Realizable second-order finite-volume schemes for the advection of moment sets of the particle size distribution

Frédérique Laurent, Tan Trung Nguyen

Realizable second-order finite-volume schemes for the advection of moment sets of the particle size distribution

F. Laurent[a,b,*], T. T. Nguyen[a,b]

[a] *Laboratoire EM2C, CNRS, CentraleSupélec, Université Paris-Saclay, Grande Voie des Vignes, 92295 Châtenay-Malabry cedex, France*
[b] *Fédération de Mathématiques de l'Ecole Centrale Paris - FR CNRS 3487*

Abstract

The accurate description and robust simulation at relatively low cost of a size polydisperse population of fine particles in a carrier fluid is still a major challenge for many applications. For this purpose, moment methods, derived from a population balance equation, represent a very interesting strategy. However, one of the major issues of such methods is the realizability: the numerical schemes have to ensure that the moment sets stay realizable, *i.e.* that an underlying distribution exists. This issue is all the more crucial that some moment vectors can be at the boundary of the moment space for practical applications, corresponding to a population of particles with only one or a few sizes. It is then investigated here for the advection operator, for which it is particularly significant. Then second order realizable kinetic finite volume schemes are designed, with two strategies for the fluxes evaluation based on the work of Kah et al. [1] and of Vikas et al. [2], which are here completely revisited, extended to take into account the boundary of the moment space and any number of moments, analyzed and compared in a Cartesian mesh context. For a potential easiest generalization to unstructured meshes, simplified but still realizable versions of these schemes are also developed. The high accuracy of all the schemes is then numerically checked on 1D and 2D test cases, with Cartesian meshes, and their robustness is shown, even when some moment vectors are at the boundary of the moment space.

Keywords: population balance equation, moment method, advection, realizable scheme, finite volume, kinetic scheme

1. Introduction

Populations of non-inertial particles in a carrier fluid are encountered in several kinds of applications (see [3] and references therein): soot in combustion applications, nanoparticles synthesis, microbubbles in biology processes, aerosol technology, ... Its evolution can be described by a population balance equation (PBE) [3, 4, 5, 6], which is a transport equation for the number density function (NDF) of the particles. This NDF depends on time, spatial location and one or several internal coordinates, which can for example describe the size of the particles. The PBE includes

*Corresponding author
Email addresses: frederique.laurent@centralesupelec.fr (F. Laurent), tan-trung.nguyen@centralesupelec.fr (T. T. Nguyen)

usually the spatial transport terms, describing for example advection and diffusion, and some localized source terms describing, at each spatial location, phenomena such as nucleation, aggregation, coagulation, breakup, growth or oxidation/dissolution. It is coupled with the equations, usually Navier-Stokes equations, describing the carrier fluid [7].

In this work, only one internal variable is considered, describing the size of the particles, assuming for example that they are spherical. In order to be able to describe the size polydispersion of the particles at a reasonable cost, the use of moment methods seems to be an interesting strategy (see for example [8, 9, 4]): only a finite set of moments of the NDF are then transported. It can also be hybridized with a discretization along the internal coordinate [10, 11, 12, 13, 14]. However, two major issues arise for moment methods. The first one is the closure of the moment equations essentially due to the source terms in the PBE. Several strategies were used: some of them provide a functional dependence of the unknown moments using the transported moment set, such as the interpolative closure (MOMIC) [15]. For the other ones, a NDF, or its corresponding measure, is reconstructed from the moment set, allowing evaluation of all the unclosed terms. This reconstruction can be for example the entropy maximization [16, 17, 18], a sum of Dirac delta function (quadrature method of moment, QMOM) [8] or a superposition of kernel density functions (kernel density element method, KDEM [19] or extended quadrature method of moment EQMOM [20, 21, 22]). The second major issue of moment methods is the realizability. Indeed, since the set of variables are the moments of a non-negative NDF (or, more rigorously, a positive measure) on $\mathbb{R}_+$ or a sub-interval of $\mathbb{R}_+$, it belongs to a space strictly included in $\mathbb{R}_+^N$, where N is the number of moments [23, 24, 25]. This space is called the moment space. The numerical methods have to ensure that the variables stay in this moment space, *i.e.* that the moments stay realizable. This issue is not always considered, thus leading to unphysical results (*e.g.* invalid moment sets). Indeed, the classical schemes for high-order transport in physical space can lead to invalid moment sets [26, 2, 27], as well as for the source terms [13, 12], even if the closure itself ensures the realizability at the continuous level. This happens all the more easily when some moment sets are at the boundary of the moment space, thus corresponding to a sum of a few weighted Dirac delta functions, as obtained through nucleation. To circumvent this issue, some authors resort to moment correction algorithms [28, 26] based on a necessary but eventually not sufficient condition for realizability in order to obtain a valid moment set. The cost of the method then increases and the correction spoils the overall accuracy.

It is then very important to develop realizable schemes, *i.e.* schemes directly preserving the realizability of the moment set. Moreover, an operator splitting strategy, solving separately the spatial transport of both phases and the source terms was shown to be efficient and well adapted to industrial-oriented codes [29]. This allows us to deal separately with the spatial transport and the source terms. Concerning the source terms, realizable schemes were already developed for moment methods where the closure is based on a reconstruction of the NDF [13, 4, 12]. A realizable scheme was also provided for the diffusion operator in the case of QMOM [27, 4]. In this work, only advection then is considered. In practice, this operator, at least when using first order explicit finite volume methods, is usually much less costly than the potentially complex source term operator, composed of one ODEs system for each considered moment vector. So, it will be very interesting to minimize the number of degree of freedom by using a high order scheme for advection, as soon as its cost is not prohibitive.

A Lagrangian type of scheme has been developed [30]. The advection of the moments is then obtained through the advection of some numerical particles for which a moment vector is affected. The resulting scheme is then naturally realizable. However, it suffers from the same drawbacks

as usual Lagrangian methods: the need of interpolation of the carrier phase properties, the non easy coupling with this phase and the complexity in term of parallelization for high performance computing. Moreover, it could need a large number of numerical particles to converge, for which the eventually costly source terms operator has to be solved. That is why only Eulerian schemes are studied here. On the one hand, a second-order realizable kinetic finite volume method has been developed in a structured mesh context [1], when the support of the NDF is compact. It was recently applied in a context of a mesh refinement [31]. It is based on a kinetic evaluation of the fluxes thanks to the use of the analytical solution at the kinetic level and on a MUSCL type of reconstruction on the canonical moments, which define a one to one relation between the interior of the moment space and the interior of an hypercube. However, it was only applied for inertial particles and for a four moments method, the algebra being otherwise difficult. On the other hand, a pseudo–second-order realizable finite volume method [2] has been developed in structured and unstructured mesh contexts. Fluxes computation is then based on a reconstruction of the moments at the cell interfaces; it is obtained thank to the Gauss quadrature of the moments, just reconstructing the weights. However, it was reduced to an even number of moments and suffers from some accuracy reduction when the quadrature points evolve strongly. In this work, the last two schemes are completely revisited, generalizing them to any number of moments and allowing them to deal with the boundary of the moment space without loosing the realizability, which is an hard task. They are also analyzed, especially looking for conditions to obtain the second order of accuracy, and they are compared. This is done for NDF of support included in $[0, +\infty)$, the case of a compact support being discussed in the appendix. Kinetic schemes are thus derived in a structured mesh context, first with a reconstruction on variables defining a one to one relation between the interior of the moment space and $(0, +\infty)^N$, where N is the number of moments in the set. Algorithms are then adapted to the case of the support included in $[0, +\infty)$ and generalized to any number of moments. The weight reconstruction is also considered, for an even or an odd number of moments, in a different way compared to [2], thus not being dependent of abscissas differences between the cells. Simplified schemes are then derived with the two kinds of reconstructions, the first one being modified for this case. It will be generalizable to unstructured meshes in a cell-centered context.

The paper is then organized as follows. In Section 2, the moment equations for the pure advection case are given, as well as the realizability constraints. Then, in Section 3, realizable finite volume kinetic schemes are provided and their orders of accuracy discussed. Some simplified version of these schemes are also given in Section 4, as well as the new constraint on the CFL to guarantee the realizability. Finally some verifications are given, considering systems with high numbers of moments, first for 1D configurations with steady or unsteady and compressible carrier phase velocity fields in Section 5 and for the 2D configuration of the Taylor-Green vortices in Section 6.

2. Moment transport equation and realizability

In this section, moment equations for the pure advection case are first recalled. Then, the space in which the moment vector lives is described, as well as the realizability conditions. This can be done directly, using the Hankel determinants. But some interesting tools are also introduced, defining a bijection between the interior of the moment space $\mathcal{M}_N$ and $(0, +\infty)^N$.

2.1. Moment equations

Let us consider the NDF, denoted $f(t, x, \xi)$, of some cloud of small particles transported by a carrier fluid. The parameter ξ, which is the size of the particles, lives in the interval $[0, +\infty)$. The

case of a compact support is discussed in Appendix A. In the case of pure advection, the population balance equation (PBE) then reduces to:

$$\partial_t f + \partial_x (u\, f) = 0. \tag{1}$$

Since the considered particles are non-inertial, u is the velocity of the carrier phase, which is a priori compressible. It is assumed to be a regular function of (t, x) in what follows.

Instead of resolving directly this PBE, one considers a finite set of moments $(m_k)_{k\in\{0,1,\dots,N\}}$ of the NDF, the k^{th} order moment m_k being defined by:

$$m_k(t, x) = \int_0^{+\infty} \xi^k f(t, x, \xi)\mathrm{d}\xi. \tag{2}$$

These moments are the solution of the following system of equations, denoting $\mathbf{m}_N = (m_0, \dots, m_N)^t$:

$$\partial_t \mathbf{m}_N + \partial_x (u\, \mathbf{m}_N) = 0. \tag{3}$$

Let us remark that this system is closed, contrary to the often encountered systems in moment methods. However, it is usually only a part of a more complex problem for which some closure is needed and can be provided through the reconstruction of a NDF from the moments (see for example [8, 1, 21]). Moreover, the moment equations seem here independent from one another. In fact, they are coupled by the fact that the vector $\mathbf{m}_N = (m_0, m_1, \dots, m_N)^t$ has to be a moment vector of a positive measure, *i.e.* has to stay in the moment space. This is the realizability condition, which is detailed in the next section.

2.2. Moment space: definition and first characterization

Let us denote $\mathcal{P}$ the space of finite positive Borel measures on $(0, +\infty)$. And for $\mu \in \mathcal{P}$, let us denote $\mathbf{m}_N(\mu)$ the vector of moments of μ of order 0 to N, assuming that they are finite:

$$\mathbf{m}_N(\mu) = (m_0(\mu), \dots, m_N(\mu))^t, \quad m_k(\mu) = \int_0^{+\infty} x^k \mathrm{d}\mu. \tag{4}$$

The moment vectors $\mathbf{m}_N = (m_0, m_1, \dots, m_N)^t$ lives in the Nth-moment space.

Definition 2.1. The Nth-moment space $\mathcal{M}_N$ on the interval $(0, \infty)$ is given by

$$\mathcal{M}_N = \{\mathbf{m}_N(\mu) \quad | \quad \mu \in \mathcal{P}\}.$$

If a moment vector $\mathbf{m}_N$ belongs to this space, it is said to be *realizable* and one then defines

$$\mathcal{P}(\mathbf{m}_N) = \left\{\mu \in \mathcal{P} \quad | \quad \mathbf{m}_N = \int_0^{+\infty} (1, x, \dots, x^N)^t \mathrm{d}\mu\right\}.$$

If $\mathbf{m}_N$ belongs to the interior of this space, it is said to be *strictly realizable*.

This Nth-moment space is convex. To characterize it, one can introduce the Hankel determinants, defined by:

$$\underline{H}_{2n+d} = \begin{vmatrix} m_d & \dots & m_{n+d} \\ \vdots & & \vdots \\ m_{n+d} & \dots & m_{2n+d} \end{vmatrix}, \tag{5}$$

with $d = 0, 1$; $n \geq 0$. Indeed, one has the following theorem, for which a proof can be found in [23, 24, 25]:

Theorem 2.1. *The vector* $\mathbf{m}_N = (m_0, m_1, \ldots, m_N)^t$ *is strictly realizable if and only if*

$$\underline{H}_k > 0, \quad k \in \{0, 1, \ldots, N\} \tag{6}$$

and if it belongs to the boundary of the moment space, then there exists $k \leq N$ *such that*

$$\underline{H}_0 > 0, \ldots, \underline{H}_{k-1} > 0, \underline{H}_k = 0, \ldots, \underline{H}_N = 0. \tag{7}$$

In the latter case, k *is denoted* $\mathcal{N}(\mathbf{m}_N)$ *and* $\mathcal{P}(\mathbf{m}_N)$ *is a singleton: a sum of* $\left\lfloor \frac{k+1}{2} \right\rfloor$ *weighted Dirac delta functions*[1]*.*

Let us remark that if $\mathcal{N}(\mathbf{m}_N)$ is odd, then one of the Dirac delta functions is centered at 0. Moreover, (7) is not a sufficient condition for a moment to be on the boundary (and not outside) of the moment space: an additional condition is then needed [32, 33].

The moment space has a rather complex geometry, as explained in the next section. Moreover, the Hankel determinants provide algebraic relations to determine if a vector belongs to the moment space but this tool is not easy to use, since we do not want to compute all these determinants. But other quantities can be derived, linked with a one-to-one mapping of the interior of the moment space, thanks to the theory of orthogonal polynomials.

2.3. Orthogonal polynomials theory and new characterization of the moment space

Let $\mathbb{P}$ be the space of real polynomials. For a positive finite Borel measure μ such that its moments are well defined and for p, q in $\mathbb{P}$, let us define the scalar product:

$$\langle p, q \rangle = \int_{\mathbb{R}} p(x)q(x)\mathrm{d}\mu. \tag{8}$$

When considering a measure μ with infinite support, a sequence $(\pi_k)_{k\geq 0}$ of orthogonal polynomials relative to this scalar product ($\langle \pi_k, \pi_p \rangle = \delta_{k,p}\langle \pi_k, \pi_k \rangle$), where π_k is of exact degree k, satisfies the following three term recurrence relation, with $\pi_0 = 1$, $\pi_{-1} = 0$ and $\beta_k > 0$ [34]:

$$\pi_{k+1}(x) = (x - \alpha_k)\pi_k(x) - \beta_k \pi_{k-1}(x). \tag{9}$$

Conversely, if the sequence of polynomials satisfies (9) with $\beta_k > 0$ for all $k \in \mathbb{N}$, then there exists a measure μ on the real line for which the polynomials are orthogonal. Moreover, this measure is supported on $[0, +\infty)$ if and only if there exists a sequence $(\zeta_k)_{k\geq 1}$ of positive numbers such that the coefficients in the recurrence relation (9) satisfy $\alpha_0 = \zeta_1$ and for all $k \geq 1$ [35]:

$$\beta_k = \zeta_{2k-1}\zeta_{2k}, \quad \alpha_k = \zeta_{2k} + \zeta_{2k+1}. \tag{10}$$

And this measure is supported on $[0, 1]$ if and only if the coefficients ζ_k form a chain sequence, *i.e.* they can be decomposed as $\zeta_k = p_k(1 - p_{k-1})$, with $p_0 = 0$ and $p_k \in (0, 1)$ for $k \geq 1$ [36].

When considering a measure with finite support, the sequence of such orthogonal polynomials is finite and only a finite number of the ζ_k (or of the p_k) can be defined. These coefficients then allow to characterize the interior of the moment space in the case where the support is included in $[0, +\infty)$.

[1] $\lfloor r \rfloor$ denotes here the largest integer less than or equal to the real number r

Indeed, for a moment vector $\mathbf{m}_N$ at the boundary of the moment space $\mathcal{M}_N$, corresponding to a measure with a finite support, one has $\zeta_{\mathcal{N}(\mathbf{m}_N)} = 0$ and the ζ_k for $k > \mathcal{N}(\mathbf{m}_N)$ are not defined.

A more geometrical point of view can also be considered [24]. For a vector $\mathbf{m}_{k-1}$ in the interior of the moment space $\mathcal{M}_{k-1}$, let us then first define

$$m_k^-(\mathbf{m}_{k-1}) = \min_{\mu \in \mathcal{P}(\mathbf{m}_{k-1})} m_k(\mu), \qquad m_k^+(\mathbf{m}_{k-1}) = \max_{\mu \in \mathcal{P}(\mathbf{m}_{k-1})} m_k(\mu), \tag{11}$$

as the lower and upper boundary of the admissible interval for the moment m_k of order k, the lower order moments $\mathbf{m}_{k-1}$ being known. In what follows, only measures supported on $[0, +\infty)$ are considered. In this case, $m_k^+ = +\infty$ and m_k^- is finite, with $m_k - m_k^- = m_0 \prod_{i=1}^k \zeta_i$. Moreover, m_k^- strongly depends on $\mathbf{m}_{k-1}$, meaning that the interval where the k^{th} order moment lives strongly depends on the value of the lower order moments. Some examples are given in Appendix B.

Then, the strict realizability is characterized either by the positivity of the Hankel determinants or by the positivity of the ζ_k and induces a link between the moments. Several algorithms allow to compute efficiently the recursion coefficient α_k and β_k and then the ζ_k from the moments: Rutishauser's QD algorithm [37, 38], Gordon's PD algorithm [39, 40] and variation of an algorithm attributed to Chebyshev and given by Wheeler in [41]. Since this last one is said to be slightly more stable in practice [41], let us detail it, as well as the reverse algorithm, giving the moments from the ζ_k and m_0.

2.4. Chebyshev algorithm for the computation of the ζ_k and reverse algorithm

Consider the matrix Z with elements $Z_{k,p} =< \pi_k x^p >$, which must be zero if $k > p$ and which satisfy $Z_{-1,p} = 0$, $Z_{0,p} = m_p$ and, thanks to the orthogonal polynomials recursion formula:

$$Z_{k+1,p} = Z_{k,p+1} - \alpha_k Z_{k,p} - \beta_k Z_{k-1,p}. \tag{12}$$

Coefficients α_k, β_k are determined by:

$$\beta_0 = m_0, \qquad \alpha_0 = \frac{m_1}{m_0}, \qquad \forall k > 0 \quad \beta_k = \frac{Z_{k,k}}{Z_{k-1,k-1}}, \quad \alpha_k = \frac{Z_{k,k+1}}{Z_{k,k}} - \frac{Z_{k-1,k}}{Z_{k-1,k-1}}, \tag{13}$$

which results from $Z_{k+1,k-1} = 0 = Z_{k+1,k}$. Thus, if the $Z_{k,p}$ are known for each $k \leq n$ and $p \leq k$, then one can compute the α_n, β_n from (13) and then the $Z_{n+1,p}$ for $p = 0, \ldots, n+1$ from (12).

In what follows, it will be also interesting to compute the moments $(m_k)_{k \in \{1,\ldots,N\}}$ from the $(\zeta_k)_{k \in \{1,\ldots,N\}}$ and from m_0. It could be done by reversing the previous algorithm, but Skibinsky [42] showed the following theorem, also proved in [24] by another method.

Theorem 2.2. *Let $S_{i,j}$ be given by $S_{i,j} = 0$, $0 \leq j < i$, $S_{0,j} = 1$, $j \geq 0$ and*

$$S_{i,j} = S_{i,j-1} + \zeta_{j-i+1} S_{i-1,j}, \qquad 1 \leq i \leq j.$$

If $\mathbf{m}_{n-1}$ is in the interior of the moment space, then

$$m_n = m_0 \sum_{i=0}^{\lfloor n/2 \rfloor} S_{i,n-i}^2 \prod_{j=1}^{n-2i} \zeta_j. \tag{14}$$

Let us remark that this theorem was proved for measures on $[0,1]$. It is easy to see that it is still available for measures on a compact support $[0,\xi_{\max}]$, thanks to relations given in Appendix A. Moreover, a truncated moment vector on $[0,+\infty)$ can always be seen as a moment vector of a measure with a compact support $[0,\xi_{\max}]$ (it can be seen by considering the corresponding quadrature), in such a way that this theorem is still valid in our case. This allows us to prove an important new property of the relation between the ζ_k and the moments.

