Quadrature method of moments for aggregation-breakage processes. *J. Colloid Interf. Sci.* **258** (2003) 322–334.
[20] M. Massot and P. Villedieu, Modélisation multi-fluide eulérienne pour la simulation de brouillards denses polydispersés. *C. R. Acad. Sci. Paris Sér. I Math.* **332** (2001) 869–874.
[21] M. Massot, M. Kumar, A. Gomez and M.D. Smooke, Counterflow spray diffusion flames of heptane: computations and experiments, in *Proceedings of the 27th Symp. (International) on Combustion*, The Comb. Institute (1998) 1975–1983.
[22] P.J. O'Rourke, *Collective drop effects on vaporizing liquid sprays.* Ph.D. thesis, University of Princeton (1981).

[23] D. Ramkrishna and A.G. Fredrickson, *Population Balances: Theory and Applications to Particulate Systems in Engineering.* Academic Press (2000).
[24] P.-A. Raviart and L. Sainsaulieu, A nonconservative hyperbolic system modeling spray dynamics. I. Solution of the Riemann problem. *Math. Mod. Meth. Appl. S.* **5** (1995) 297–333.
[25] M. Rüger, S. Hohmann, M. Sommerfeld and G. Kohnen, Euler/Lagrange calculations of turbulent sprays : the effect of droplet collisions and coalescence. *Atomization Spray.* **10** (2000) 47–81.
[26] B. van Leer, Towards the ultimate conservative difference scheme v. a second order sequel to godunov's method. *J. Comput. Phys.* **32** (1979) 101–136.
[27] P. Villedieu and J. Hylkema, Une méthode particulaire aléatoire reposant sur une équation cinétique pour la simulation numérique des sprays denses de gouttelettes liquides. *C. R. Acad. Sci. Paris Sér. I Math.* **325** (1997) 323–328.
[28] F.A. Williams, Spray combustion and atomization. *Phys. Fluids* **1** (1958) 541–545.
[29] F.A. Williams, *Combustion Theory (Combustion Science and Engineering Series).* F.A. Williams Ed., Reading, MA: Addison-Wesley (1985).
[30] D.L. Wright, R. McGraw and D.E. Rosner, Bivariate extension of the quadrature method of moments for modeling simultaneous coagulation and sintering of particle populations. *J. Colloid Interf. Sci.* **236** (2001) 242–251.