Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

ASYMPTOTICALLY STABLE AND TIME DIMINISHING SCHEMES FOR RAREFIED GAS DYNAMIC

Nicolas Crouseilles

INRIA-Rennes Bretagne Atlantique, IPSO Project
and IRMAR (Université de Rennes 1)
35042 Rennes, France

Giacomo Dimarco

Department of Mathematics and Computer Science
University of Ferrara
Ferrara, Italy

Mohammed Lemou

CNRS and IRMAR (Université de Rennes 1)
and INRIA-Rennes Bretagne Atlantique, IPSO Project
35042 Rennes, France

Abstract. In this work, we introduce a new class of numerical schemes for rarefied gas dynamic problems described by collisional kinetic equations. The idea consists in reformulating the problem using a micro-macro decomposition and successively in solving the microscopic part by using asymptotically stable Monte Carlo methods. We consider two types of decompositions, the first leading to the Euler system of gas dynamics while the second to the Navier-Stokes equations for the macroscopic part. In addition, the particle method which solves the microscopic part is designed in such a way that the global scheme becomes computationally less expensive as the solution approaches the equilibrium state as opposite to standard methods for kinetic equations which computational cost increases with the number of interactions. At the same time, the statistical error due to the particle part of the solution decreases as the system approach the equilibrium state. This causes the method to degenerate to the sole solution of the macroscopic hydrodynamic equations (Euler or Navier-Stokes) in the limit of infinite number of collisions. In a last part, we will show the behaviors of this new approach in comparisons to standard Monte Carlo techniques for solving the kinetic equation by testing it on different problems which typically arise in rarefied gas dynamic simulations.

1. **Introduction.** The numerical simulation of kinetic equations involving many different applications ranging from rarefied gas dynamic and plasma physics to socio-economic models is a very active field of research. In this work, we focus our attention on the development of a new class of methods for rarefied gas dynamic problems. To this aim, concerning gas flow simulations, because of the intrinsic multiscale nature of many problems, it may happen that Euler or Navier-Stokes

2000 *Mathematics Subject Classification.* 65M06, 35B25, 82C80, 82D10, 41A60.

Key words and phrases. Euler equations; Navier-Stokes equations; kinetic equations; asymptotic preserving schemes; time diminishing schemes; micro-macro decomposition; particle methods.

The authors have been supported by the ANR project Moonrise.

models are not sufficient in describing several different phenomena while kinetic models are very often the most adapted approaches. One of the key difference between fluid and kinetic models [9] is the high dimensionality of the mesoscopic approach which describes the state of the system by studying the time evolution of the probability of a particle to be in a given state in the six dimensional phase space at a given time [9, 20]. On the other hand, the fluid models evolve macroscopic quantities such as density, temperature and mean velocity, which depend only on time and on the three dimensional physical space variables. It is a matter of fact that the kinetic description is much richer than the hydrodynamic one but the price to pay is very often too expensive numerical simulations which avoid the use of these approaches in practice.

For this reason, the numerical techniques for solving kinetic equations are often designed in such a way that the computational cost is as low as possible. However, it is undeniable that, even if many progresses have been done in the recent past [20], the goal of simulating realistic problems by means of kinetic models with deterministic techniques has not still be achieved. Thus, in practice, the most frequently used methods are based on probabilistic techniques, the most known method belonging to this category being the Direct Simulation Monte Carlo method (DSMC) [1, 3, 8, 28, 29]. However, this kind of approaches which work very well for stationary problems (thanks typically to time averaging techniques) are only poorly accurate if few particles are used or computationally too expensive if many particles are employed for unsteady problems. To overcome problems related to the low convergence rate and numerical noise of Monte Carlo schemes, several methods have been designed in the recent past. We quote in particular the review papers [8, 30] for an overview on efficient and low variance Monte Carlo methods and some recent papers concerning the use of these techniques for rarefied gas dynamic problems [12, 15, 16, 18, 17, 22, 23, 32] and plasma physics [10, 6, 7].

A second important drawback of kinetic approaches is that the collision term becomes stiff when interactions become important (*i.e.* when the system is close to the hydrodynamic limit). Thus, it turns out that standard explicit schemes lose their efficiency due to the necessity of using very small time steps to solve the collisional scale. In addition, in many situations, the fluid limit may occur in some regions of the domain and at some given times, while the kinetic regime is the most probable one in the rest of the domain. In these cases, one is normally obliged to resolve the micro scales in order to remain stable and consistent, but this requires very small time steps and phase space cells. On the other side, simulations have to be performed on macroscopic lengths, which makes the problem very challenging. Hence, domain decomposition approaches [5, 13, 14, 27, 31, 21] can be adopted. However, the connection of the different models demands specific development as well as the interface identification is not always a simple task to solve. Thus, as an alternative, the problem has been recently addressed by designing the so-called asymptotic preserving schemes [4, 11, 19, 24, 25, 26]. These methods are able to overcome the above numerical restrictions and automatically degenerate to consistent discretizations of the limiting models when the parameters which characterize the microscopic behaviors goes to zero. However, while the computational cost related to the solution of the microscale dynamics is in this way overcame, the dimensionality of the problem remains unchanged even in the limit case.