Corollary 2.3. *The moment m_n can be written*

$$m_n = m_0 \left[\prod_{k=1}^{n} \zeta_k + P_n(\zeta_1, \ldots, \zeta_{n-1}) \right], \tag{15}$$

where $P_n(\zeta_1, \ldots, \zeta_{n-1})$ is a polynomial function of degree n.

Proof. It is easy to prove, by recursion, that $S_{i,j}$, for $i \leq j$ is a polynomial function of degree i of $(\zeta_k)_{k=1,\ldots,j-i+1}$, in such a way that $S^2_{i,n-i}$ is a polynomial function of degree $2i$ of $(\zeta_k)_{k=1,\ldots,n-2i+1}$. Then the property (15) follows immediately from (14). □

Let us remark that the term $m_0 P_n(\zeta_1, \ldots, \zeta_{n-1})$ corresponds to $m_k^-(\mathbf{m}_{k-1})$, since $m_n - m_n^-(\mathbf{m}_{n-1}) = m_0 \prod_{i=1}^{n} \zeta_i$, as shown in [24]. The polynomials P_n are given in Appendix C for $n = 1, \ldots, 7$.

3. Realizable finite volume kinetic schemes

The major issue of the numerical scheme developed here is to ensure the realizability of the vector $\mathbf{m}_N$. Indeed, Wright [26] showed that independent transport of moments with algorithms of order greater than one in space can result in the generation of invalid moment sets.

Let us develop a realizable numerical scheme for the 1D configuration. A generalization on 2D or 3D problems is straightforward for Cartesian meshes thanks to the method of lines.

3.1. General form of the finite volume kinetic scheme

Let us introduce a discretization of the spatial domain into cells $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ of center $x_j = \frac{x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}}}{2}$ and of width $\Delta x = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$. A time discretization $t^0 = 0 < t^1 < \cdots < t^n < \ldots$ is also used, with the time steps $\Delta t^n = t^{n+1} - t^n$. The properties of the averaged value of the moments on a cell is used here to define a numerical scheme.

3.1.1. Equations on the averaged value of the moments over a cell

One defines the characteristics $X(t;s,y)$ as the solution of

$$\begin{cases} \mathrm{d}_t X(t;s,y) = u(t, X(t;s,y)), \\ X(s;s,y) = y. \end{cases}$$

One also defines $J(t;s,y)$ the derivative of $y \mapsto X(t;s,y)$, for fixed values of t and s.

Proposition 3.1. *For $\mathbf{m}_N = (m_0, \ldots, m_N)^t$ a moment vector solution of (3), the mean value of each moment of order k at time t^n, denoted $\overline{m}^n_{k,j} = \dfrac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_k(t^n, x)\,dx$, satisfies*

$$\overline{m}^{n+1}_{k,j} = \overline{m}^n_{k,j} - \frac{\Delta t^n}{\Delta x} \left[\overline{F}^n_{k,j+\frac{1}{2}} - \overline{F}^n_{k,j-\frac{1}{2}} \right], \tag{16}$$

with

$$\overline{F}^n_{k,j+\frac{1}{2}} = \frac{1}{\Delta t^n}\int_{X(t^n;t^{n+1},x_{j+\frac{1}{2}})}^{x_{j+\frac{1}{2}}} m_k(t^n,x)dx. \tag{17}$$

Proof. Let us first remark that $\partial_t J(t;t^{n+1},x) = J(t;t^{n+1},x)\partial_x u(t,X(t;t^{n+1},x))$. Then, the function $t \mapsto m_k(t,X(t;t^{n+1},x))J(t;t^{n+1},x)$ is constant, since its derivative is:

$$\left[\partial_t m_k + u\partial_x m_k + m_k\partial_x u\right](t,X(t;t^{n+1},x))J(t;t^{n+1},x) = 0.$$

Using the values of this function at time t^n and t^{n+1}, we obtain the k^{th} component of the solution of (3) between t^n and t^{n+1}:

$$m_k(t^{n+1},x) = m_k(t^n,X(t^n;t^{n+1},x))J(t^n;t^{n+1},x).$$

Then, its averaged value at time t^{n+1} can be written, using the change of variables $\xi = X(t^n;t^{n+1},x)$:

$$\overline{m}^{n+1}_{k,j} = \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_k(t^n,X(t^n;t^{n+1},x))J(t^n;t^{n+1},x)\mathrm{d}x = \frac{1}{\Delta x}\int_{X(t^n;t^{n+1},x_{j-\frac{1}{2}})}^{X(t^n;t^{n+1},x_{j+\frac{1}{2}})} m_k(t^n,\xi)\mathrm{d}\xi.$$

This concludes the proof. □

This results allows us to develop a kinetic scheme.

3.1.2. Kinetic scheme

Let us denote $\mathbf{m}^n_j = (m^n_{0,j},\dots,m^n_{N,j})^t$ an approximation at time t^n of the averaged value of the moment vector over the cell j and $X_{j+\frac{1}{2}}$ an approximation of $X(t^n;t^{n+1},x_{j+\frac{1}{2}})$. In what follows, the CFL like number defined by:

$$\mathrm{CFL} = \max_{n,j}\left(u^n_{j,\max}\frac{\Delta t^n}{\Delta x}\right), \quad u^n_{j,\max} = \max\left\{u(t,x),\ x\in[x_{j-\frac{1}{2}},x_{j+\frac{1}{2}}],\ t\in[t^n,t^{n+1}]\right\} \tag{18}$$

is assumed to be smaller than one in such a way that $X(t^n;t^{n+1},x_{j+\frac{1}{2}})$ is in the cell j or $j+1$. Moreover, the same property is assumed for $X_{j+\frac{1}{2}}$.

First, let us define the approximation of $X(t^n;t^{n+1},x_{j+\frac{1}{2}})$ that will be used here. For a constant and uniform fluid velocity, the exact value is used

$$X_{j+\frac{1}{2}} = x_{j+\frac{1}{2}} - u\Delta t^n. \tag{19}$$

Otherwise, one need a good approximation of the velocity and a resolution of the ODE defining the characteristics with an at least second order of accuracy to obtain third order approximation of $X(t^n;t^{n+1},x_{j+\frac{1}{2}})$. For a stationary fluid velocity, a linear reconstruction of this velocity is done inside the cell. In what follows, we will assume that the fluid velocity is known at the interfaces and its value inside the cell is given by a linear interpolation. This leads to:

$$X_{j+\frac{1}{2}} = x_{j+\frac{1}{2}} - \frac{u(x_{j+\frac{1}{2}})}{\delta}\left(1-e^{-\delta\Delta t^n}\right), \tag{20}$$

where, denoting $\delta_j = \frac{u\left(x_{j+\frac{1}{2}}\right)-u\left(x_{j-\frac{1}{2}}\right)}{\Delta x}$: $\delta = \delta_j$ if $u(x_{j+\frac{1}{2}}) \geq 0$ and $\delta = \delta_{j+1}$ if $u(x_{j+\frac{1}{2}}) < 0$. Let us remark that $X_{j+\frac{1}{2}} < x_{j+\frac{1}{2}}$ in the first case and $X_{j+\frac{1}{2}} > x_{j+\frac{1}{2}}$ in the second one. For an

unstationary fluid velocity, a linear temporal and spatial interpolation of the fluid velocity can be used. Using an explicit second order Runge et Kutta method for the resolution of the ODEs defining the characteristics, one obtains:

$$X_{j+\frac{1}{2}} = x_{j+\frac{1}{2}} - \frac{\Delta t^n}{2}\left[(1-\delta\Delta t^n)\,u\left(t^{n+1}, x_{j+\frac{1}{2}}\right) + u\left(t^n, x_{j+\frac{1}{2}}\right)\right], \tag{21}$$

where, denoting $\delta_j = \frac{u\left(t^n, x_{j+\frac{1}{2}}\right) - u\left(t^n, x_{j-\frac{1}{2}}\right)}{\Delta x}$: $\delta = \delta_j$ if $u(t^{n+1}, x_{j+\frac{1}{2}}) \geq 0$ and $\delta = \delta_{j+1}$ if $u(t^{n+1}, x_{j+\frac{1}{2}}) < 0$. It leads to a third order approximation of $X(t^n; t^{n+1}, x_{j+\frac{1}{2}})$:

Lemma 3.2. *Let us assume that the CFL number is smaller than one and that the parameter $X_{j+\frac{1}{2}}$ is defined by (19), (20) or (21), depending of the effective dependence of $u(t,x)$ in t and x. Then $X_{j+\frac{1}{2}}$ is a third order approximation of $X(t^n; t^{n+1}, x_{j+\frac{1}{2}})$:*

$$X_{j+\frac{1}{2}} - X(t^n; t^{n+1}, x_{j+\frac{1}{2}}) = O(\Delta x^3).$$

Moreover, $X_{j+\frac{1}{2}} - X(t^n; t^{n+1}, x_{j+\frac{1}{2}}) - X_{j-\frac{1}{2}} + X(t^n; t^{n+1}, x_{j-\frac{1}{2}}) = O(\Delta x^4)$ and $x_{j+\frac{1}{2}} - X_{j+\frac{1}{2}} - x_{j-\frac{1}{2}} + X_{j-\frac{1}{2}} = O(\Delta x^2)$.

Proof. A Taylor expansion for $X(t^n; t^{n+1}, x_{j+\frac{1}{2}})$ leads to:

$$\begin{aligned} X(t^n; t^{n+1}, x_{j+\frac{1}{2}}) &= x_{j+\frac{1}{2}} + \int_{t^{n+1}}^{t^n} u(t, X(t; t^{n+1}, x_{j+\frac{1}{2}}))\mathrm{d}t \qquad (22)\\ &= x_{j+\frac{1}{2}} - u(t^{n+1}, x_{j+\frac{1}{2}})\Delta t^n + \frac{(\Delta t^n)^2}{2}\left[\partial_t u(t^{n+1}, x_{j+\frac{1}{2}}) + u(t^{n+1}, x_{j+\frac{1}{2}})\partial_x u(t^{n+1}, x_{j+\frac{1}{2}})\right] + O((\Delta t^n)^3). \end{aligned}$$

Formula (19) is exact for a constant fluid velocity and, since δ is a first order approximation of $\partial_x u(t^n, x_{j+\frac{1}{2}})$, the Taylor expansion of (20) in the case of a stationary fluid velocity and of (21) in the general case leads to the same zeroth, first and second order terms as in (22). Finally, both last results are shown from the same kind of developments, eventually adding the third order term. □

Following the property of the exact solution given in Proposition 3.1, the scheme that we consider here is given by the recursion formula:

$$\mathbf{m}_j^{n+1} = \mathbf{m}_j^n - \frac{\Delta t^n}{\Delta x}\left[\mathbf{F}_{j+\frac{1}{2}}^n - \mathbf{F}_{j-\frac{1}{2}}^n\right], \tag{23}$$

with

$$\mathbf{F}_{j+\frac{1}{2}}^n = \frac{1}{\Delta t^n}\int_{X_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathbf{m}^n(x)\mathrm{d}x, \tag{24}$$

the function $\mathbf{m}^n(x) = (m_0^n(x), \ldots, m_N^n(x))^t$ being defined from the moments $\mathbf{m}_j^n$, in a way described in Section 3.2.

3.1.3. Scheme properties

The properties of the scheme will depends on this reconstruction of the moments over the cells. We impose here the two following properties, assuming that $\mathbf{m}_j^n$ is in the moment space $\mathcal{M}_N$:

- [P1] The averaged value of the reconstructed moment vector $\mathbf{m}^n(x)$ over each cell j is $\mathbf{m}_j^n$:

$$\frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathbf{m}^n(x)\mathrm{d}x = \mathbf{m}_j^n.$$

- [P2] For each x, the vector $\mathbf{m}^n(x)$ is in the moment space $\mathcal{M}_N$.

For the accuracy, a third property is introduced:

- [P3] From an averaged moment vector $\overline{\mathbf{m}}_j = \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \overline{\mathbf{m}}(x)\mathrm{d}x$, the reconstruction $\mathbf{m}(x)$ is second order accurate: there exists a bounded function $\varphi_{\overline{\mathbf{m}}}$ such that

$$\mathbf{m}(x) = \overline{\mathbf{m}}(x) + \Delta x^2 \varphi_{\overline{\mathbf{m}}}(x), \tag{25}$$

with, for all j:

$$\forall \delta \in]-\Delta x, \Delta x[\qquad \mathcal{D}_j(\varphi_{\overline{\mathbf{m}}}) \stackrel{\text{def}}{=} \int_{x_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}+\delta} \varphi_{\overline{\mathbf{m}}}(x)\mathrm{d}x - \int_{x_{j-\frac{1}{2}}}^{x_{j-\frac{1}{2}}+\delta} \varphi_{\overline{\mathbf{m}}}(x)\mathrm{d}x = O(\Delta x^2).$$

This leads to the following theorem:

Theorem 3.3. *Let us assume that the CFL number is smaller than one. The finite volume scheme defined by (23,24) with the properties [P1] and [P2] of the moment reconstruction is realizable. With the additional property [P3] and using the value of $X_{j+\frac{1}{2}}$ defined by (19), (20) or (21), then a second order is obtained for the consistency error.*

Proof. The first property [P1], with the definition (24) of the fluxes implies:

$$\mathbf{m}_j^{n+1} = \frac{1}{\Delta x}\int_{X_{j-\frac{1}{2}}}^{X_{j+\frac{1}{2}}} \mathbf{m}^n(x)\mathrm{d}x. \tag{26}$$

Since the characteristics cannot cross themselves, one have $X_{j-\frac{1}{2}} < X_{j+\frac{1}{2}}$. Then [P2] ensures that $\mathbf{m}_j^{n+1}$ is in the moment space.
Moreover, from the exact solution $\mathbf{m} = (m_k)_{k\in\{0,1,\dots,N\}}$ of (3), the consistency error is defined by

$$\epsilon_j^n = \frac{\overline{\mathbf{m}}_j^{n+1} - \overline{\mathbf{m}}_j^n}{\Delta t^n} + \frac{\widetilde{\mathbf{F}}_{j+\frac{1}{2}}^n - \widetilde{\mathbf{F}}_{j-\frac{1}{2}}^n}{\Delta x},$$

where $\overline{\mathbf{m}}_j^{n+1} = (\overline{m}_{k,j}^n)_{k\in\{0,1,\dots,N\}}$ is the exact averaged moment vector, $\overline{m}_{k,j}^n = \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_k(t^n,x)\mathrm{d}x$, and the flux vector $\widetilde{\mathbf{F}}_{j+\frac{1}{2}}^n = \frac{1}{\Delta t^n}\int_{X_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathbf{m}^n(x)\mathrm{d}x$ is defined with the reconstruction $\mathbf{m}^n(x)$ obtained from the $\overline{\mathbf{m}}_j^n$. Thanks to Proposition 3.1, this can be written:

$$\epsilon_j^n = \frac{1}{\Delta t^n \Delta x}\left[I_1 + I_2\right],$$

with

$$I_1 = \int_{X_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (\mathbf{m}^n(x) - \mathbf{m}(t^n,x))\mathrm{d}x - \int_{X_{j-\frac{1}{2}}}^{x_{j-\frac{1}{2}}} (\mathbf{m}^n(x) - \mathbf{m}(t^n,x))\mathrm{d}x$$

and

$$I_2 = \int_{X_{j+\frac{1}{2}}}^{X\left(t^n;t^{n+1},x_{j+\frac{1}{2}}\right)} \mathbf{m}(t^n,x)\mathrm{d}x - \int_{X_{j-\frac{1}{2}}}^{X\left(t^n;t^{n+1},x_{j-\frac{1}{2}}\right)} \mathbf{m}(t^n,x)\mathrm{d}x.$$

Thanks to [P3], the first term I_1 can be written:

$$\begin{aligned} I_1 &= \Delta x^2 \left[\int_{X_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \varphi_{\mathbf{m}(t^n,.)}(x)\mathrm{d}x - \int_{X_{j-\frac{1}{2}}}^{x_{j-\frac{1}{2}}} \varphi_{\mathbf{m}(t^n,.)}(x)\mathrm{d}x \right] \\ &= -\Delta x^2 \left[\int_{x_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}+\delta} \varphi_{\mathbf{m}(t^n,.)}(x)\mathrm{d}x - \int_{x_{j-\frac{1}{2}}}^{x_{j-\frac{1}{2}}+\delta} \varphi_{\mathbf{m}(t^n,.)}(x)\mathrm{d}x - \int_{X_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}+X_{j-\frac{1}{2}}-x_{j-\frac{1}{2}}} \varphi_{\mathbf{m}(t^n,.)}(x)\mathrm{d}x \right], \end{aligned}$$

with $\delta = X_{j-\frac{1}{2}} - x_{j-\frac{1}{2}}$, which is in $]-\Delta x, \Delta x[$. Then, thanks to the property [P3] and Lemma 3.2, $I_1 = O(\Delta x^4)$. The second term I_2 can be written:

$$\begin{aligned} I_2 &= \mathbf{m}(t^n, X_{j+\frac{1}{2}})\left[X\left(t^n;t^{n+1},x_{j+\frac{1}{2}}\right) - X_{j+\frac{1}{2}}\right] - \mathbf{m}(t^n, X_{j-\frac{1}{2}})\left[X\left(t^n;t^{n+1},x_{j-\frac{1}{2}}\right) - X_{j-\frac{1}{2}}\right] + O(\Delta x^6) \\ &= \mathbf{m}(t^n, X_{j+\frac{1}{2}})\left[X\left(t^n;t^{n+1},x_{j+\frac{1}{2}}\right) - X_{j+\frac{1}{2}} - X\left(t^n;t^{n+1},x_{j-\frac{1}{2}}\right) + X_{j-\frac{1}{2}}\right] + O(\Delta x^4). \end{aligned}$$

The result of Lemma 3.2 allows to conclude the proof. □

Let us remark that a classical MUSCL type reconstruction for each moment separately satisfies the properties [P1] and [P3] except around extrema of the considered moments (see Appendix D), but this would inevitably lead to unrealizable moments $\mathbf{m}^n(x)$, due to the complex shape of the moment space [1], in which case the property [P2] would not be satisfied.

3.2. Reconstruction of the moments in the cell and flux computation

To complete the scheme (23,24) a spatial reconstruction of the moments from their mean value is provided, verifying at least properties [P1] and [P2]. Different kinds of such reconstructions can be found in the literature. The constant reconstruction is the simplest, but leads to low order schemes. It is given in Section 3.2.1.

Another one is based on a direct reconstruction at the level of the NDF. This was done in the completely different context of linear kinetic equations in slab geometry [43, 44]. In this context, the internal variable had a compact support and the advection velocity depended on the internal variable ξ in such a way that the moment equations had to be closed; this was done through the reconstruction of the NDF from the moments by entropy maximization [45]. Then, for each value of the internal variable, a spatial MUSCL type reconstruction of this NDF was used. And a quadrature was used to compute the fluxes, in such a way that the spatial reconstruction had to be done for a finite number of values of the internal variable. This is applied in our context in a quite different way presented in Appendix F, using EQMOM [21, 12] instead of the entropy maximization, which is not really adapted to NDF defined on $[0, +\infty)$. However, it does not allow to deal with the boundary of the moment space, since the NDF has to be a real function, and it induces a high supplementary cost related to the reconstruction (the computation time is then multiplied by at least 65 when using 5 moments, compared to the schemes described in this section). Moreover, it is also shown in Appendix F that the maximum principle on the moments can be lost.

Still in the case of compact support, the moments can be spatially reconstructed through a MUSCL type reconstruction of the canonical moments (see Appendix A for their definition and their link with the ζ_k). This was done in the context of population of inertial particles [1], their velocity belonging to the internal variables and being then treated differently compare to our application. Moreover, this reconstruction was limited to a model considering only four moments. In Section 3.2.2, since we are interested in moments on $[0, +\infty)$, instead of the canonical moments, the variables $(\zeta_k)_{k\in\{1,\dots,N\}}$ are reconstructed. Moreover, as in [1], the property [P1] is not directly verified, since the reconstructed moments are no more affine in the cell and the same kind of strategy is used to recover it, with eventually additional corrections here to ensure robustness, even for moment vectors close to the boundary of the moment space. And here, the reconstruction is done for any number of moments, eventually at the boundary of the moment space.

The last type of spatial reconstruction of the moments was done in [2]. It used a spatial reconstruction of the only weights of the quadratures associated with the moment vectors, and was used in another context, the internal variable being the velocity and evolving in $\mathbb{R}$. However, it was limited to an even number of moments, in the interior of the moment space and led to only a pseudo–second-order of accuracy, due to the fact that the variation of the abscissas was not taken into account in the reconstruction. In Section 3.2.3, a variant of this reconstruction is introduced, adapted to any number of moments, eventually at the boundary of the moment space and allowing a real second-order of accuracy for half the lowest order moments.

3.2.1. The constant reconstruction based kinetic scheme

From the mean value $m^n_{k,j}$ of the moment in the cell, the first idea is to consider a constant value of the moments into the cell: $\mathbf{m}^n(x) = \mathbf{m}^n_j$ for $x \in]x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}[$. The flux can then be written:

$$\mathbf{F}^n_{j+\frac{1}{2}} = \frac{\mathbf{m}^n_j}{\Delta t^n}\left(x_{j+\frac{1}{2}} - X_{j+\frac{1}{2}}\right)^+ - \frac{\mathbf{m}^n_{j+1}}{\Delta t^n}\left(X_{j+\frac{1}{2}} - x_{j+\frac{1}{2}}\right)^+.$$

where $u^+ = \max\{0, u\}$. Its convergence with a first order of accuracy is easily shown with the value of $X_{j+\frac{1}{2}}$ defined by (19), (20) or (21).

3.2.2. The ζ reconstruction based kinetic scheme (ζ kinetic scheme)

Let us denote $(\zeta^n_{k,j})_{k\in\{1,N\}}$ the ζ_k corresponding to the moment vector $\mathbf{m}^n_j$. For a fixed value of n, instead of reconstructing directly $\mathbf{m}^n(x)$, the corresponding $(\zeta_k(x))_{k\in\{1,\dots,N\}}$ as well as $m^n_0(x)$ are reconstructed. For m^n_0, a classical MUSCL type of reconstruction is used:

$$m^n_0(x) = m^n_{0,j} + D^n_{0,j}(x - x_j) \text{ for } x \in]x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}[.$$

For each k, a piecewise affine reconstruction $\zeta^n_k(x)$ is defined by:

$$\zeta^n_k(x) = \bar{\zeta}^n_{k,j} + D^n_{k,j}(x - x_j) \text{ for } x \in]x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}[.$$

Each $\bar{\zeta}^n_{k,j}$ is computed to ensure [P1]. Thanks to Corollary 2.3, this leads to:

$$\begin{aligned}\Delta x\, m^n_{k,j} =& \bar{\zeta}^n_{k,j}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m^n_0(x)\prod_{i=1}^{k-1}\zeta^n_i(x)\mathrm{d}x + D^n_{k,j}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}(x - x_j)m^n_0(x)\prod_{i=1}^{k-1}\zeta^n_i(x)\mathrm{d}x\\ &+ \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m^n_0(x)P_k(\zeta^n_1(x), \dots, \zeta^n_{k-1}(x))\mathrm{d}x.\end{aligned}$$

Let us write $\bar{\zeta}^n_{k,j} = a_{k,j} + b_{k,j} D^n_{k,j}$, where $a_{k,j}$ and $b_{k,j}$ are defined by:

$$a_{k,j} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_0^n(x) \prod_{i=1}^{k-1} \zeta_i^n(x) \mathrm{d}x = \Delta x \, m^n_{k,j} - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_0^n(x) P_k(\zeta_1^n(x), \ldots, \zeta_{k-1}^n(x)) \mathrm{d}x, \tag{27}$$

$$b_{k,j} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_0^n(x) \prod_{i=1}^{k-1} \zeta_i^n(x) \mathrm{d}x = - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (x - x_j) m_0^n(x) \prod_{i=1}^{k-1} \zeta_i^n(x) \mathrm{d}x. \tag{28}$$

Let us remark that $|b_{k,j}|$ is smaller than $\frac{\Delta x}{2}$. Numerically, the term $P_k(\zeta_1^n(x), \ldots, \zeta_{k-1}^n(x))$ is computed by using the reverse algorithm (see Section 2.4), setting $\zeta_k = 0$. Moreover, the integrals, which are integrals of polynomial functions, are computed using a Gauss-Legendre quadrature with $\left\lceil \frac{N}{2} \right\rceil + 1$ points, which is enough to obtain their exact value.

For m_0, a minmod limiter is used:

$$D^n_{0,j} = minmod \left(\frac{m^n_{0,j+1} - m^n_{0,j}}{\Delta x}, \frac{m^n_{0,j} - m^n_{0,j-1}}{\Delta x} \right). \tag{29}$$

Moreover, as in [46, 1], the slopes $D^n_{k,j}$ are such that $\zeta^n_k(x)$ stays between $\min\{\zeta^n_{k,j-1}, \zeta^n_{k,j}, \zeta^n_{k,j+1}\}$ and $\max\{\zeta^n_{k,j-1}, \zeta^n_{k,j}, \zeta^n_{k,j+1}\}$ in each cell j, as soon as it is the case for $a_{k,j}$. One can use [1]:

$$D^n_{k,j} = \begin{cases} \min \left(\dfrac{|\zeta^n_{k,j+1} - a_{k,j}|}{\Delta x + 2b_{k,j}}, \dfrac{|a_{k,j} - \zeta^n_{k,j-1}|}{\Delta x - 2b_{k,j}}, \dfrac{2a_{k,j}}{\Delta x - 2b_{k,j}} \right) & \text{if } \zeta^n_{k,j-1} < \zeta^n_{k,j} < \zeta^n_{k,j+1}, \\ -\min \left(\dfrac{|\zeta^n_{k,j+1} - a_{k,j}|}{\Delta x + 2b_{k,j}}, \dfrac{|a_{k,j} - \zeta^n_{k,j-1}|}{\Delta x - 2b_{k,j}}, \dfrac{2a_{k,j}}{\Delta x + 2b_{k,j}} \right) & \text{if } \zeta^n_{k,j-1} > \zeta^n_{k,j} > \zeta^n_{k,j+1}. \end{cases} \tag{30}$$

The last condition (in the min) ensures the non-negativity of the reconstruction, as soon as $a_{k,j}$ is non-negative. Unfortunately, $a_{k,j}$ can eventually be negative, meaning that, even with a constant reconstruction for $\zeta^n_k(x)$, one can obtain unrealizable moments. In this case, a correction has to be done on the slopes $D^n_{i,j}$, with $i < k$: they are reduced successively for $i = k_0, ..., k-1$ where k_0 is the maximal value such that $a_{k,j}$ is positive if the $D^n_{i,j}$ were equal to zero for $i \in [k_0, k-1]$. For that, starting with $i = k_0$ till $i = k-1$, each $D^n_{i,j}$ is multiplied by 0.9 once or several times (at most 5 times here) and if it is not sufficient, it is set to zero. In the worst case, all slopes $D^n_{i,j}$ are then set to zero, for $i \geq 1$. But when needed, only one limitation is often sufficient. Let us remark that the consequence of such kind of a posteriori corrections, if one or several slopes is set to zero, is the reduction of the polynomial degree of $m_k(x)$, as done in other kind of method such as MOOD [47], in a quite different way. The algorithm detailing the reconstruction procedure and especially the corrections ensuring the positivity of $a_{k,j}$ is given in Appendix E.

The verification of property [P3] for this reconstruction is complex. Thanks to the polynomial form of the functions, it can be shown for regular initial conditions that, except near the boundary of the moment space (where one of several $\zeta^n_{i,j}$ is small), $a_{k,j} = \zeta^n_{k,j} + O(\Delta x^2)$ and $b_{k,j} = O(\Delta x^2)$, in such a way that $\bar{\zeta}^n_{k,j} = \zeta^n_{k,j} + O(\Delta x^2)$. This means that, for a small enough value of Δx and for moments far enough from the boundary of the moment space, the parameter $a_{k,j}$ should be positive in such a way that no correction has to be done. That is why this case was not encountered in [1] where the boundary of the moment space were not dealt with. In this context, *i.e.* far from the boundary of the moment space, one can then show that, except near any extrema of the function $\zeta_k(x)$, $\mathbf{m}(x) = \overline{\mathbf{m}}(x) + \Delta x^2 \varphi_{\overline{\mathbf{m}}}(x)$, where $\varphi_{\overline{\mathbf{m}}}$ is bounded. However, the additional condition on

$\varphi_{\overline{\mathbf{m}}}$ is hard to show, even if we guess that it is valid, except near extrema of any ζ_k or near the boundary of the moment space. The order of accuracy of the scheme will be checked numerically on some examples.

The fluxes can then be written, thanks to Corollary 2.3:

$$F^n_{k,j+\frac{1}{2}} = \frac{1}{\Delta t^n}\int_{X_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_0^n(x)\left[\prod_{j=1}^{k}\zeta_j^n(x) + P_k(\zeta_1^n(x),\dots,\zeta_{k-1}^n(x))\right]\mathrm{d}x, \tag{31}$$

They are computed using a Gauss-Legendre quadrature with $\left\lceil\frac{N}{2}\right\rceil+1$ points, which gives their exact value since one has to integrate polynomial of degree at most $N+1$. The corresponding scheme is called "ζ reconstruction based kinetic scheme", abbreviated to "ζ kinetic scheme". Let us remark that the resolution of the equation on m_0 does not depend on the other moments, even if some corrections are needed for the positivity of the $a_{k,j}$. Moreover, the corresponding scheme is TVD for the constant velocity case. It is also easy to see that a maximum principle is obtained for $\zeta_1 = m_1/m_0$ in this case.

3.2.3. The quadrature weights reconstruction based kinetic scheme (QW kinetic scheme)

Similarly to what was done in [2], the reconstruction is based here on the quadrature of the moment sets, but in a different way. Let us then define the quadratures weights $(w_{\alpha,j})_{\alpha\in\{1,\dots,p\}}$ and abscissas $(\xi_{\alpha,j})_{\alpha\in\{1,\dots,p\}}$ of the moment set $(m^n_{k,j})_{k\in\{0,\dots,N\}}$, such that:

$$m^n_{k,j} = \sum_{\alpha=1}^{p} w_{\alpha,j}\xi^k_{\alpha,j}, \qquad k\in\{0,\dots,N\},$$

with $p = \left\lfloor\frac{N}{2}\right\rfloor+1$. If the number of moments is even, it is the classical Gauss quadrature with $N=2p-1$. Then, abscissas and weights are deduced from the eigenvalues and eigenvectors of the Jacobi matrix with coefficients α_k and $\sqrt{\beta_k}$ [48], where α_k and β_k are the coefficient of the three term recurrence relation (9). Otherwise, $N=2(p-1)$ and a Gauss-Radau quadrature is used, meaning that an abscissa is set to zero: $\xi_{j,1}=0$. They are computed in the same way as for the classical Gauss quadrature, except that α_{p-1} is set to $-\beta_{p-1}\frac{\Pi_{p-1}(0)}{\Pi_p(0)}$ [48].

Like in [2], the reconstruction of the moments inside the cell j is done through a reconstruction of the weights:

$$m^n_k(x) = \sum_{\alpha=1}^{p} w_\alpha(x)\xi^k_{\alpha,j}, \qquad w_\alpha(x) = w_{\alpha,j} + D^n_{\alpha,j}(x-x_j), \quad x\in[x_{j-\frac{1}{2}},x_{j+\frac{1}{2}}].$$

The averaged value of $m^n_k(x)$ on the j^{th} cell is then automatically $m^n_{k,j}$. But to define the slopes, unlike in [2], one uses here a quadrature of the neighbor cells with the same abscissas:

$$m^n_{k,j-1} = \sum_{\alpha=1}^{p} w^-_{\alpha,j}\xi^k_{\alpha,j}, \quad m^n_{k,j+1} = \sum_{\alpha=1}^{p} w^+_{\alpha,j}\xi^k_{\alpha,j}, \quad k\in\{0,\dots,p-1\}. \tag{32}$$

This allows also to consider moments at the boundary of the moment space, where p is then reduced, which was not possible with the method used in [2], where the case of the neighboring cells with a representation by a quadrature with a different number of abscissas could not be taken into

account easily. Here, the weights $(w^-_{\alpha,j})_{\alpha\in\{1,\dots,p\}}$ and $(w^+_{\alpha,j})_{\alpha\in\{1,\dots,p\}}$ are well defined since they are the solution of a linear system with a Vandermonde matrix of coefficients $\xi_{\alpha,j}$, which are distinct. If a regular initial condition is considered for (3) and if $(m^n_{k,j})_{k\in\{0,\dots,N\}}$ is far from the boundary of the moment space, then these weights $(w^-_{\alpha,j})_{\alpha\in\{1,\dots,p\}}$ and $(w^+_{\alpha,j})_{\alpha\in\{1,\dots,p\}}$ are usually non-negative since the moment sets for $j-1$, j and $j+1$ are close to each other and the weights $w_{\alpha,j}$ are far from zero. But if it is not the case, the most negative $w^+_{\alpha,j}$ or $w^-_{\alpha,j}$ is set to zero and the corresponding linear system in (32) is solved for $k\in\{0,\dots,p-2\}$, thus eliminating the corresponding abscissa $\xi_{\alpha,j}$. This operation is reproduced till all the $w^\pm_{\alpha,j}$ are non-negative. The corresponding algorithm is detailed in Appendix E. At worst, only one abscissa will stay in (32). But for the distributions used in this paper, when needed, only one or two abscissas has to be eliminated when using 10 moments. One then defines the slopes by using a minmod limiter:

$$D^n_{\alpha,j}=\frac{1}{2}\left(sgn(w^+_{\alpha,j}-w_{\alpha,j})+sgn(w_{\alpha,j}-w^-_{\alpha,j})\right)\min\left(\frac{|w^+_{\alpha,j}-w_{\alpha,j}|}{\Delta x},\frac{|w_{\alpha,j}-w^-_{\alpha,j}|}{\Delta x}\right).$$

In order to verify the property [P3] let us rewrite this slope in the following way, in the same way as in Appendix D:

$$D^n_{\alpha,j}=\frac{w^+_{\alpha,j}-w_{\alpha,j}}{\Delta x}\Phi(\theta_{\alpha_j}),\qquad \theta_{\alpha,j}=\frac{w_{\alpha,j}-w^-_{\alpha,j}}{w^+_{\alpha,j}-w_{\alpha,j}}.$$

The function Φ then corresponds to the classical minmod flux limiter [49]: $\Phi(\theta)=minmod(1,\theta)$. For the k^{th} order moment, this corresponds to the following reconstruction, for $x\in(x_{j-\frac{1}{2}},x_{j+\frac{1}{2}})$:

$$m^n_k(x)=m^n_{k,j}+(x-x_j)D^n_{k,j},\qquad D^n_{k,j}=\sum_{\alpha=1}^p D^n_{\alpha,j}\xi^k_{\alpha,j}=\sum_{\alpha=1}^p\frac{w^+_{\alpha,j}-w_{\alpha,j}}{\Delta x}\xi^k_{\alpha,j}\Phi(\theta_{\alpha,j}).$$

Thanks to the definition of the weights, the corresponding slope can be rewritten, if $k\leq p-1$:

$$D^n_{k,j}=\frac{m^n_{k,j+1}-m^n_{k,j}}{\Delta x}+\sum_{\alpha=1}^p\frac{w^+_{\alpha,j}-w_{\alpha,j}}{\Delta x}\xi^k_{\alpha,j}\left[\Phi(\theta_{\alpha,j})-1\right].$$

With only the first term of the slope, the reconstruction of the moment would correspond to a second order approximation without any slope limiter. Moreover, let us denote $\widetilde{w}_{\alpha,j}(x)$ the weights of a quadrature corresponding to the analytical solution with imposed abscissas $(\xi_{\alpha,j})_{k\in\{1,\dots,p\}}$: $\overline{m}_k(x)=\sum_{\alpha=1}^p\widetilde{w}_{\alpha,j}(x)\xi^k_{\alpha,j}$ for $k=0,\dots,p-1$. Except near the extrema of this weights, the term $\Phi(\theta_{\alpha,j})-1$ is first order accurate. Then, in this case, the reconstruction $m^n_k(x)$ from the analytical averaged moments is of the form given in [P3], $m_k(x)=\overline{m}_k(x)+\Delta x^2\varphi_{\overline{m}_k}(x)$. Moreover, for example if $\delta\in]0,\Delta x[$:

$$\mathcal{D}_j(\varphi_{\overline{m}_k})=O(\Delta x^2)+\frac{\delta(\Delta x-\delta)}{2\Delta x^2}\left\{\sum_{\alpha=1}^p\frac{w^+_{\alpha,j+1}-w_{\alpha,j+1}}{\Delta x}\xi^k_{\alpha,j+1}\left[\Phi(\theta_{\alpha,j+1})-1\right]-\sum_{\alpha=1}^p\frac{w^+_{\alpha,j}-w_{\alpha,j}}{\Delta x}\xi^k_{\alpha,j}\left[\Phi(\theta_{\alpha,j})-1\right]\right\}.$$

Thanks to the use of the $\widetilde{w}_{\alpha,j}(x)$, the previous expression can be written:

$$\begin{aligned}\mathcal{D}_j(\varphi_{\overline{m}_k})=&O(\Delta x^2)+\frac{\delta(\Delta x-\delta)}{2\Delta x^2}\left\{\sum_{\alpha=1}^p\left[\widetilde{w}'_{\alpha,j+1}(x_{j+\frac{1}{2}})-\widetilde{w}'_{\alpha,j}(x_{j+\frac{1}{2}})\right]\xi^k_{\alpha,j+1}\left[\Phi(\theta_{\alpha,j+1})-1\right]\right.\\&\left.+\sum_{\alpha=1}^p\widetilde{w}'_{\alpha,j}(x_{j+\frac{1}{2}})\left[\xi^k_{\alpha,j+1}-\xi^k_{\alpha,j}\right]\left[\Phi(\theta_{\alpha,j+1})-1\right]+\sum_{\alpha=1}^p\widetilde{w}'_{\alpha,j}(x_{j+\frac{1}{2}})\xi^k_{\alpha,j+1}\left[\Phi(\theta_{\alpha,j+1})-\Phi(\theta_{\alpha,j})\right].\right\}.\end{aligned}$$

Except near the extrema of the weight $\widetilde{w}_{\alpha,j}(x)$, the term $\Phi(\theta_{\alpha,j+1})-\Phi(\theta_{\alpha,j})$ is second order accurate. Moreover, except near the boundary of the moment space, the quadrature reconstruction is stable, in such a way that $\xi^k_{\alpha,j+1}-\xi^k_{\alpha,j}$ and $\widetilde{w}'_{\alpha,j+1}(x_{j+\frac{1}{2}})-\widetilde{w}'_{\alpha,j}(x_{j+\frac{1}{2}})$ are then also first order terms. Thus, except near the boundary of the moment space and near the extrema of the quadrature weights, the property [P3] is checked for the moments of order 0 to $p-1$.

Finally, the fluxes are written:

$$\begin{aligned}\mathbf{F}^n_{j+\frac{1}{2}} =& \frac{\sum_{\alpha=1}^p \left(w_{\alpha,j}+\frac{1}{2}D^n_{\alpha,j}(X_{j+\frac{1}{2}}-x_{j-\frac{1}{2}})\right)\boldsymbol{\Xi}_{\alpha,j}}{\Delta t^n}\left(x_{j+\frac{1}{2}}-X_{j+\frac{1}{2}}\right)^+ \\ &- \frac{\sum_{\alpha=1}^p \left(w_{\alpha,j+1}-\frac{1}{2}D^n_{\alpha,j+1}(x_{j+\frac{3}{2}}-X_{j+\frac{1}{2}})\right)\boldsymbol{\Xi}_{\alpha,j+1}}{\Delta t^n}\left(X_{j+\frac{1}{2}}-x_{j+\frac{1}{2}}\right)^+.\end{aligned}$$

with $\boldsymbol{\Xi}_{\alpha,j}=(1,\xi_{\alpha,j},\dots,\xi^N_{\alpha,j})^T$. This scheme is called "quadrature weights (QW) reconstruction based kinetic scheme", abbreviated to "QW kinetic scheme". Let us remark that, since the limitation is done separately for each weight, the global reconstruction can not guarantee a maximum principle on any moments. One can however remark that the $\xi_{\alpha,j}$ are bounded by the maximal and minimal values of the abscissas corresponding to the initial moments. Indeed, the new moments $\mathbf{m}^{n+1}_j$ correspond to a sum of weighted Dirac delta functions at $\xi_{\alpha,j}$ and $\xi_{j\pm1,\alpha}$ with $\alpha=1,\dots,p$. Then, the new abscissas are necessarily inbetween these values.