In this paper, we propose a new approach to solve kinetic collisional equations by using hybrid Monte Carlo techniques which addresses both the problem of the

high dimensionality and of large numerical noise. This work, inspired by [16, 18], modifies a recent approach [10] in order to design a numerical scheme which is more efficient. In particular, as opposite to standard techniques, this new approach reduces the computational cost as the equilibrium state (which can be described by hydrodynamic like systems) is reached. In addition, the statistical error due to the particle part of the solution decreases as the number of interactions increases, realizing a variance reduction method which effectiveness depends on the regime studied.

The solution we propose is the following: we couple a Monte Carlo method for the solution of the kinetic equation with a finite volume method for the solution of the macroscopic equations in each point of the computational domain. Successively, we construct our algorithm in such a way that the computation of the microscopic solution can be avoided by using a simple asymptotic preserving (AP) method ; in such a way, the computational cost of the solution only depends at each time step on the quantity of solution which is out of the equilibrium state. In order to provide the correct solution for all the regimes, the macroscopic moment equations are coupled to the kinetic equation through a kinetic correction term, which takes into account departures from thermodynamical equilibrium. We both consider as a macroscopic equations the compressible Euler and the Navier-Stokes equations. As the equilibrium is approached, the number of particle in the Monte Carlo method is reduced which causes the computational cost to diminish and to be equivalent to the computational cost of a classic numerical method for the hydrodynamic equations in the limit. Thus, the method is automatically costly diminishing without imposing any artificial transition to pass from the microscopic to the macroscopic model at the contrary to domain decomposition techniques in which a transition region should be artificially imposed. In this sense, the schemes proposed here realize an automatic transition from kinetic to hydrodynamic which only depends on the real physics and not on numerical artifacts. Moreover, at the contrary to standard domain decomposition methods in which the cost is low only in regions of pure equilibrium, with these schemes the computational cost continuously diminishes passing from one regime to the other, while at the contrary to standard AP schemes the reduction of the complexity is not only due to the overcoming of the stiffness of the collisional scale, but also to the reduction of the dimensionality.

The remainder of the paper is organized as follows. In the next section we recall some basics about the collisional kinetic equations and their fluid limit. A reformulation of the kinetic equation which permits to design the numerical scheme is described in section 3, while the scheme itself is described in Section 4. Numerical results are presented in Section 5 which show the better performances of this approach with respect to classical Monte Carlo methods. Section 6 is used to draw some conclusions and suggest future developments. In Appendix are reported some detailed computations used through the paper.

2. The kinetic BGK equation, the compressible Euler and Navier-Stokes limits.

We consider a simplified kinetic equation in which the interactions between particles are described by relaxation towards the local Maxwellian equilibrium [9]. This is the so-called BGK model and it is considered to be a valid alternative to

the more complex Boltzmann operator in fluids which are not far from the thermodynamical equilibrium. It reads [9, 20]

$$\partial_t f + v \cdot \nabla_x f = \frac{1}{\varepsilon} Q(f), \tag{1}$$

where $f(x, v, t)$ is the density distribution function, $(x, v) \in \Omega \times \mathbb{R}^{d_v}$, with $\Omega \subset \mathbb{R}^{d_x}$, d_x and $d_v = 1, 2$, or 3 and $t > 0$ the time. We take $d_x = d_v = d$ even if all results extend to the case in which the two spaces have different dimensions. The constant ε is the Knudsen number which measures the time scale of collisions. Collisions are replaced by relaxation

$$Q_{BGK}(f) = \nu\,(M[f] - f), \tag{2}$$

where $\nu = \nu(\rho, T) > 0$ is a given relaxation frequency and measures the average time between two collisions, while ρ and T are the density and temperature of the gas defined below. We indicate the so-called local Maxwellian by $M[f]$ or $M[U]$ to stress the fact that it depends on the distribution function f but only through its moments U. It takes the form

$$M[U](v) = M[\rho, u, T](v) = \frac{\rho}{(2\pi T)^{d/2}} \exp\left(\frac{-|v-u|^2}{2T}\right), \tag{3}$$

where u is the mean velocity. The vector $U = (\rho, u, T)$ is the vector of the macroscopic quantities obtained by integrating the distribution function over the velocity space multiplied by the so-called collision invariants $m(v) = (1, v, |v|^2/2)^T$