3.2.4. Dealing with the boundary of the moment space

It can happen that, at some point, the moment vector is at the boundary of the moment space, with eventually for one of the neighbor cells, a moment vector at another boundary (different value of $\mathcal{N}(\mathbf{m}_N)$) or in the interior of the moment space. In this case, the developed schemes can locally loose their second order of accuracy. But the key point is their realizability, which can be very difficult to preserve in this case, which was not taken into account in [1] or [2]. However, here, the developed schemes are still realizable if some precautions are taken. For the ζ reconstruction, the ζ_k for $k>\mathcal{N}(\mathbf{m}_N)$ are not defined but are set to zero and eventually some corrections are done on the slopes to guaranty the positivity of $a_{k,j}$, as explained in Section 3.2.2. For the QW reconstruction, the value of p has just to be reduced to $\left\lfloor\frac{\mathcal{N}(\mathbf{m}_N)+1}{2}\right\rfloor$. Let us remark that for the quadrature, such kind of adaptation of the number of weights were introduced [50], but here it is directly based on the evaluation of $\mathcal{N}(\mathbf{m}^n_N)$, through the computation of the ζ_k, as for the ζ reconstruction. Since they are deduced from the coefficients of the three term recurrence relation used for the computation of the quadrature thanks to (10), they are computed in any cases, with at worst a marginal increase of the cost.

Numerically, an additional difficulty comes from the detection of the boundary of the moment space, *i.e.* the evaluation of $\mathcal{N}(\mathbf{m}_N)$, the computation of the ζ_k being ill-conditioned near this boundary. Indeed, when a ζ_k is positive but very small, the computation of the next ζ_i can give anything. Then, a small parameter ϵ is introduced here, and $\mathcal{N}(\mathbf{m}_N)$ is replaced by $\mathcal{N}_\epsilon(\mathbf{m}_N)$ defined as the minimum value of k such that $\zeta_k<\epsilon$ (set to $N+1$ if all the ζ_k are greater than ϵ). Moreover, for $i\geq\mathcal{N}_\epsilon(\mathbf{m}_N)$, ζ_i is then set to zero and the corresponding vector is then projected: the moments are computed from the new set of $(\zeta_k)_{k\in\{1,\dots,N\}}$. In practice, we used in this work $\epsilon=10^{-7}$.

4. A realizable simplified finite volume scheme

The accuracy of the kinetic schemes comes from the accuracy of the spatial reconstruction as well as the correct evaluation of the characteristics, to characterize the part of the cell containing the particles that will be transferred to another one during the time step. For this last point, a generalization to multi-dimensional unstructured meshes is not evident. In this section, we then consider semi-discretized finite volume schemes of the form:

$$\mathrm{d}_t \mathbf{m}_j(t) = -\frac{1}{\Delta x}\left[\mathbf{F}_{j+\frac{1}{2}}(t) - \mathbf{F}_{j-\frac{1}{2}}(t)\right], \qquad \mathbf{F}_{j+\frac{1}{2}}(t) = u(t, x_{j+\frac{1}{2}})\mathbf{m}_{j+\frac{1}{2}}(t), \tag{33}$$

where only the reconstructed value $\mathbf{m}_{j+\frac{1}{2}}(t)$ of the moment at the interface $x_{j+\frac{1}{2}}$ is considered.

The system (33) will then be solved thanks to a strong stability preserving (SSP) explicit Runge-Kutta method [51], a second order one here. It is a convex combination of Euler explicit time steps. The realizability then only have to be shown for the explicit Euler method:

$$\mathbf{m}_j^{n+1} = \mathbf{m}_j^n - \frac{\Delta t^n}{\Delta x}\left[\mathbf{F}^n_{j+\frac{1}{2}} - \mathbf{F}^n_{j-\frac{1}{2}}\right], \tag{34}$$

with $\mathbf{F}^n_{j+\frac{1}{2}} = u(t^n, x_{j+\frac{1}{2}})\mathbf{m}^n_{j+\frac{1}{2}}$. Moreover, to deal with the boundary of the moment space, a numerical projection is also done after each of these Euler explicit time steps, as described in Section 3.2.4.

Two types of reconstruction are done to define the moment vector $\mathbf{m}_{j+\frac{1}{2}}$ at the interface, as for the kinetic scheme: a first one using the ζ_k variables and a second one using the weights of the quadrature. The realizability of the scheme is then shown in each case.

4.1. The ζ reconstruction based simplified scheme (ζ simplified scheme)

Here, the reconstruction is quite different from the one of Section 3.2.2 in order to be able to ensure the realizability with the explicit Euler method. Indeed, from the idea of Berthon [52], the cell j is split in three parts of size $\Delta x/3$ here, on which the reconstruction of the moment vector is constant. It is denoted $\mathbf{m}_j^-$ for the left part, $\mathbf{m}_j^*$ for the middle part and $\mathbf{m}_j^+$ for the right part, in such a way that:

$$\mathbf{m}_j^n = \frac{1}{3}\left(\mathbf{m}_j^- + \mathbf{m}_j^* + \mathbf{m}_j^+\right).$$

The values of the ζ_k corresponding to the moment vectors $\mathbf{m}_j^n$, $\mathbf{m}_j^-$, $\mathbf{m}_j^*$ and $\mathbf{m}_j^+$ are denoted respectively $\zeta^n_{k,j}$, $\zeta^-_{k,j}$, $\zeta^*_{k,j}$ and $\zeta^+_{k,j}$. Let us remark that Berthon [52] used this kind of reconstruction in a completely different context of the Euler equations (*i.e.* considering moments of order 0, 1 and 2 in velocity, on $\mathbb{R}$). This has then to be adapted in our context of high order moments in size. Moreover, here, the role of $\mathbf{m}_j^*$ is just to guaranty the realizability of the scheme by doing a reconstruction of realizable moments in the entire cell (property [P2]). But this kind of reconstruction is not used in the context the kinetic scheme of Section 3, since it would necessarily induce a loss of accuracy, the reconstruction being only accurate at the bounds of the cell.

The reconstruction is then done in the following way. For the left and right values, a classical MUSCL type of reconstruction is used:

$$m_0^{\pm} = m^n_{0,j} \pm D^n_{0,j}\frac{\Delta x}{2}, \qquad \zeta^{\pm}_{k,j} = \zeta^n_{k,j} \pm D^n_{k,j}\frac{\Delta x}{2}, \quad k \in \{1, \dots, N\}, \tag{35}$$

with the minmod limiters:

$$D^n_{0,j} = minmod\left(\frac{m^n_{0,j+1} - m^n_{0,j}}{\Delta x}, \frac{m^n_{0,j} - m^n_{0,j-1}}{\Delta x}\right), \tag{36}$$

$$D^n_{k,j} = minmod\left(\frac{\zeta^n_{k,j+1} - \zeta^n_{k,j}}{\Delta x}, \frac{\zeta^n_{k,j} - \zeta^n_{k,j-1}}{\Delta x}\right). \tag{37}$$

Then, the values of $\mathbf{m}^+_j$ and $\mathbf{m}^-_j$ can be computed. However, the vector $\mathbf{m}^*_j$, given by $3\mathbf{m}^n_j - \mathbf{m}^+_j - \mathbf{m}^-_j$, is not necessarily realizable. In this case, some values of the $D^n_{k,j}$, for $k > 0$, have to be reduced. To ensure the realizability of $(m^*_{0,j}, m^*_{1,j})$, the following limitation on $D^n_{1,j}$ has to be added:

$$D^n_{0,j} D^n_{1,j} < \frac{2}{\Delta x^2} m^n_{1,j}. \tag{38}$$

For the higher order moments, it is harder to derive analytical formulas. Instead, if $\mathbf{m}^*_j$ is not in the moment space, a correction has to be done on the slopes $D^n_{i,j}$, with $i < k$: they are reduced successively for $i = k_0, ..., k-1$ where k_0 is the maximal value such that $(m^*_{0,j}, \dots, m^*_{k,j})$ is realizable if the $D^n_{i,j}$ were equal to zero for $i \in [k_0, k-1]$. For that, starting with $i = k_0$ till $i = k-1$, each $D^n_{i,j}$ is divided by 2 or set to zero if it is not sufficient. The corresponding algorithm is detailed in Appendix E.

Once the reconstruction is done, the value of the fluxes is obtained by

$$\mathbf{F}^n_{j+\frac{1}{2}} = \max\{u(t^n, x_{j+\frac{1}{2}}), 0\}\mathbf{m}^+_j + \min\{u(t^n, x_{j+\frac{1}{2}}), 0\}\mathbf{m}^-_{j+1}. \tag{39}$$

The corresponding scheme is called "ζ reconstruction based simplified scheme", abbreviated to "ζ simplified scheme". Under some limitation on the CFL number, it is realizable, as shown in the following theorem.

Theorem 4.1. *From realizable moment vectors $(\mathbf{m}^n_j)_j$, the equations (34,39) define a realizable moment vector $\mathbf{m}^{n+1}_j$ if the CFL number is smaller than 1/3.*

Proof. The vector $\mathbf{m}^{n+1}_j$ is given by:

$$\mathbf{m}^{n+1}_j = \mathbf{m}^n_j - \lambda_j \mathbf{m}^+_j - \tilde{\mu}_{j-1}\mathbf{m}^-_j + \mu_j \mathbf{m}^-_{j+1} + \tilde{\lambda}_{j-1}\mathbf{m}^+_{j-1},$$

where $\lambda_j = \frac{\Delta t^n}{\Delta x}\max\{u(t^n, x_{j+\frac{1}{2}}), 0\}$, $\mu_j = -\frac{\Delta t^n}{\Delta x}\min\{u(t^n, x_{j+\frac{1}{2}}), 0\}$, $\tilde{\lambda}_{j-1} = \frac{\Delta t^n}{\Delta x}\max\{u(t^n, x_{j-\frac{1}{2}}), 0\}$, $\tilde{\mu}_{j-1} = -\frac{\Delta t^n}{\Delta x}\min\{u(t^n, x_{j-\frac{1}{2}}), 0\}$ are coefficients between 0 and 1/3. The last two terms define some moment vectors and the rest can be written $\left(\frac{1}{3} - \lambda_j\right)\mathbf{m}^+_j + \left(\frac{1}{3} - \tilde{\mu}_{j-1}\right)\mathbf{m}^-_j + \frac{1}{3}\mathbf{m}^*_j$ and is also realizable. □

4.2. The quadrature weights reconstruction based simplified scheme (QW simplified scheme)

The same reconstruction of the weights is done here as for the kinetic scheme in Section 3.2.3. Using the same notation, the flux is then given by:

$$\begin{aligned}\mathbf{F}_{j+\frac{1}{2}}(t) = &\max\{u(t, x_{j+\frac{1}{2}}), 0\}\sum_{\alpha=1}^{p}\left(w_{\alpha,j} + \frac{\Delta x}{2}D^n_{\alpha,j}\right)\mathbf{\Xi}_{\alpha,j} \\ &+ \min\{u(t, x_{j+\frac{1}{2}}), 0\}\sum_{\alpha=1}^{p}\left(w_{\alpha,j+1} - \frac{\Delta x}{2}D^n_{\alpha,j+1}\right)\mathbf{\Xi}_{\alpha,j+1}.\end{aligned} \tag{40}$$

The corresponding scheme is called "quadrature weights (QW) reconstruction based simplified scheme", abbreviated to "QW simplified scheme" and is then realizable under some less restrictive limitation on the CFL number than for the ζ-simplified scheme.

Theorem 4.2. *From realizable moment vectors $(\mathbf{m}_j^n)_j$, the equations (34,40) define a realizable moment vector $\mathbf{m}_j^{n+1}$ if the CFL number is smaller than 1/2.*

Proof. The vector $\mathbf{m}_j^{n+1}$ can be written:

$$\mathbf{m}_j^{n+1} = \sum_{\alpha=1}^{p} \left[w_{\alpha,j} - \lambda_j \left(w_{\alpha,j} + \frac{\Delta x}{2} D_{\alpha,j}^n \right) - \tilde{\mu}_{j-1} \left(w_{\alpha,j} - \frac{\Delta x}{2} D_{\alpha,j}^n \right) \right] \mathbf{\Xi}_{\alpha,j} + \mu_j \sum_{\alpha=1}^{p} \left(w_{\alpha,j+1} - \frac{\Delta x}{2} D_{\alpha,j+1}^n \right) \mathbf{\Xi}_{\alpha,j+1} + \tilde{\lambda}_{j-1} \sum_{\alpha=1}^{p} \left(w_{\alpha,j-1} + \frac{\Delta x}{2} D_{\alpha,j-1}^n \right) \mathbf{\Xi}_{\alpha,j-1},$$

with the same definition of λ_j, μ_j, $\tilde{\lambda}_{j-1}$ and $\tilde{\mu}_{j-1}$ as in the proof of Theorem 4.1 but now belonging to $[0, 1/2]$. The only thing to check is the non-negativity of the weights in the first summation. If $w_{\alpha,j}$ is zero, then $D_{\alpha,j}^n = 0$ and this term is zero. Otherwise, $w_{\alpha,j}$ is positive and this term can be written:

$$w_{\alpha,j} \left[1 - \left(1 + \frac{\Delta x}{2w_{\alpha,j}} D_{\alpha,j}^n \right) \lambda_j - \left(1 - \frac{\Delta x}{2w_{\alpha,j}} D_{\alpha,j}^n \right) \tilde{\mu}_{j-1} \right],$$

which is non-negative, since necessarily $|D_{\alpha,j}^n| \leq \frac{2w_{\alpha,j}}{\Delta x}$ and λ_j and $\tilde{\mu}_{j-1}$ are smaller than 1/2. □

5. Verification: comparisons with analytical solution in 1D configurations

Two configurations are studied here for which analytical solutions are available. For the first one, moments are transported by a constant fluid velocity in a periodic domain, whereas for the second one, moments are transported by an unsteady and compressible fluid velocity. In both cases, three initial conditions can be considered, defined either from the NDF or from the ζ_k: a regular one with moments in the interior of the moment space, an oscillating one where the ζ_k oscillate at different frequencies to test the robustness of the methods and a multi-modal one, reaching the boundary of the moment space and typical of what can be obtained in physical systems were nucleation and aggregation occur. These initial conditions are first described before comparing the results of our schemes to the analytical solution in both test cases. Then, the schemes are tested in 1D configurations with a constant or a compressible fluid velocity, considering a high number of moments, equal to 10 unless mentioned otherwise.

5.1. Initial moments

5.1.1. Regular initial NDF

A regular initial condition is defined through the following beta-NDF:

$$f_0(\xi, x) = 16x^2(1-x)^2 \frac{\xi^{\lambda(x)}(1-\xi)^{\mu(x)}}{\beta(\lambda(x), \mu(x))},$$

with

$$\lambda(x) = \frac{7}{2} + \frac{3}{2}\sin(2\pi x), \quad \mu(x) = \frac{7}{2} - \frac{3}{2}\cos(2\pi x)$$

and the initial moments are defined by $m_k^0(x) = \int_0^1 \xi^k f_0(\xi, x) \mathrm{d}\xi$. These moments for $k \leq 6$, as well as the corresponding ζ_k are plotted in Fig. 1. They are regular and periodic.

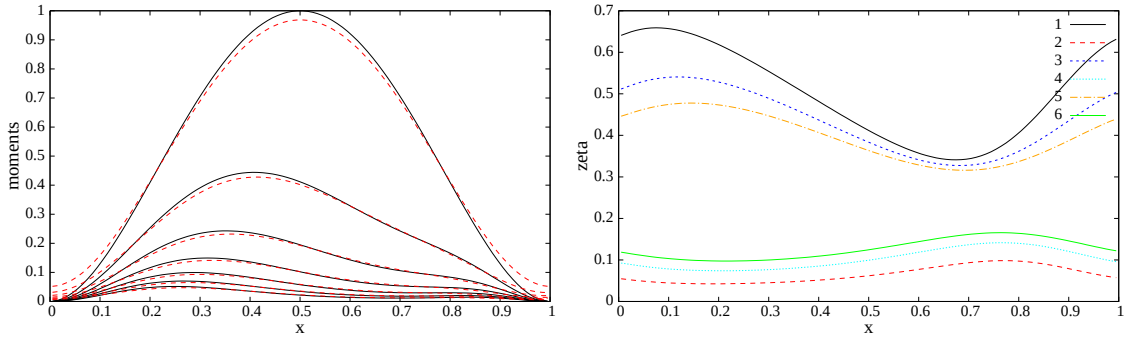

Figure 1: Constant fluid velocity test case with the regular initial solution; Right: space evolution of the moments of order 0 to 6 (top to bottom lines) for the exact initial solution (solid black lines) and for the simulation with the first order kinetic scheme at time $t = 2$ (red dashed lines). Left: initial values of the ζ_k, $k \in \{1, \ldots, 6\}$.

5.1.2. Oscillating initial ζ_k

An initial solution can be also defined from a choice of the ζ_k and of m_0. Here, oscillating functions are used for ζ_k, with a frequency increasing with k, whereas m_0 is a polynomial function:

$$\zeta_k^0(x) = \frac{x}{2}\left[1.01 + \cos\left(\frac{\pi k}{2}x\right)\right], \qquad m_0^0(x) = 16x^2(1-x)^2. \tag{41}$$

This allows to recover all the initial moments, which are in the interior of the moment space, but quite close to the boundary at some points.

5.1.3. Multi-modal initial NDF

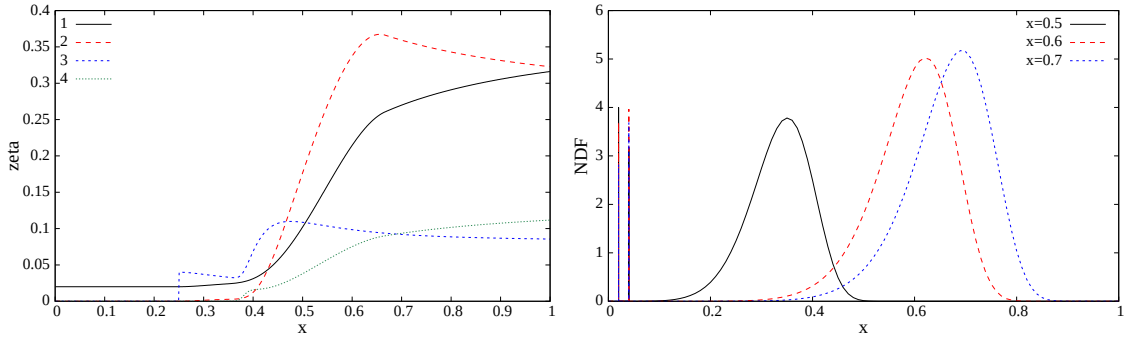

Figure 2: Left: ζ_k for k=1,2,3,4 as functions of the spatial location x. Right: initial distribution $f_0(\xi, x)$ as function of ξ for $x = 0.5$, $x = 0.6$ and $x = 0.7$ (for the Dirac delta functions, the absolute value of the heights of the scaled weights is arbitrary).

To test the ability of the methods to deal with the boundary of the moment space and the transition with the interior of this space, another initial distribution is introduced. It is detailed in Appendix G and the corresponding moments are regular (C^2) functions of x. Then, the distribution is only one Dirac delta function at $\xi = \xi_1 = 0.02$ for $x \in \left[0, \frac{1}{4}\right]$, in such a way that only ζ_1 is not zero, as seen in Fig. 2(left). It represents a distribution obtained through nucleation. For $x \in \left[\frac{1}{4}, \frac{1}{3}\right]$,

it is a sum of two Dirac delta functions at $\xi = \xi_1$ and $\xi = 2\xi_1$ and ζ_2 and ζ_3 are positive, with a discontinuity at $x = \frac{1}{3}$ for ζ_2. The second peak represents particles obtained by aggregation of two initial nuclei. A continuous Rosin-Rammler distribution is added for the rest of the domain, in such a way that the moments are then in the interior of the moment space. The obtained distribution is plotted in Fig. 2(right) for three different positions. Moreover, the corresponding moments are plotted in Fig. 7.