$$U = \langle mf \rangle = \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix},$$

where $\langle mf \rangle := \int_{\mathbb{R}^d} f(v) m(v) dv$ and $E = \frac{1}{2}\rho u^2 + \frac{d}{2}\rho T$ is the total energy. Hence, $M(U)$ and f share the same first three moments in v. The kinetic equation is completed by boundary and initial conditions for f. When the gas is dense and temperature is sufficiently low, the Knudsen number is typically very small. In this case, the gas appears macroscopically in equilibrium. This is the fluid limit model [2, 9] obtained taking the limit $\varepsilon \to 0$ in (1). Let us briefly describe this limit. Multiplying (1) by $m(v)$ and integrating with respect to v yields

$$\partial_t \langle mf \rangle + \nabla_x \cdot \langle vmf \rangle = 0.$$

This is equivalent to the following system

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix} + \nabla_x \cdot \begin{pmatrix} \rho u \\ \langle (v \otimes v) f \rangle \\ \frac{1}{2}\langle |v|^2 v f \rangle \end{pmatrix} = 0.$$

When ε goes to zero in (1), the distribution function tends to the local Maxwellian $M(U)$ given by (3). The previous system can then be closed, and the momentum $\langle (v \otimes v) f \rangle = \langle (v \otimes v) M(U) \rangle$ and energy fluxes $\langle |v|^2 v f \rangle = \langle |v|^2 v M(U) \rangle$ can be expressed as a function of U

$$\partial_t U + \nabla_x \cdot \begin{pmatrix} \rho u \\ \rho u \otimes u + pI \\ (E + p) u \end{pmatrix} = 0, \tag{4}$$

with $p = \rho T$ the pressure. Using the Chapman-Enskog expansion $f = \sum_{n=0}^{\infty} \varepsilon^n f_n$, higher order fluid models can be derived. We specifically consider in this work

the first order approximation to the distribution function f, which means we take $f = f_0 + \varepsilon f_1$. This choice gives after simple computations $f_0 = M(U)$ and $f_1 = -(I - \Pi_M)v \cdot \nabla_x M(U)$, with Π_M the orthogonal projection to be precised in the next Section. This leads to the usual compressible Navier-Stokes equations

$$\partial_t U + \nabla_x \cdot \begin{pmatrix} \rho u \\ \rho u \otimes u + pI \\ (E+p)u \end{pmatrix} = -\varepsilon \begin{pmatrix} 0 \\ \nabla_x \cdot \sigma \\ \nabla_x \cdot (\sigma u + \mathcal{Q}) \end{pmatrix}, \tag{5}$$

where $\sigma = -\mu(\nabla_x u + (\nabla_x u)^T - (2/d)\nabla_x \cdot uI)$ is the stress tensor and $\mathcal{Q} = k\nabla_x T$ the heat flux.

3. **Derivation of micro-macro models.** This section is devoted to the derivation of two different micro-macro models starting from equation (1) which are the basis of the numerical methods developed next. The first micro-macro model consider the compressible Euler equation as the macroscopic model while the second micro-macro decomposition employs the Navier-Stokes equations as a macroscopic model. Concerning the first case, the first step we write the distribution function f according to the following decomposition

$$f = M + g, \quad \text{with } M = \frac{\rho}{(2\pi T)^{d/2}} \exp\left(-\frac{|v-u|^2}{2T}\right).$$

Now, since $M = M(U)$ and f shares the same first three moments, we have (recalling $m(v) = (1, v, |v|^2/2)^T$)

$$U(t,x) = \int_{\mathbb{R}^d} m(v) f(t,x,v) dv = \int_{\mathbb{R}^d} m(v) M(t,x,v) dv,$$

which in particular implies $\langle mg \rangle = 0$. Let now $\mathscr{T}$ be the transport operator $\mathscr{T} f = v \cdot \nabla_x f$, then the kinetic equation (1) writes

$$\partial_t M + \partial_t g + \mathscr{T} M + \mathscr{T} g = -\frac{\nu}{\varepsilon} g. \tag{6}$$

Denoting now by Π_M the orthogonal projection in $L^2(M^{-1}dv)$ endowed with the weighted scalar product $(\varphi, \psi)_M = \langle \varphi \psi M^{-1} \rangle$ onto the following space

$$\mathscr{N}(Q) = \text{Span}\left\{M, vM, |v|^2 M\right\}, \tag{7}$$

with $\mathscr{N}(Q)$ the null space of the operator Q, the explicit expression of the projection operator can be analytically computed. This is given by (see [4]), for all $\varphi \in L^2(M^{-1}dv)$

$$\Pi_M(\varphi) = \frac{1}{\rho}\left[\langle \varphi \rangle + \frac{(v-u)\langle (v-u)\varphi \rangle}{T} + \left(\frac{|v-u|^2}{2T} - \frac{1}{2}\right)\left\langle \left(\frac{|v-u|^2}{T} - 1\right)\varphi \right\rangle\right] M. \tag{8}$$