5.2. Results with a constant fluid velocity

To test the different schemes, let us first consider a case with a constant velocity. With no restriction, one assumes $u = 1$. Moreover, we consider the spatial domain $[0, 1]$ with periodic boundary conditions. The analytical solution at time $t = 2$ or $t = 5$ is then equal to the initial condition.

5.2.1. Numerical accuracy in a regular case

When considering the regular initial NDF, with moments far from the boundary of the moment space, simulations are done with the three kinetic schemes on uniform meshes with a CFL number equal to 0.8 and are compared with the analytical solution at the final time $t = 2$. First, using a 100 points spatial discretization, the moments of order 0 to 6 obtained at the final time with the first order scheme are plotted in Fig. 1(left), showing its numerical diffusion whereas the other kinetic schemes lead to very precise results (not shown here).

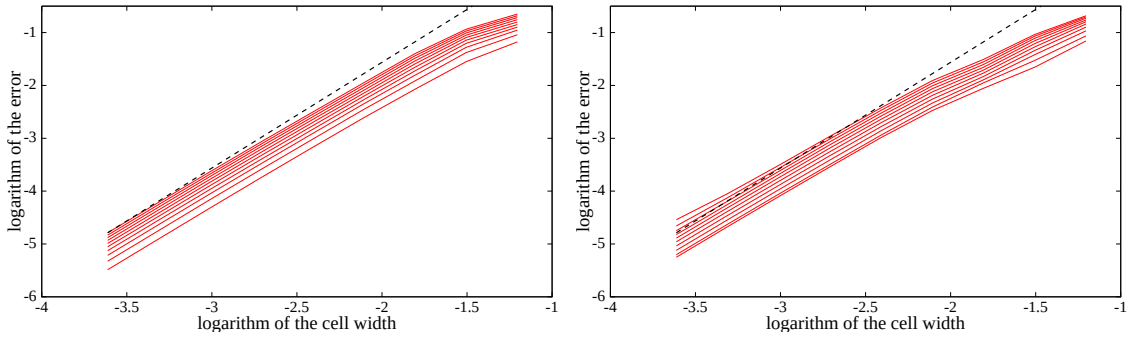

Figure 3: Constant fluid velocity test case with the regular initial NDF: L_1 norm of the error on the moments of order 0 to 9 (solid red lines, bottom to top) with the ζ kinetic scheme (left) and the QW kinetic scheme (right) for a CFL equal to 0.8. The same line of slope 2 is represented by black dashed lines on the two figures.

To evaluate more precisely the accuracy of the methods, the L_1 norm of the errors, divided by the L_1 norm of the corresponding moments, are plotted as a function of the cell width in Fig. 3, using from 16 to 4096 cells. The L_1 norm is used since we saw that a local loss of accuracy can happen near extrema of the reconstructed variables or near zones where the moments are at the boundary of the moment space. Both the methods are numerically almost second order accurate: the order is about 1.93 for all moments with the ζ kinetic scheme and 1.91 for the moments of order 0 to 5 with the QW kinetic scheme. Moreover, the ζ kinetic scheme does not introduce any dependance on the number of considered moments in this case where no corrections are needed on the slope to ensure the positivity of the $a_{k,j}$. On the contrary, for the QW kinetic scheme, the number of considered moments has an influence on the accuracy on all moments. It can be seen in Fig. 4, where the same normalized L_1 norm of the errors are plotted when using 6 and 5

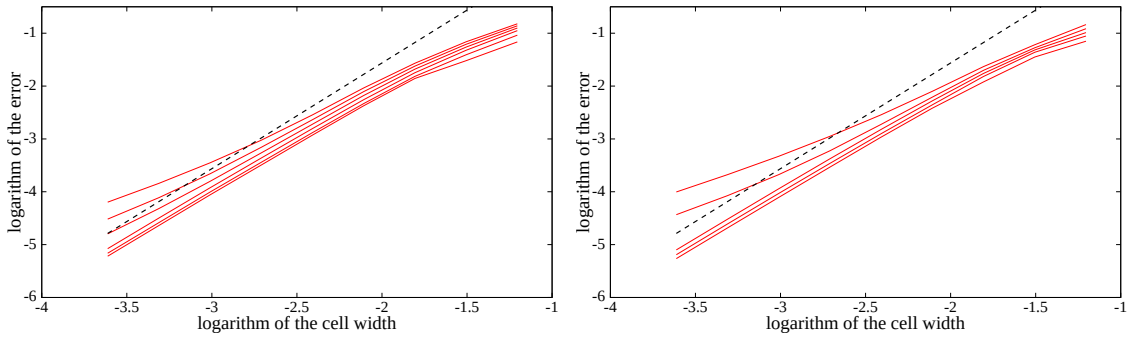

Figure 4: Constant fluid velocity test case with the regular initial solution: L_1 norm of the error on the moments of order 0 to N (solid red lines, bottom to top) with the QW kinetic scheme for a CFL equal to 0.8 and for $N = 5$ (left) and $N = 4$ (right). The same line of slope 2 is represented by black dashed lines on the two figures.

moments with this scheme. Moreover, as explained in Section 3.2.3, the second order of accuracy is only guarantied for half the first order moments and indeed, here, one can see that for half of the highest order moments, the order of accuracy degenerates to 1 when a fine enough mesh is used (this degenerescence did not yet really appear for the finest mesh used in the case with 10 moments). Let us also remark that, for this case, no correction is needed to ensure the positivity of the weights of the reconstruction (*i.e.* no elimination of weights for the neighbor cells), as expected since the moments are far from the boundary of the moment space.

When considering the simplified schemes, the CFL number is set to 0.3. The behavior of the schemes are similar (and not represented here): both the methods are also numerically almost second order accurate: the order is about 1.93 for all moments with the ζ simplified scheme and about 1.9 for the moments of order 0 to 6 with the QW simplified scheme. They are also compared

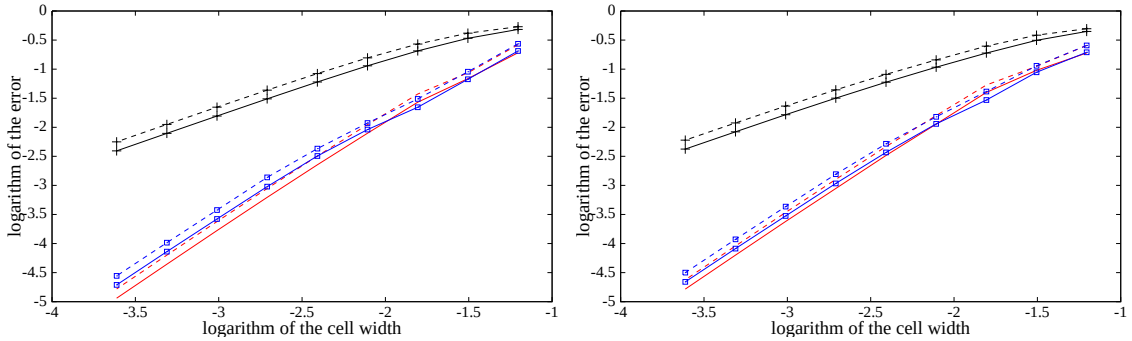

Figure 5: Constant fluid velocity test case with the regular initial NDF: L_1 norm of the error on the moments of order 0 (left) and 1 (right) with the kinetic (solid lines) of the simplified (dashed lines) schemes for a CFL equal to 0.3, using the constant (black +), the ζ (red) or the QW (blue □) reconstruction.

to the kinetic schemes, using the same CFL equal to 0.3. The errors on the moments of order 0 and 1 are then plotted in Fig. 5. One can first remark that the accuracy of the kinetic schemes is smaller than for a CFL equal to 0.8, with more than half an order of magnitude for the difference. It can be checked in this case that the results for the simplified schemes are very little sensitive to the CFL number. Moreover, the accuracy obtained with the non-constant reconstructions of the

moments are much higher than with the constant reconstruction, allowing to reduce drastically the number of cells needed for a given accuracy (about a factor 10 on the cell number for an error of 1%). This is an important point for more complex problems where a quite costly ODE system has to be solved on each cell for the resolution of the source terms. Moreover, the ζ reconstruction gives the best results for most of the discretizations.

5.2.2. Verification in extreme cases

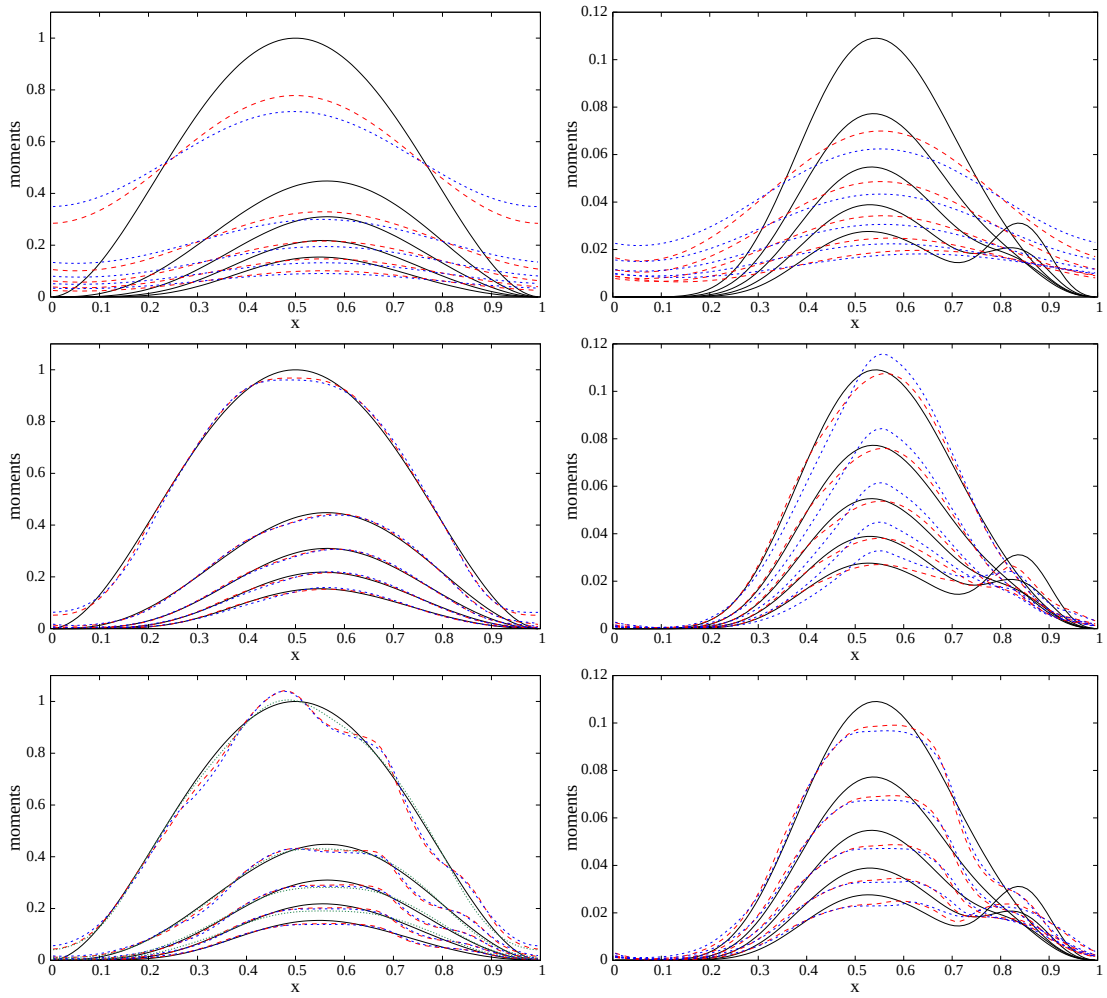

Figure 6: Constant fluid velocity test case with the oscillating initial ζ_k: space evolution of the moments of order 0 to 4 (left, top to bottom lines) and of order 5 to 8 (right, top to bottom lines for $x < 0.7$) at time $t = 5$ for the exact solution (solid black lines) and for the simulation with the constant reconstruction (top), with the ζ reconstruction (middle) and the QW reconstruction (bottom) for the kinetic scheme (red dashed line or green dots for QW with 4 moments) and the simplified scheme (blue dotted lines) with 100 points and a CFL equal to 0.3.

Let us consider more challenging test cases. The final time, $t = 5$, is larger here to amplify the phenomena. The oscillating initial ζ_k is first considered. One can see that the ζ schemes, using

100 points and a CFL number equal to 0.3, give accurate results on the moments of order 0 to 4 (see Fig. 6 middle). However, for the highest order moments, the maximum principle is no more respected and the error is larger than for the first order moments but still smaller than with the other reconstructions, at least for the kinetic scheme. With the QW schemes using 10 moments and the same discretization (Fig. 6 bottom), the results are a little bit less good: the maximum principle is no more respected for the first order moments with some kind of oscillations around the analytical solution. However, these oscillations do not blow up and the accuracy is still better than with the first order scheme, for which the numerical diffusion flatten a lot the results (Fig. 6 top). Moreover, the accuracy of the QW schemes for the first order moments is improved by the use of a smaller number of moments.

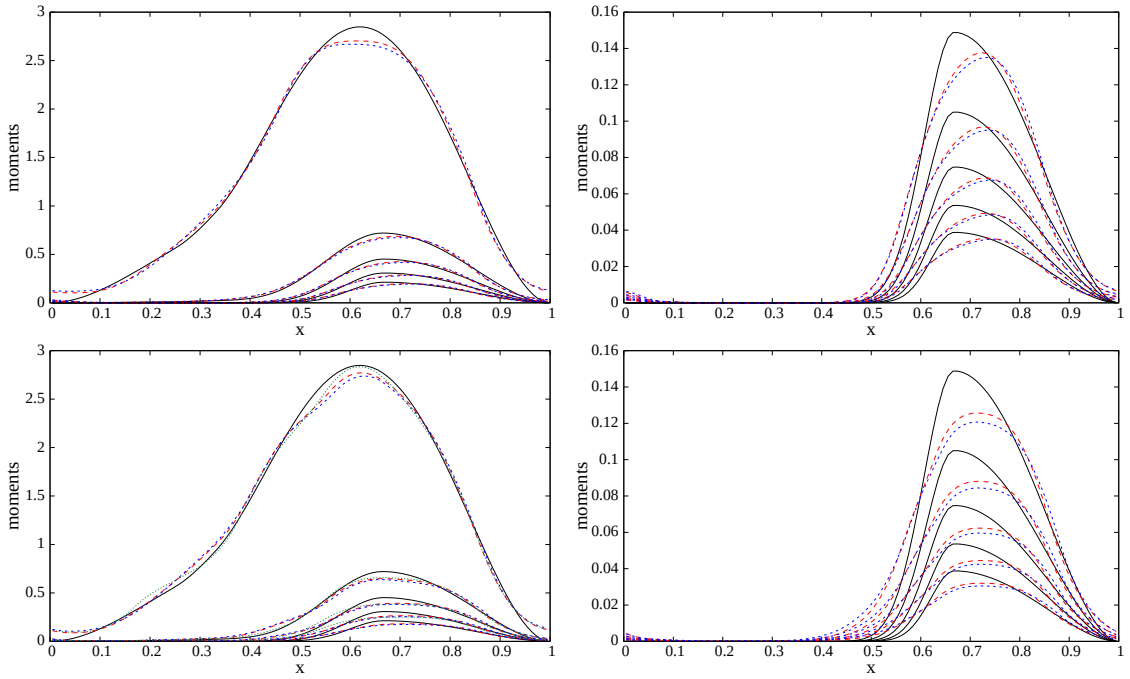

Figure 7: Constant fluid velocity test case with the multi-modal initial NDF: space evolution of the moments of order 0 to 4 (left, top to bottom lines) and of order 5 to 8 (right, top to bottom lines) at time $t = 5$ for the exact solution (solid black lines) and for the simulation with the ζ reconstruction (top) and the QW reconstruction (bottom) for the kinetic scheme (red dashed line or green dots for QW with 4 moments) and the simplified scheme (blue dotted lines) with 100 points and a CFL equal to 0.3.

Computations were done also with the multi-modal initial NDF, representing a more realistic NDF. The analytical solution as well as the solution computed with the ζ and QW schemes with 100 points and a CFL number equal to 0.3 are represented in Fig. 7. Both schemes do not encounter any problem simulating this test case, even if the moment vector is at the boundary of the moment space in some part of the domain. Moreover, in this more realistic case, the maximum principles are respected and the ζ reconstruction gives the most accurate results when using 10 moments, but the use of only 4 moments allow to improve the accuracy for the QW schemes. As it will be shown for the unsteady case, the difference of accuracy on the zeroth and first order moments is higher between the ζ and QW reconstructions when a finer discretization is used for simulations with 10

moments and a second order of accuracy is still recovered. Moreover, all methods are still always much more accurate than the schemes using a constant reconstruction.

5.3. Results with an unsteady and compressible fluid velocity

An unsteady fluid velocity u and the corresponding characteristics are given, for $x \in [0, 1]$ by:

$$u(t, x) = \frac{1 - x}{1 + t}, \qquad X(t; s, x) = 1 + \frac{x - 1}{1 + t - s}.$$

The analytical solution is given by:

$$\mathbf{m}_N(t, x) = (1 + t)\mathbf{m}_N^0 \left(1 + (x - 1)(1 + t)\right).$$

For the simulations, we use $N = 9$ and the final time is $t = 1$.

5.3.1. Numerical accuracy

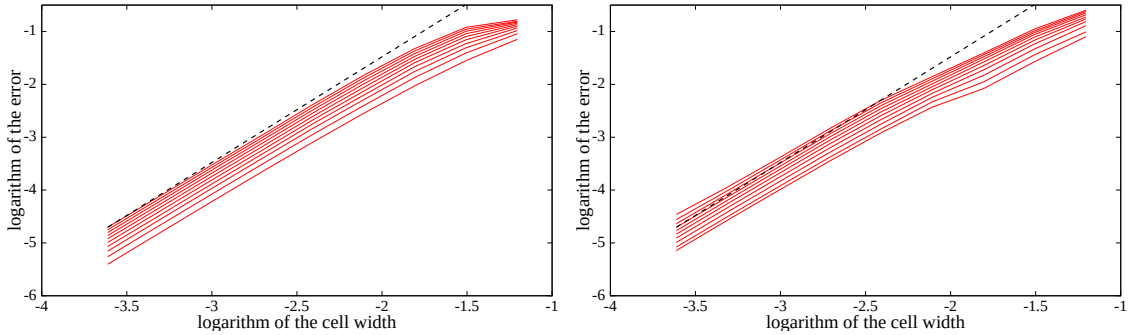

Figure 8: Unsteady fluid velocity test case with the regular initial NDF: L_1 norm of the error on the moments of order 0 to 9 (solid red lines, bottom to top) with the ζ kinetic scheme (left) and the QW kinetic scheme (right) for a CFL equal to 0.8. The same line of slope 2 is represented by black dashed lines on the two figures.

The same kind of simulations are done in the unsteady case as with the stationary one. For the ζ and QW kinetic schemes and the regular initial NDF, the errors on the moments are shown in Fig. 8. The numerical order of accuracy is about 1.93 when using the ζ reconstruction and about 1.91 for the moments of order 0 to 5, when using the QW reconstruction. Moreover, in this case, the influence of the CFL is much smaller than with the constant fluid velocity, probably due to the interpolation error of the fluid velocity in the case of the kinetic schemes. All methods then give similar results compared to the constant fluid velocity test case, the ζ reconstruction based schemes being still a little bit more accurate.

The difference between the two kinds of reconstructions is higher when considering the multi-modal initial NDF (see Fig. 9 left), this difference being reduced when considering only 4 moments with the QW reconstruction based schemes. In any cases, the accuracy stays much higher than with the constant reconstruction. In Fig. 9(right), it can be seen that the ζ and QW kinetic schemes well capture the mean value of the distribution (corresponding to ζ_1), as well as its variance (corresponding to $\zeta_1\zeta_2$) with only 100 cells. The corresponding plot for the simplified scheme is very similar and not shown here. Moreover, one can see in Fig. 10 that the second order of accuracy is still obtained in this case, for both the ζ and the QW kinetic schemes.

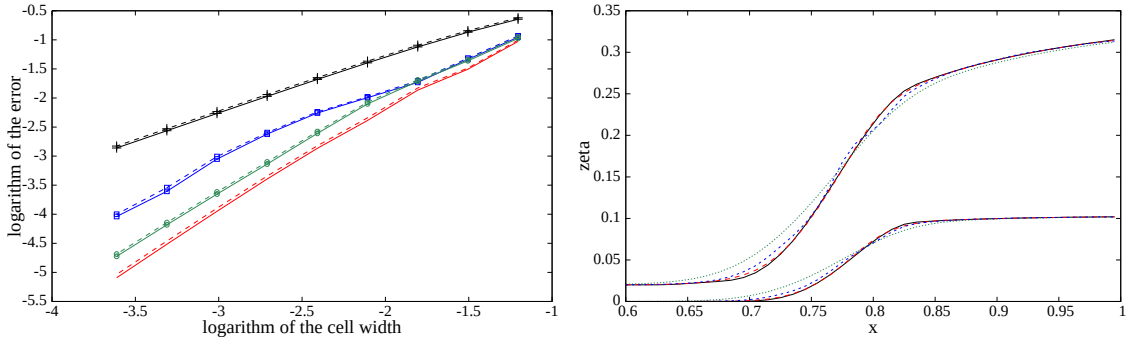

Figure 9: Unsteady fluid velocity test case with the multi-modal initial NDF. Left: L_1 norm of the error on the moments of order 0 with the kinetic (solid lines) of the simplified (dashed lines) schemes for a CFL equal to 0.3, using the constant (black +), the ζ (red) or the QW (blue $\square$ when $N = 9$ and green $\circ$ when $N = 3$) reconstruction. Right: space evolution of the mean value ζ_1 (top curves) and the variance $\zeta_1\zeta_2$ (bottom curves) of the NDF at time $t = 1$ for the exact solution (solid black lines) and for the simulation with the kinetic scheme and a CFL equal to 0.8, using the constant reconstruction (dashed red lines), the ζ reconstruction (blue dashed line) and the QW reconstruction (dotted green line), with 100 points.

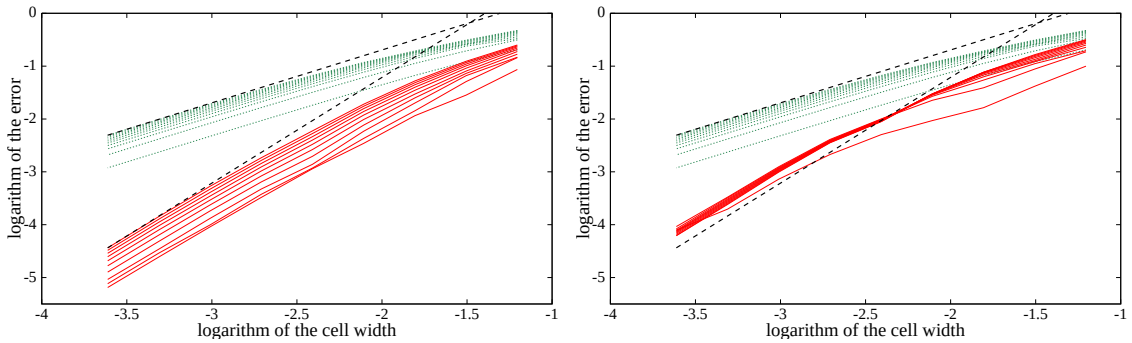

Figure 10: Unsteady fluid velocity test case with the multi-modal initial NDF: L_1 norm of the error on the moments of order 0 to 9 (solid red lines, bottom to top) with the ζ kinetic scheme (left) and the QW kinetic scheme (right) and with the first order scheme (both figures, dotted green lines) for a CFL equal to 0.8. The same lines of slope 1 and 2 are represented by black dashed lines on the two figures.

5.3.2. Computational time

In the case of the multi-modal initial NDF, the computational time of the different schemes with the use of several numbers of moments are given in Table 1, when using a CFL number equal to 0.3 for all methods. This case is chosen since it is representative of physical cases and induces a fair comparison of the methods: some corrections of the slopes are needed for the ζ schemes as well as some abscissa eliminations for a few points for the QW schemes.

It shows that the computational time for the high order schemes is about 2.5 to 3 times higher than for the corresponding first order scheme, for the same discretization, except for the simplified QW scheme and for the ζ kinetic scheme with $N \geq 8$ where it is about 4 to 5 times higher. Indeed, when considering the ζ kinetic scheme, Gauss-Legendre quadrature with $\left\lceil \frac{N}{2} \right\rceil + 1$ points are used to compute all the integrals, thus still increasing the cost when the number of moments increase. Moreover, for the simplified schemes, the computational times are larger due to the use of the second order SSP Runge-Kutta method, thus needing two reconstructions per time step. However,

N	3	4	5	6	7	8	9
1^{st} order kinetic scheme	1.0	1.1	1.1	1.2	1.3	1.3	1.3
ζ kinetic scheme	2.5	2.9	3.3	3.8	4.0	5.1	6.2
QW kinetic scheme	2.4	2.9	2.8	3.3	3.4	3.9	4.0
1^{st} order simplified scheme	1.4	1.5	1.5	1.6	1.7	1.7	1.8
ζ simplified scheme	3.3	3.4	3.5	3.8	3.9	4.2	4.4
QW simplified scheme	4.4	5.1	5.0	6.0	6.2	7.4	7.5

Table 1: Unsteady fluid velocity test case with the multi-modal initial NDF: normalized computational time for a CFL number equal to 0.3.

the ζ simplified scheme is still competitive, at least for the same CFL, due to the very simple reconstruction. But the CFL cannot be increased beyond 1/3, whereas the global computational time of the other methods can be decreased by an increasing of the CFL number. When considering a global problem, including source terms, the time step can however have to be limited by the coupling characteristic time between the operators [53, 29].

Finally, due to the possible reduction of the number of cells, the cost of the schemes developed here is lower than the one of the first order scheme for the same accuracy. Moreover, the number of degrees of freedom being then reduced, it implies a reduction of the global cost of the complete problem with source terms.

6. Results with 2D, steady and incompressible fluid velocity

Let us consider a 2D configuration of particles in a vortex. The unsteady fluid velocity is then defined for $x \in [0, 1/2]$ and $y \in [0, 1/2]$ by:

$$u_x(t, x, y) = \sin(2\pi x)\cos(2\pi y), \qquad u_y(t, x, y) = -\cos(2\pi x)\sin(2\pi y).$$

Let us introduce the distance r to the point $(1/8, 1/8)$. A population of particles is initially present in the vortex, in a disk defined by $r < 1/8$. Its distribution is given by $f_0(\xi, 1 - 8r)$, where f_0 corresponds to the multi-modal NDF defined by (G.1), with a weight $w_3(x)$ now being equal to 1 for $x \in [2/3, 1]$. This population is then transported by the fluid till the time $t = 0.8$. Due to the incompressibility of the fluid phase, the value of the moments are conserved along the characteristics. A reference solution is then computed by solving reverse characteristics from each position, using a high order ODE solver with time step adaptation. For example, the zeroth order moments obtained with this method is given in Fig. 11(top left).

Simulations are done for moments of order 0 to 9, using a 200×200 uniform discretization and a CFL number equal to 0.3. To solve the 2D problem, a dimensional splitting is used: a 1D scheme is used alternatively on one-dimensional problems in the x and y directions. To obtain a second order of accuracy with this splitting method but using the same CFL number for each operator, steps of length Δt are used on each problem, but alternating the order of these steps in alternate time steps [49].

Since the kinetic and the simplified schemes give very similar results, only the results with the kinetic scheme are presented here. The zeroth order moment is plotted in Fig. 11. One can see the good level of accuracy of the ζ and QW kinetic schemes, especially compared to the first order scheme.

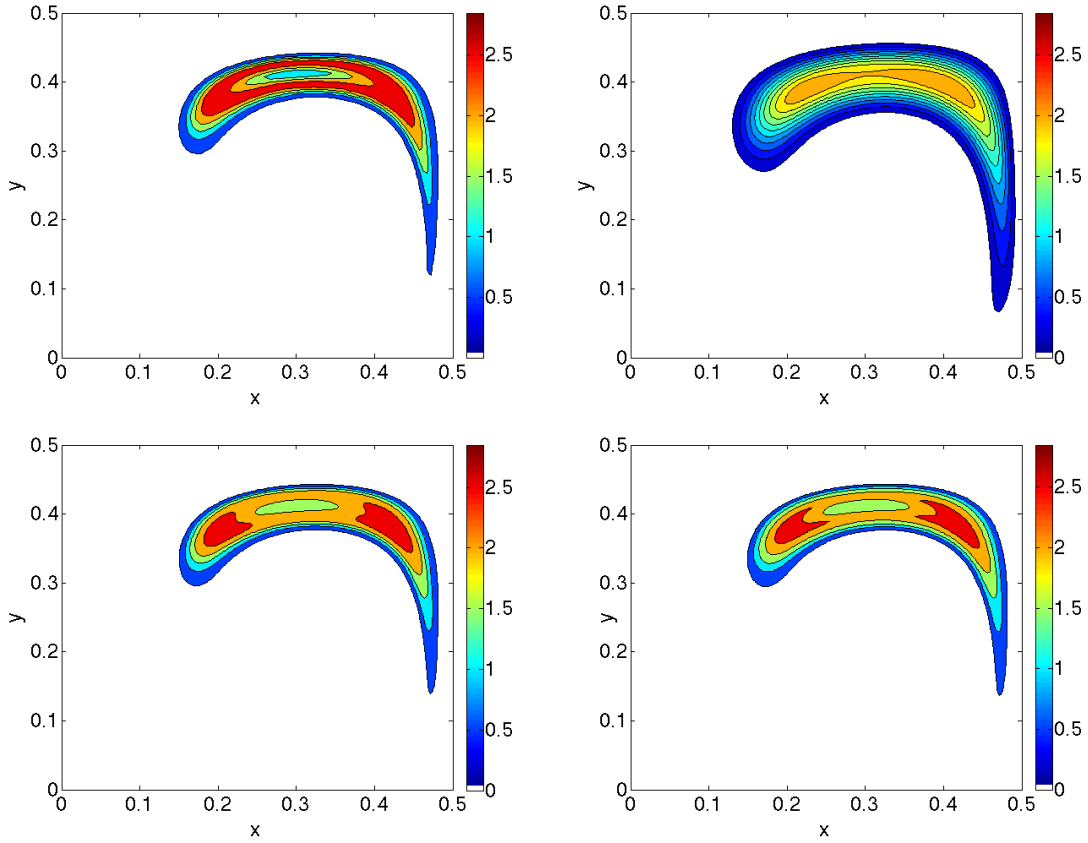

Figure 11: 2D test case: moment m_0 for the reference solution (top left), for the simulations on a 200×200 uniform mesh with a CFL number equal to 0.3, using the first order kinetic scheme (top right), the ζ kinetic scheme (bottom left) and the QW kinetic scheme (bottom right).

For going further in the comparisons, the mean value of the distribution, *i.e.* m_1/m_0 and its variance, *i.e.* $\zeta_1\zeta_2$ are plotted in Fig. 12 as functions of x at $y = 0.4$. The high numerical diffusion of the first order scheme is still remarkable, whereas the ζ and QW kinetic schemes allow to capture accurately these quantities.

Finally, for the kinetic schemes, the L_1 norm of the errors, divided by the L_1 norm of the corresponding moments, are plotted as a function of the cell width in Fig. 13, using from 100 to 1600 cells per direction. This shows the convergence of the methods and their high accuracy compared to the first order scheme, even if the second order of accuracy is not attained for the considered meshes, the solution having high gradients.

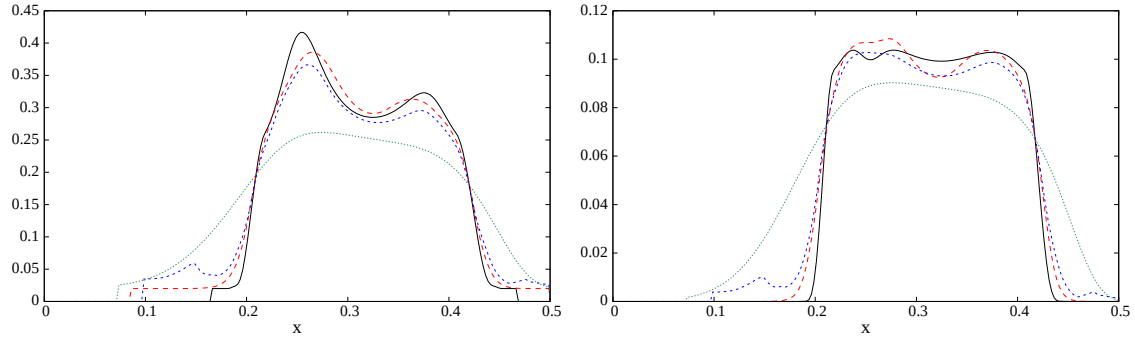

Figure 12: 2D test case: mean value (left) and variance (right) of the distribution as a function of x at $y = 0.4$ for the reference solution (black solid line), for the simulations on a 200×200 uniform mesh and a CFL number equal to 0.3, with the first order kinetic scheme (green dots), with the ζ kinetic scheme (red dashed line) and with the QW kinetic scheme (blue dotted line)

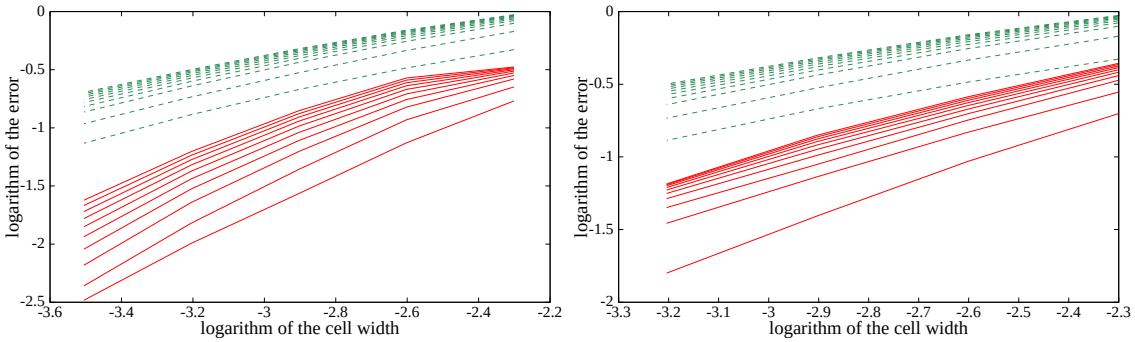

Figure 13: 2D test case: L_1 norm of the error on the moments of order 0 to 9 (solid red lines, bottom to top) with the ζ kinetic scheme (left) and the QW kinetic scheme (right) and with the first order scheme (both figures, dashed green lines) for a CFL equal to 0.8.

7. Conclusion

In this paper, we have provided four realizable accurate finite volume schemes for the transport of moments by a given velocity field, in Cartesian mesh context: for the ones, the flux computation is based on a follow-up of the characteristics (kinetic schemes) and for the other, it is based on the value of the moments at the interface (simplified schemes). In any case, a spatial reconstruction of the moments is needed and done by reconstructing variables that only have to be non-negative: either the corresponding ζ_k variables or the weights of the corresponding quadrature. Unlike previous developed realizable Eulerian schemes, they are able to deal with moment vectors of all sizes, at least till 10 moments, possibly at the boundary of the moment space, which can occur in practical applications. The verifications have been done on various test cases, 1D and 2D, with steady or unsteady and compressible or incompressible fluid velocities. The developed schemes then showed their high accuracy, compared to the first order scheme, and their second order of accuracy for all moments in the case of the ζ reconstruction and for half the lowest order ones for the QW reconstruction. The kinetic schemes are more accurate than the simplified schemes, especially due to the fact that a higher CFL number can be used. But their generalization to unstructured meshes

seems less easy than for the simplified schemes. Moreover, the ζ reconstruction leads to slightly better accuracy than the QW one in most cases. And the accuracy on the QW schemes is much more sensitive to the number of considered moments and can eventually decrease when this number increases. Finally, for a given accuracy, the cost of all the developed schemes is usually lower than the one of the first order scheme. And since they all allow a high reduction of the number of degrees of freedom, the global cost of the complete problem with source terms can be drastically reduced by using these schemes, compared to the first order ones, thus showing the great interest of such kind of schemes. Moreover, we are studying an implementation of a simplified scheme in the open-source computational fluid dynamics toolbox OpenFOAM as part of the Open-QBMM project [54, 55].

Acknowledgments

This work was supported by a grant from French National Research Agency: ASMAPE project (ANR-13-TDMO-02 ASMAPE, PI IFPen: O. Colin; PI for EM2C Lab.: F. Laurent).

The authors would like also to thank Professor R. O. Fox for several helpful discussions and for the precious feedback he provided about this paper.

Appendix A. Case of a compact support

Let us consider here the case where the support of the NDF is included in a known compact interval of the form $[0, \xi_{\max}]$. The moments are then denoted m_k and and the corresponding "dimensionless" moments are $\tilde{m}_k = \dfrac{m_k}{\xi_{\max}^{k+1}}$. They correspond to moments on the support $[0, 1]$ and are obtained through the change of variables $\xi \mapsto \dfrac{\xi}{\xi_{\max}}$. Then, a second kind of Hankel determinant has to be defined:

$$\overline{H}_{2n+d} = \begin{vmatrix} \tilde{m}_{1-d} - \tilde{m}_{2-d} & \dots & \tilde{m}_n - \tilde{m}_{n+1} \\ \vdots & \vdots & \vdots \\ \tilde{m}_n - \tilde{m}_{n+1} & \dots & \tilde{m}_{2n-1+d} - \tilde{m}_{2n+d} \end{vmatrix},$$

with $d = 0, 1$; $n \geq 0$. Similarly to the case with support in $[0, +\infty)$, one has the following characterization of the moment space [24]: the vector $\mathbf{m}_N = (m_1, \dots, m_N)^t$ is realizable (*i.e.* in the moment space) if and only if

$$\underline{H}_k \geq 0 \text{ and } \overline{H}_k \geq 0, \quad k \in \{0, 1, \dots, N\}.$$

Moreover, it is strictly realizable (*i.e.* in the interior of the moment space) if and only if these Hankel determinants are positive.

From the ζ_k corresponding to the moments m_k (then the $\zeta_k/\xi_{\max}$ correspond to the $\tilde{m}_k$), one can defined the canonical moments p_k by $\zeta_k = \xi_{\max} p_k (1 - p_{k-1})$. Their geometrical interpretation can be given from the m_k^- and m_k^+, which are now in $[0, m_0 \xi_{\max}^k]$ [24]:

$$p_k = \frac{m_k - m_k^-(\mathbf{m}_{k-1})}{m_k^+(\mathbf{m}_{k-1}) - m_k^-(\mathbf{m}_{k-1})}. \tag{A.1}$$

If $m_k(\mu)$ is equal to $m_k^+(\mathbf{m}_{k-1})$ or $m_k^-(\mathbf{m}_{k-1})$ for $\mu \in \mathcal{P}(\mathbf{m}_{k-1})$, (*i.e.* p_k is 0 or 1), the measure μ is a sum of weighted Dirac distributions and $\mathbf{m}_k$ belongs to the boundary of the moment space.

In this case of compact support, the QW schemes can be applied without any change. However, the ζ-schemes have to be adapted: the reconstruction has then to be done on the canonical moments to ensure the realizability for the corresponding support. In the case of the reconstruction for the kinetic scheme, the reconstructed ζ_k used for the flux computation in (31) are then $\zeta_k^n(x) = \xi_{\max} p_k^n(x)(1-p_{k-1}^n(x))$ where $p_k^n(x) = \bar{p}_{k,j}^n + D_{k,j}^n(x-x_j)$ on the j^{th} cell and $\bar{p}_{k,j}^n = a_{k,j} + b_{k,j} D_{k,j}^n$. The coefficients $a_{k,j}$ and $b_{k,j}$ are now given by:

$$a_{k,j}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_0^n(x)\xi_{\max}^k \prod_{i=1}^{k-1} p_i^n(x)(1-p_i^n(x))\mathrm{d}x = \Delta x\, m_{k,j}^n - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_0^n(x) P_k(\zeta_1^n(x),\dots,\zeta_{k-1}^n(x))\mathrm{d}x,$$

$$b_{k,j}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} m_0^n(x)\xi_{\max}^k \prod_{i=1}^{k-1} p_i^n(x)(1-p_i^n(x))\mathrm{d}x = -\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (x-x_j) m_0^n(x)\xi_{\max}^k \prod_{i=1}^{k-1} p_i^n(x)(1-p_i^n(x))\mathrm{d}x.$$

The same kind of slope limiters and corrections are used as fo the ζ kinetic scheme. For the simplified scheme, the reconstructed ζ_k are $\zeta_{k,j}^{\pm} = \xi_{\max} p_{k,j}^{\pm}(1-p_{k-1,j}^{\pm})$ with $p_{k,j}^{\pm} = p_{k,j}^n \pm \frac{D_{k,j}^n}{2}$, the slope $D_{k,j}^n$ being obtained through a minmod limitation from the $p_{k,i}^n$. The values of the fluxes are then given by (39). Let us remark that the corresponding schemes can be used in the more general case of a support in $[0,+\infty)$ by using for $\xi_{\max}$ the largest value of the abscissas obtained by a quadrature on each moment set at time t^n. Moreover, this value can only decrease with the transport schemes and the moments are then always bounded.

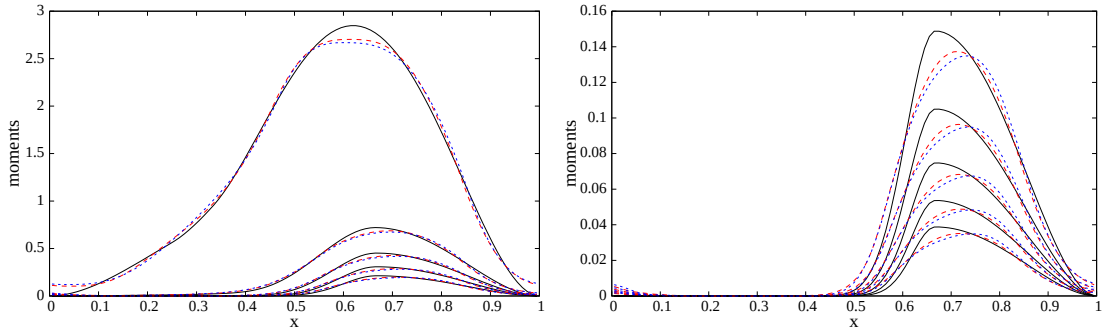

Figure A.14: Constant fluid velocity test case with the multi-modal initial NDF: space evolution of the moments of order 0 to 4 (left, top to bottom lines) and of order 5 to 8 (right, top to bottom lines) at time $t = 5$ for the exact solution (solid black lines) and for the simulation with the canonical moments reconstruction for the kinetic scheme (red dashed lines) and the simplified scheme (blue dotted lines) with 100 points and a CFL equal to 0.3.

These schemes are used in the constant fluid velocity test case with the multi-modal initial NDF. The 10 moments obtained by the simulation with 100 cells are plotted in Fig. A.14. It shows a similar accuracy compared to the ζ schemes.

Appendix B. Examples of realizability constraints for moments on $[0,+\infty)$

The first constraints (6) for the strict realizability of a moment vector $(m_0, m_1, \dots)^t$ can be written, making appear the values of $m_k^-(\mathbf{m}_{k-1})$: $m_0 > 0$, $m_1 > 0$,

$$m_2 > \frac{m_1^2}{m_0}, \qquad m_3 > \frac{m_2^2}{m_1}, \qquad m_4 > \frac{m_0 m_3^2 - 2m_1 m_2 m_3 + m_2^3}{m_2 m_0 - m_1^2}, \dots$$

Appendix C. Moments as functions of the ζ_k

The moments can be written from the ζ_k: $m_k = m_0 \left[P_k(\zeta_1, \dots, \zeta_{k-1}) + \prod_{j=1}^{k} \zeta_j \right]$, as shown in Corollary 2.3. The first polynomial functions P_k are given by:

$$\begin{aligned} P_1 =& 0, \\ P_2(\zeta_1) =& \zeta_1^2, \\ P_k(\zeta_1, \dots, \zeta_{k-1}) =& \zeta_1 \left[(\zeta_1 + \zeta_2)^{k-1} + \zeta_2 \zeta_3 Q_k(\zeta_1, \dots, \zeta_{k-1}) \right], \quad k \geq 3, \end{aligned}$$

with

$$\begin{aligned} Q_3(\zeta_1, \zeta_2) =& 0, \\ Q_4(\zeta_1, \zeta_2, \zeta_3) =& 2(\zeta_1 + \zeta_2) + \zeta_3, \\ Q_5(\zeta_1, \dots, \zeta_4) =& 3(\zeta_1 + \zeta_2)^2 + 2(\zeta_1 + \zeta_2)(\zeta_3 + \zeta_4) + (\zeta_3 + \zeta_4)^2 + \zeta_2 \zeta_3, \\ Q_6(\zeta_1, \dots, \zeta_5) =& 4(\zeta_1 + \zeta_2)^3 + 2(\zeta_1 + \zeta_2)(\zeta_3 + \zeta_4)^2 + (\zeta_3 + \zeta_4)^3 + 2\zeta_2 \zeta_3 (\zeta_3 + \zeta_4) \\ & + 2\zeta_4 \zeta_5 (\zeta_1 + \zeta_2 + \zeta_3 + \zeta_4) + \zeta_4 \zeta_5^2 + 3\zeta_3 (\zeta_1 + \zeta_2)(\zeta_1 + 2\zeta_2) \\ & + 3\zeta_4 (\zeta_1 + \zeta_2)^2, \\ Q_7(\zeta_1, \dots, \zeta_6) =& 5(\zeta_1 + \zeta_2)^4 + 2(\zeta_1 + \zeta_2)(\zeta_3 + \zeta_4)^3 + (\zeta_3 + \zeta_4)^4 \\ & + 3\zeta_2 \zeta_3 (\zeta_3 + \zeta_4)^2 + 2\zeta_4 \zeta_5 \zeta_6 (\zeta_1 + \zeta_2 + \zeta_3 + \zeta_4 + \zeta_5) + \zeta_4 \zeta_5 \zeta_6^2 \\ & + 6\zeta_3 \zeta_4 (\zeta_1 + \zeta_2)(\zeta_1 + 2\zeta_2) + 3\zeta_4 \zeta_5 (\zeta_1 + \zeta_2)^2 + 4(\zeta_1 + \zeta_2)^3 \zeta_4 \\ & + 2(\zeta_1 + \zeta_2)^2 (2\zeta_1 + 5\zeta_2)\zeta_3 + 3\zeta_4 \zeta_5 (\zeta_3 + \zeta_4)^2 \\ & + 4(\zeta_1 + \zeta_2)(\zeta_3 + \zeta_4)\zeta_4 \zeta_5 + \zeta_3^2 (10\zeta_2^2 + 12\zeta_1 \zeta_2 + 3\zeta_1^2) \\ & + 3(\zeta_1 + \zeta_2)^2 \zeta_4^2 + 2\zeta_4 \zeta_5^2 (\zeta_1 + \zeta_2 + \zeta_3) + \zeta_4 \zeta_5^3 + 3\zeta_4^2 \zeta_5^2 + 2\zeta_2 \zeta_3 \zeta_4 \zeta_5. \end{aligned}$$

Appendix D. Property [P3] for classical MUSCL reconstruction

Let us consider just one moment, denoted m, without any indice here. Let us denote $\overline{m}(x)$ the exact solution of the transport equation at a given time and $m_j = \frac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \overline{m}(x) \mathrm{d}x$ its averaged value in the cell j, for any j. Let us then consider the affine reconstruction $m(x)$ from the m_j, corresponding to a MUSCL scheme:

$$m(x) = m_j + (x - x_j) D_j; \qquad x \in (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}).$$

The slope D_j is then an approximation of the derivative of $\overline{m}$, using a limiter [56]:

$$D_j = \frac{m_{j+1} - m_j}{\Delta x} \Phi(\theta_j), \quad \theta_j = \frac{m_j - m_{j-1}}{m_{j+1} - m_j},$$

where the function Φ defines the flux limiter, such that $\Phi(1) = 1$ and Φ is bounded and Lipschitz continuous at $\theta = 1$. Using Taylor expansions of $\overline{m}$ around x_j, for example

$$m_j = \overline{m}_j + \frac{\Delta x^2}{24} \overline{m}_j'' + O(\Delta x^3),$$

with $\overline{m}_j = \overline{m}(x_j)$ and $\overline{m}''_j = \overline{m}''(x_j)$, it is easy to see that except near extreme points of $\overline{m}$: $\theta_j = 1 + O(\Delta x)$ and $\theta_{j+1} - \theta_j = O(\Delta x^2)$ and

$$m(x) - \overline{m}(x) = O(\Delta x^2) \quad \text{for} \quad x \in (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}).$$

Without loss of generality, let us assume that $0 \leq \delta < \Delta x$ and let us denote

$$I_{j-\frac{1}{2}} = \frac{1}{\Delta x^2} \int_{x_{j-\frac{1}{2}}}^{x_{j-\frac{1}{2}}+\delta} (m(x) - \overline{m}(x)) \mathrm{d}x = \frac{1}{\Delta x^2} \left[\delta m_j + D_j \frac{\delta}{2}(\delta - \Delta x) - \int_{x_{j-\frac{1}{2}}}^{x_{j-\frac{1}{2}}+\delta} \overline{m}(x) \mathrm{d}x \right].$$

We then have to see if $I_{j-\frac{1}{2}} - I_{j-\frac{1}{2}} = O(\Delta x^2)$. First, using Taylor expansion:

$$I_{j-\frac{1}{2}} = \frac{\delta(\Delta x - \delta)}{2\Delta x^2} \left\{ \overline{m}'_j \left[\Phi(\theta_j) - 1 \right] + \frac{\Delta x}{2} \overline{m}''_j \Phi(\theta_j) \right\} + \overline{m}''_j \left[\frac{\delta}{24} - \frac{\left(-\frac{\Delta x}{2} + \delta\right)^3 + \left(\frac{\Delta x}{2}\right)^3}{6\Delta x^2} \right] + O(\Delta x^2),$$

in such a way that

$$I_{j+\frac{1}{2}} - I_{j-\frac{1}{2}} = \frac{\delta(\Delta x - \delta)}{2\Delta x^2} \left\{ \overline{m}'_j \left[\Phi(\theta_{j+1}) - \Phi(\theta_j) \right] + \frac{\Delta x}{2} \overline{m}''_j \left[\Phi(\theta_{j+1}) - 1 + \Phi(\theta_{j+1}) - \Phi(\theta_j) \right] \right\} + O(\Delta x^2).$$

Then, except near extreme points of $\overline{m}$, the property [P3] is verified with this kind of reconstruction. Moreover, let us remark that the use of the minmod limiter allows the reconstruction to have a smaller total variation than $\overline{m}$.

Appendix E. Algorithms

The reconstruction used for the ζ kinetic scheme is detailed in Algorithm 1, especially the corrections done in the case $a_{k,j} < 0$. More precisely, this algorithm computes the parameters of the reconstruction in the cell j ($\bar{\zeta}^n_{k,j}$ for $k \in \{1, \ldots, N\}$ and the slopes $D^n_{k,j}$ for $k \in \{0, \ldots, N\}$), providing the zeroth order moments and the values of the ζ_k corresponding to $\mathbf{m}^n_j$, $\mathbf{m}^n_{j-1}$ and $\mathbf{m}^n_{j+1}$. For that, the formula (27) leading to the computation of $a_{k,j}$ from the reconstruction of m_0 and the ζ_i for $i < k$ is denoted:

$$a_{k,j} = \Psi(m^n_{0,j}, D^n_{0,j}, a_{1,j}, b_{1,j}, D^n_{1,j}, \ldots, a_{k-1,j}, b_{k-1,j}, D^n_{k-1,j}).$$

Algorithm 1: Reconstruction in the cell j for the ζ kinetic scheme

Data: $m^n_{0,j}, m^n_{0,j-1}, m^n_{0,j+1}, \left(\zeta^n_{k,j}, \zeta^n_{k,j-1}, \zeta^n_{k,j+1}\right)_{k\in\{1,\dots,N\}}$

Result: Reconstruction parameters: $\left(\bar{\zeta}^n_{k,j}\right)_{k\in\{1,\dots,N\}}$ and $\left(D^n_{k,j}\right)_{k\in\{0,\dots,N\}}$

Compute $D_{0,j}$ from (29);
for $k \leftarrow 1$ **to** N **do**
  Compute $a_{k,j}$, $b_{k,j}$ from (27,28);
  Compute $D_{k,j}$ from (30) ;
  if $a_{k,j} < 0$ **then**
    $k_0 \leftarrow \max\left\{i, \Psi\left(m^n_{0,j}, D^n_{0,j}, \dots, a_{i-1,j}, b_{i-1,j}, D^n_{i-1,j}, a_{i,j}, b_{i,j}, 0, \dots, a_{k-1,j}, b_{k-1,j}, 0\right) > 0\right\}$;
    $n_c \leftarrow 0$;
    while $a_{k,j} < 0$ **do**
      if $n_c < 5$ **then**
        $n_c \leftarrow n_c + 1$;
        $D^n_{k_0,j} \leftarrow 0.9\, D^n_{k_0,j}$;
        $\bar{\zeta}^n_{k_0,j} \leftarrow a_{k_0,j} + b_{k_0,j} D^n_{k_0,j}$;
        Compute $a_{p,j}$, $b_{p,j}$ and $D^n_{p,j}$ from (27,28) and (30) for $p = k_0 + 1, \dots, k$;
      else
        $n_c \leftarrow 0$;
        $D^n_{k_0,j} \leftarrow 0$;
        $\bar{\zeta}^n_{k_0,j} \leftarrow a_{k_0,j}$;
        $k_0 \leftarrow k_0 + 1$;
  $\bar{\zeta}^n_{k,j} \leftarrow a_{k,j} + b_{k,j} D^n_{k,j}$;

The reconstruction used for the ζ simplified scheme is detailed in Algorithm 2, especially the corrections done in the case where $\mathbf{m}^*_j$ is not in the interior of the same moment space as $\mathbf{m}^n_j$, *i.e.* where $\mathcal{N}_\epsilon(\mathbf{m}^*_j) < \mathcal{N}_\epsilon(\mathbf{m}^n_j)$.

Algorithm 2: Reconstruction in the cell j for the ζ simplified scheme

Data: $m^n_{0,j}, m^n_{0,j-1}, m^n_{0,j+1}, \left(\zeta^n_{k,j}, \zeta^n_{k,j-1}, \zeta^n_{k,j+1}\right)_{k\in\{1,\dots,N\}}$ and $\mathcal{N}_\epsilon(\mathbf{m}^n_j)$

Result: $\mathbf{m}^-_j$ and $\mathbf{m}^+_j$

Compute the slopes $\left(D^n_{k,j}\right)_{k\in\{0,\dots,\mathcal{N}_\epsilon(\mathbf{m}^n_j)-1\}}$ from (36,37), adding the limitation (38) if $k=1$;

Compute $m^\pm_{0,j}$ and $\left(\zeta^\pm_{k,j}\right)_{k\in\{1,\dots,\mathcal{N}_\epsilon(\mathbf{m}^n_j)-1\}}$ from (35);

Set $\left(D^n_{k,j}\right)_{k\in\{\mathcal{N}_\epsilon(\mathbf{m}^n_j),\dots,N\}}$ and $\left(\zeta^\pm_{k,j}\right)_{k\in\{\mathcal{N}_\epsilon(\mathbf{m}^n_j),\dots,N\}}$ to zero;

Compute the moments $\mathbf{m}^\pm_j$ from $m^\pm_{0,j}$ and $(\zeta^\pm_{k,j})_{k\in\{1,\dots,N\}}$ (reverse algorithm);

$\mathbf{m}^*_j \leftarrow 3\mathbf{m}^n_j - \mathbf{m}^+_j - \mathbf{m}^-_j$;

Compute $\mathcal{N}_\epsilon(\mathbf{m}^*_j)$;

if $\mathcal{N}_\epsilon(\mathbf{m}^*_j) < \mathcal{N}_\epsilon(\mathbf{m}^n_j)$ **then**

  for $p \leftarrow 2$ **to** $\mathcal{N}_\epsilon(\mathbf{m}^n_j)$ **do**

    Compute the moments $\mathbf{m}^\pm_j$ from $m^\pm_{0,j}$, $(\zeta^\pm_{k,j})_{k\in\{1,\dots,p\}}$ and$(\zeta^n_{k,j})_{k\in\{p+1,\dots,N\}}$;

    $\mathbf{m}^*_j \leftarrow 3\mathbf{m}^n_j - \mathbf{m}^+_j - \mathbf{m}^-_j$;

    Compute $\mathcal{N}_\epsilon(\mathbf{m}^*_j)$;

    if $\mathcal{N}_\epsilon(\mathbf{m}^*_j) < \mathcal{N}_\epsilon(\mathbf{m}^n_j)$ **then**

      $D^n_{p,j} \leftarrow 0.5\, D^n_{p,j}$;

      Compute $\zeta^\pm_{p,j}$ from (35);

      Compute the moments $\mathbf{m}^\pm_j$ from $m^\pm_{0,j}$, $(\zeta^\pm_{k,j})_{k\in\{1,\dots,p\}}$ and$(\zeta^n_{k,j})_{k\in\{p+1,\dots,N\}}$;

      $\mathbf{m}^*_j \leftarrow 3\mathbf{m}^n_j - \mathbf{m}^+_j - \mathbf{m}^-_j$;

      Compute $\mathcal{N}_\epsilon(\mathbf{m}^*_j)$;

    if $\mathcal{N}_\epsilon(\mathbf{m}^*_j) < \mathcal{N}_\epsilon(\mathbf{m}^n_j)$ **then**

      $D^n_{p,j} \leftarrow 0$;

      $\zeta^\pm_{p,j} \leftarrow \zeta^n_{p,j}$;

The reconstruction used for the QW schemes is detailed in Algorithm 3. For that, let us denote $V\left((\xi_\alpha)_{\alpha\in\{1,\dots,n\}}\right)$ the Vandermonde matrix corresponding to the coefficients $(\xi_1,\dots,\xi_n)$, *i.e.* the matrix of coefficients ξ_j^{i-1} on the row $i\in\{1,\dots,n\}$ and column $j\in\{1,\dots,n\}$.

Algorithm 3: Reconstruction in the cell j for the QW schemes

Data: $\mathbf{m}_j^n$, $\mathbf{m}_{j-1}^n$, $\mathbf{m}_{j-1}^n$

Result: $\left(w_{\alpha,j}, \xi_{\alpha,j}, D_{\alpha,j}^n\right)_{\alpha\in\{1,\dots,\lfloor\frac{N}{2}\rfloor+1\}}$

$p \leftarrow \left\lfloor \frac{\mathcal{N}_\epsilon(\mathbf{m}_j^n)}{2} \right\rfloor + 1$;

Compute $(w_{\alpha,j}, \xi_{\alpha,j})_{\alpha\in\{1,\dots,p\}}$ from $(m_{k,j}^n)_{k\in\{0,\dots,\mathcal{N}_\epsilon(\mathbf{m}_j^n)-1\}}$ (Chebychev algorithm);

$\mathcal{C} = \{1,\dots,p\}$;

$\left(w_{\alpha,j}^{\pm}\right)_{\alpha\in\mathcal{C}}^t \leftarrow V\left((\xi_\alpha)_{\alpha\in\mathcal{C}}\right)^{-1}\left(m_{0,j\pm1}^n,\dots,m_{p-1,j\pm1}^n\right)^t$;

while *any* $\left(w_{\alpha,j}^{\pm}\right)_{\alpha\in\mathcal{C}} < 0$ **do**

  $p \leftarrow p-1$;

  Determine α_0 such that $w_{\alpha_0,j}^{\pm} = \min\{w_{\alpha,j}^{\pm}, \alpha\in\mathcal{C}\}$;

  $w_{\alpha_0,j}^{\pm} \leftarrow 0$;

  $\mathcal{C} = \mathcal{C} - \{\alpha_0\}$;

  $\left(w_{\alpha,j}^{\pm}\right)_{\alpha\in\mathcal{C}}^t \leftarrow V\left((\xi_\alpha)_{\alpha\in\mathcal{C}}\right)^{-1}\left(m_{0,j\pm1}^n,\dots,m_{p-1,j\pm1}^n\right)^t$;

Appendix F. Reconstruction at the level of the NDF

Another way to reconstruct the moments $\mathbf{m}^n(x)$ is to consider a spatial reconstruction at the NDF level [43, 44]. Let us then reconstruct a NDF f_j^n from the moment set $(m_{k,j}^n)_{k\in\{0,\dots,N\}}$, using EQMOM or entropy maximization and let us define

$$f^n(x,\xi) = f_j^n(\xi) + D_j^n(\xi)(x-x_j), \quad x\in[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}].$$

The slope is defined thanks to a minmod limiter:

$$D_j^n(\xi) = \frac{1}{2}\left(sgn(f_{j+1}^n(\xi)-f_j^n(\xi)) + sgn(f_j^n(\xi)-f_{j-1}^n(\xi)\right)\min\left(\frac{|f_{j+1}^n(\xi)-f_j^n(\xi)|}{\Delta x}, \frac{|f_j^n(\xi)-f_{j-1}^n(\xi)|}{\Delta x}\right)$$

and the reconstructed moments for the kinetic scheme are then $m_k^n(x) = \int_0^{+\infty}\xi^k f^n(x,\xi)\mathrm{d}\xi$. Practically, a quadrature on the ξ variable is used to compute the fluxes. But here, to ensure the [P1] property, this quadrature has to correspond to the measure $f_j(\xi)\mathrm{d}\xi$, *i.e.*

$$m_{k,j}^n = \int_0^{+\infty}\xi^k f_j(\xi)\mathrm{d}\xi = \sum_{\alpha=1}^{N_q}\bar{w}_{\alpha,j}\xi_{\alpha,j}^k = \sum_{\alpha=1}^{N_q} w_{\alpha,j}\xi_{\alpha,j}^k f_j^n(\xi_{\alpha,j}), \qquad k\in\{0,\dots,N\},$$

where we define $w_{\alpha,j} = \bar{w}_{\alpha,j}/f_j^n(\xi_{\alpha,j})$, since $f_j^n(\xi_{\alpha,j})$ cannot be zero. In the case of EQMOM reconstruction, the weights $\bar{w}_{\alpha,j}$ and abscissas $\xi_{\alpha,j}$ can correspond to the secondary quadrature [21, 12]. This then leads to:

$$m_k^n(x) = \int_0^{+\infty}\xi^k\left[f_j^n(\xi) + D_j^n(\xi)(x-x_j)\right]\mathrm{d}\xi = \sum_{\alpha=1}^{N_q} w_{\alpha,j}\xi_{\alpha,j}^k\left(f_j^n(\xi_{\alpha,j}) + D_j^n(\xi_{\alpha,j})(x-x_j)\right).$$

The corresponding kinetic scheme is then realizable and the flux are written:

$$\mathbf{F}^n_{j+\frac{1}{2}} = \frac{1}{\Delta t^n}\sum_{\alpha=1}^{N_q}\left(w_{\alpha,j}f^n_j(\xi_{\alpha,j}) + \frac{D^n_{\alpha,j}}{2}(X_{j+\frac{1}{2}} - x_{j-\frac{1}{2}})\right)\mathbf{\Xi}_{\alpha,j}\left(x_{j+\frac{1}{2}} - X_{j+\frac{1}{2}}\right)^+ \tag{F.1}$$
$$- \frac{1}{\Delta t^n}\sum_{\alpha=1}^{N_q}\left(w_{\alpha,j+1}f^n_j(\xi_{\alpha,j+1}) - \frac{D^n_{\alpha,j+1}}{2}(x_{j+\frac{3}{2}} - X_{j+\frac{1}{2}})\right)\mathbf{\Xi}_{\alpha,j+1}\left(X_{j+\frac{1}{2}} - x_{j+\frac{1}{2}}\right)^+.$$

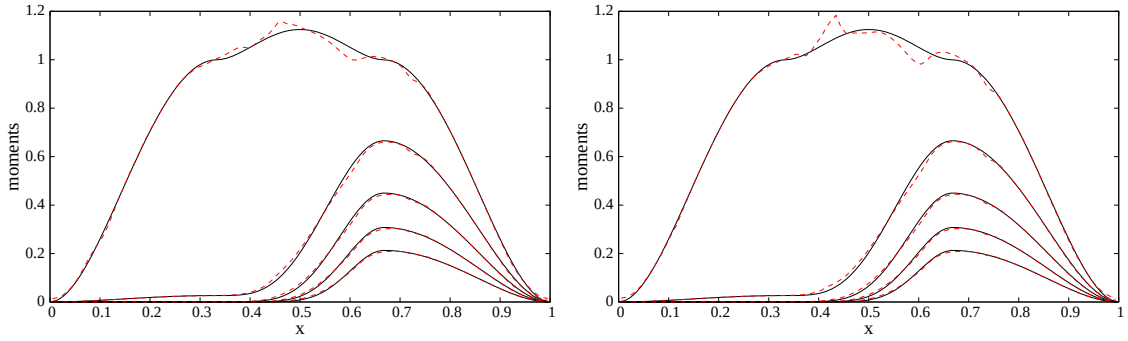

Figure F.15: Constant fluid velocity test case with a multi-modal initial NDF: space evolution of the moments of order 0 to 6 (top to bottom) at time $t = 1$ for the exact solution (solid black lines) and for the simulation with the kinetic scheme using a reconstruction of the NDF with the gamma-EQMOM (left, red dashed lines) and the log-normal-EQMOM (right, red dashed lines) with 100 cells and a CFL equal to 0.8.

However, there are three drawbacks for this kind of scheme. First, it depends on the choice of the NDF reconstruction, since several reconstructions are possible, and it also adds the cost of the reconstruction itself, which is not negligible. However, such kind of reconstruction is usually needed for the other operators present in the complete physical problem. The second and more problematic drawback comes from the impossibility to deal with the boundary of the moment space with this kind of method. Finally, the maximum principle on the moments may not be preserved. Indeed, for example, $f^n(x,\xi)$ is smaller than $\max\{f^n_j(\xi), f^n_{j-1}(\xi), f^n_{j+1}(\xi)\}$ for any value of ξ. But the moments of the distribution $\xi \mapsto \max\{f^n_j(\xi), f^n_{j-1}(\xi), f^n_{j+1}(\xi)\}$ has no reason to be limited by $\max\{m^n_{k,j}, m^n_{k,j+1}, m^n_{k,j-1}\}$. This can be illustrated on a multi-modal but regular initial NDF:

$$f_0(\xi,x) = 9x^2(2-3x)^2\mathbb{1}_{[0,2/3]}(x)R(\xi,\lambda_1,k_1) + 9(3x-1)^2(1-x)^2\mathbb{1}_{[1/3,1]}(x)R(\xi,\lambda(x),k(x)),$$

where R is the Rosin-Rammler pdf and with $\lambda_1 = 0.03$, $k_1 = 2$. The parameters $\lambda(x)$ and $k(x)$ are plotted in Fig. G.16(right). Here, both the gamma and the log-normal EQMOM reconstructions [21, 22, 12] are used to computed fluxes defined by (F.1). The result of the simulations for 5 moments with the corresponding kinetic scheme with 100 cells and a CFL number equal to 0.8 are plotted in Fig. F.15. The maximal principle is not respected for the zeroth order moment and the results depend on the reconstruction, with a slightly better behavior here when using the gamma-EQMOM.

Appendix G. Multi-modal initial NDF

The multi-modal NDF is defined by:

$$f_0(\xi,x) = w_1(x)\delta_{\xi_1}(\xi) + w_2(x)\delta_{2\xi_1}(\xi) + w_3(x)R(x,\lambda(x),k(x)), \tag{G.1}$$

with $\xi_1 = 0.02$ and where R is the Rosin-Rammler pdf:

$$R(\xi, \lambda, k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^{k}\right).$$

The weights w_1, w_2 and w_3 are regular (C^2) and fourth order polynomial by part functions. They are plotted in Fig. G.16(left). The parameters λ and k of the Rosin-Rammler pdf are regular (C^2) and third order polynomials by parts functions plotted in Fig. G.16(right).

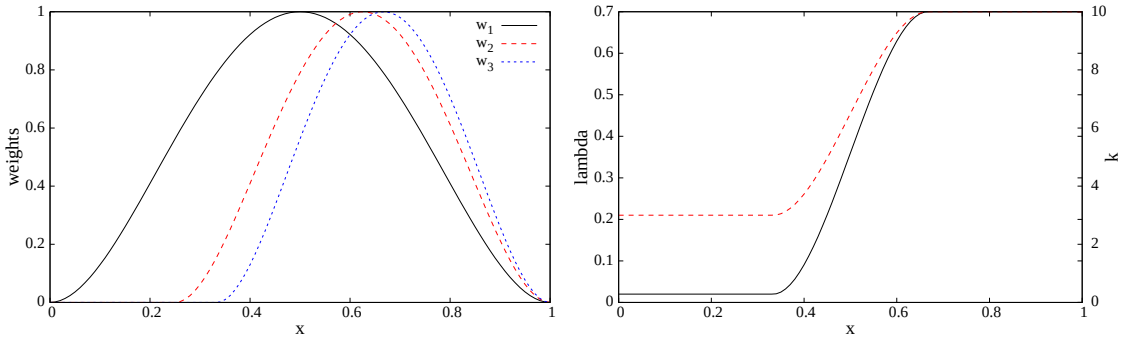

Figure G.16: Left: weights w_1, w_2 and w_3 as functions of the spatial location x. Right: functions $\lambda(x)$ (black solid line) and $k(x)$ (red dashed line).

[1] D. Kah, F. Laurent, M. Massot, S. Jay, A high order moment method simulating evaporation and advection of a polydisperse liquid spray, Journal of Computational Physics 231 (2) (2012) 394–422.

[2] V. Vikas, Z. J. Wang, A. Passalacqua, R. O. Fox, Realizable high-order finite-volume schemes for quadrature-based moment methods, Journal of Computational Physics 230 (13) (2011) 5328 – 5352.

[3] D. Ramkrishna, M. R. Singh, Population balance modeling: Current status and future prospects, Annual Review of Chemical and Biomolecular Engineering 5 (2014) 123–46.

[4] D. L. Marchisio, R. O. Fox, Computational Models for Polydisperse Particulate and Multiphase Systems, Cambridge University Press, Cambridge, UK, 2013.

[5] S. Rigopoulos, Population balance modelling of polydispersed particles in reactive flows, Progress in Energy and Combustion Science 36 (2010) 412–443.

[6] F. Sporleder, Z. Borka, J. Solsvik, H. A. Jakobsen, On the population balance equation, Reviews in Chemical Engineering 28 (2012) 149–169.

[7] A. Zucca, D. L. Marchisio, A. A. Barresi, R. O. Fox, Implementation of the population balance equation in CFD codes for modelling soot formation in turbulent flames, Chemical Engineering Science 61 (1) (2006) 87–95.

[8] R. McGraw, Description of aerosol dynamics by the quadrature method of moments, Aerosol Science and Technology 27 (1997) 255–265.

[9] J. Barrett, N. Webb, A comparison of some approximate methods for solving the aerosol general dynamic equation, Journal of Aerosol Science 29 (1998) 31 – 39.

[10] F. Laurent, A. Sibra, F. Doisneau, Two-size moment multi-fluid model: a robust and high-fidelity description of polydisperse moderately dense evaporating sprays, Commun. Comput. Phys. (2016) 1–41.Accepted, available online at https://hal.archives-ouvertes.fr/hal-01169730.

[11] A. Sibra, J. Dupays, A. Murrone, F. Laurent, M. Massot, Simulation of reactive polydisperse sprays strongly coupled to unsteady flows in solid rocket motors: Efficient strategy using eulerian multi-fluid methods, submitted, available online at https://hal.archives-ouvertes.fr/hal-01063816 (2015).

[12] T. Nguyen, F. Laurent, R. Fox, M. Massot, Solution of population balance equations in applications with fine particles: Mathematical modeling and numerical schemes, Journal of Computational Physics 325 (2016) 129 – 156.

[13] M. Massot, F. Laurent, D. Kah, S. de Chaisemartin, A robust moment method for evaluation of the disappearance rate of evaporating sprays, SIAM J. Appl. Math. 70 (8) (2010) 3203–3234.

[14] M. M. Attarakih, C. Drumm, H.-J. Bart, Solution of the population balance equation using the sectional quadrature method of moments (sqmom), Chemical Engineering Science 64 (4) (2009) 742 – 752, 3rd International Conference on Population Balance Modelling.

[15] M. Frenklach, S. J. Harris, Aerosol dynamics modeling using the method of moments, Journal of Colloid and Interface Science 118.

[16] A. Tagliani, Hausdorff moment problem and maximum entropy: a unified approach, Appl. Math. Comput. 105 (2-3) (1999) 291–305.

[17] L. R. Mead, N. Papanicolaou, Maximum entropy in the problem of moments, J. Math. Phys. 25 (8) (1984) 2404–2417.

[18] A. Vié, F. Laurent, M. Massot, Size-velocity correlations in hybrid high order moment/multifluid methods for polydisperse evaporating sprays: Modeling and numerical issues, Journal of Computational Physics 237 (2013) 177–210.

[19] G. A. Athanassoulis, P. N. Gavriliadis, The truncated Hausdorff moment problem solved by using kernel density functions, Probabilistic Engineering Mechanics 17 (3) (2002) 273–291.

[20] C. Chalons, R. O. Fox, M. Massot, A multi-Gaussian quadrature method of moments for gas-particle flows in a LES framework, in: Proceedings of the Summer Program 2010, Center for Turbulence Research, Stanford University, Stanford, 2010, pp. 347–358.

[21] C. Yuan, F. Laurent, R. O. Fox, An extended quadrature method of moments for population balance equations, Journal of Aerosol Science 51 (2012) 1–23.

[22] E. Madadi-Kandjani, A. Passalacqua, An extended quadrature-based moment method with log-normal kernel density functions, Chemical Engineering Science 131 (2015) 323–339.

[23] J. A. Shohat, J. D. Tamarkin, The Problem of Moments, American Mathematical Society Mathematical surveys, vol. II, American Mathematical Society, New York, 1943.

[24] H. Dette, W. J. Studden, The theory of canonical moments with applications in statistics, probability, and analysis, Wiley Series in Probability and Statistics: Applied Probability and Statistics, John Wiley & Sons Inc., New York, 1997, a Wiley-Interscience Publication.

[25] W. Gautschi, Orthogonal polynomials: applications and computation, Acta numerica 5 (1996) 45–119.

[26] D. L. Wright, Numerical advection of moments of the particle size distribution in Eulerian models, Journal of Aerosol Science 38 (3) (2007) 352–369.

[27] V. Vikas, Z. J. Wang, R. O. Fox, Realizable high-order finite-volume schemes for quadrature-based moment methods applied to diffusion population balance equations, Journal of Computational Physics 249 (2013) 162–179.

[28] R. McGraw. Correcting moment sequences for errors associated with advective transport [online] (2006).

[29] F. Doisneau, A. Sibra, J. Dupays, A. Murrone, F. Laurent, M. Massot, Numerical strategy for unsteady two-way coupling in polydisperse sprays: application to solid rocket motor instabilities, Journal of Propulsion and Power 30 (3) (2014) 727–748.

[30] A. Attili, F. Bisetti, Application of a robust and efficient lagrangian particle scheme to soot transport in turbulent flames, Computers & Fluids 84 (2013) 164 – 175.

[31] M. Essadki, S. de Chaisemartin, S. Jay, M. Massot, F. Laurent, A. Larat, Adaptive mesh refinement for polydisperse spray simulation, Oil & Gas Science and Technology (2016) 1–24.In Press.

[32] R. E. Curto, L. A. Fialkow, Recursiveness, positivity, and truncated moment problems, Houston J. Math. 17 (4) (1991) 603–635.

[33] J. B. Lasserre, Moments, positive polynomials and their applications, Vol. 1 of Imperial College Press Optimization Series, Imperial College Press, London, 2010.

[34] J. Favard, Sur les polynomes de Tchebicheff, Comptes Rendus des Séances de l'Académie des Sciences, Paris 200 (1935) 2052–2053.

[35] T. S. Chihara, An introduction to orthogonal polynomials, Gordon and Breach Science Publishers, New York-London-Paris, 1978, mathematics and its Applications, Vol. 13.

[36] H. S. Wall, Analytic Theory of Continued Fractions, D. Van Nostrand Company, Inc., New York, N. Y., 1948.

[37] H. Rutishauser, Der Quotienten-Differenzen-Algorithmus, Z. Angew. Math. Physik 5 (1954) 233–251.

[38] P. Henrici, The quotient-difference algorithm, Nat. Bur. Standards Appl. Math. Ser. no. (49) (1958) 23–46.

[39] R. G. Gordon, Error bounds in equilibrium statistical mechanics, Journal of Mathematical Physics 9 (5) (1968) 655–663.

[40] R. G. Gordon, Error bounds in spectroscopy and nonequilibrium statistical mechanics, Journal of Mathematical Physics 9 (7) (1968) 1087–1092.

[41] J. Wheeler, Modified moments and Gaussian quadrature, Rocky Mountain J. Math 4 (2) (1974) 287–296.

[42] M. Skibinsky, Extreme nth moments for distributions on [0, 1] and the inverse of a moment space map, J. Appl. Probability 5 (1968) 693–701.

[43] C. K. Garrett, C. D. Hauck, A Comparison of Moment Closures for Linear Kinetic Transport Equations: The Line Source Benchmark, Transport Theory and Statistical Physics 42 (6-7) (2013) 203–235.

[44] F. Schneider, J. Kall, G. Alldredge, A realizability-preserving high-order kinetic scheme using WENO reconstruction for entropy-based moment closures of linear kinetic equations in slab geometry, Kinet. Relat. Models 9 (1) (2016) 193–215.

[45] L. Mead, N. Papanicolaou, Maximum entropy in the problem of moments, J. Math. Phys. 25 (8) (1984) 2404–2417.

[46] F. Bouchut, S. Jin, X. Li, Numerical approximations of pressureless and isothermal gas dynamics, SIAM J. Numer. Anal. 41 (1) (2003) 135–158.

[47] S. Clain, S. Diot, R. Loubre, A high-order finite volume method for systems of conservation lawsmulti-dimensional optimal order detection (mood), Journal of Computational Physics 230 (10) (2011) 4028 – 4050.

[48] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical recipes, 3rd Edition, Cambridge University Press, Cambridge, 2007, the art of scientific computing.

[49] R. LeVeque, Finite volume methods for hyperbolic problems, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2002.

[50] C. Yuan, R. O. Fox, Conditional quadrature method of moments for kinetic equations, J. Comp. Physics 230 (22) (2011) 8216–8246.

[51] S. Gottlieb, D. Ketcheson, C.-W. Shu, Strong Stability Preserving Runge-Kutta and Multistep Time Discretizations, World Scientific, 2011.

[52] C. Berthon, Stability of the MUSCL schemes for the Euler equations, Commun. Math. Sci. 3 (2) (2005) 133–157.

[53] S. Descombes, M. Duarte, T. Dumont, V. Louvet, M. Massot, Adaptive time splitting method for multi-scale evolutionary partial differential equations, Confluentes Mathematici 3 (2011) 1–31, http://hal.archives-ouvertes.fr/hal-00587036.

[54] OpenQBMM, An open-source implementation of quadrature-based moment methods (2015). URL `www.openqbmm.org`

[55] A. Passalacqua, F. Laurent, E. Madadi-Kandjani, J. C. Heylmun, R. O. Fox, An open-source quadrature-based population balance solver for OpenFOAM, in preparation (2016).

[56] R. J. LeVeque, Numerical methods for conservation laws, 2nd Edition, Birkhäuser Verlag, Basel, 1992.