Introducing now the so-called Euler fluxes $F(U)$ corresponding to the fluxes of the compressible Euler equations (4)

$$F(U) = \begin{pmatrix} \rho u \\ \rho u \otimes u + pI \\ (E+p)u \end{pmatrix}, \tag{9}$$

the micro-macro model for the unknowns (g, U), equivalent to the kinetic equation (1), can be written as follows

$$\partial_t g + (I - \Pi_M)\mathscr{T} g = \frac{\nu}{\varepsilon}\left[-g - \frac{\varepsilon}{\nu}(I - \Pi_M)\mathscr{T} M\right], \tag{10a}$$

$$\partial_t U + \nabla_x \cdot F(U) + \nabla_x \cdot \langle vmg\rangle = 0. \tag{10b}$$

Let us consider now the second decomposition. It reads

$$f = f_0 + f_1 + g, \text{ with } f_0 = M, \text{ and } f_1 = -\varepsilon(I - \Pi_M)\mathscr{T} M. \tag{11}$$

Now, since M and f shares the same first three moments, we still have $U(t,x) = \langle mf\rangle = \langle mM\rangle$, which means $\langle mf_1\rangle = \langle mg\rangle = 0$. Injecting the decomposition (11) into the kinetic equation (1) and applying the projection operator Π_M gives

$$\partial_t M + \Pi_M \mathscr{T} M + \Pi_M \mathscr{T}(f_1 + g) = 0,$$

which is equivalent to the following equation on the moments U

$$\partial_t U + \nabla_x \cdot F(U) + \varepsilon \mathscr{D} U + \nabla_x \cdot \langle vmg\rangle = 0,$$

where $\mathscr{D}U = \nabla_x \cdot \langle vm(I - \Pi_M)\mathscr{T} M\rangle$ corresponds to the Navier-Stokes terms (right hand side of (5)). Now, injecting the decomposition (11) into (1) and applying $(I - \Pi_M)$ leads to the microscopic part

$$\partial_t g + (I - \Pi_M)\mathscr{T} g = -\frac{\nu}{\varepsilon}\Big[g + f_1 + \varepsilon(I - \Pi_M)\mathscr{T} M + \varepsilon(I - \Pi_M)\mathscr{T} f_1 + \varepsilon \partial_t f_1\Big],$$

and using the definition of f_1, we finally get the following model

$$\partial_t g + (I - \Pi_M)\mathscr{T} g = -\frac{\nu}{\varepsilon}\Big[g + \varepsilon(I - \Pi_M)(\partial_t f_1 + \mathscr{T} f_1)\Big], \tag{12a}$$

$$\partial_t U + \nabla_x \cdot F(U) + \varepsilon \mathscr{D} U + \nabla_x \cdot \langle vmg\rangle = 0, \tag{12b}$$

$$f_1 = -(I - \Pi_M)\mathscr{T} M. \tag{12c}$$

In the following section we introduce the numerical schemes based on the two cases: system (10) will be referred as Case 1 whereas system (12) will be referred as Case 2.

4. A new Time Diminishing Asymptotic Preserving class of methods for kinetic equations.

The goal is to design a class of schemes for solving kinetic-type equations which avoids the resolution of the small scale dynamics induced by the particle interactions and which cost diminishes as the equilibrium state is approached. The class of scheme derived in this paper is based on a particle approach and on the following requests

- It should be faster than standard deterministic methods (like finite volume or semi-Lagrangian) designed for solving kinetic equations.
- The statistical error should be smaller than the one of standard Monte Carlo schemes.
- The collisional scale should not impose time steps limitations.
- The computational cost should diminish as the equilibrium state is approached. As the same time, the variance should diminish as the number of interactions increases.
- The computational cost should be less than that of domain decomposition approaches.
- The scheme should not need the introduction of artificial interfaces or criteria to pass from one regime to the other one.

The idea can be summarized as follow, at each time step t^n

- Solve the kinetic equation for the perturbation part (10a) or (12a) using a particle method. This will give g^{n+1}, the perturbation values at time $(n+1)$.
- Solve the macroscopic part (10b) or (12b) with a finite volume method where particles are used to evaluate the perturbation terms $\langle vmg^{n+1}\rangle$. This gives the moments value U^{n+1}.
- Modify the perturbation g at time $n+1$ to ensure the zero-moments property at the particle level.
- Eliminate particles with same speed and different sign in order to reduce the global computational cost.

Is is necessary to present this ? The main point where the particles are discard is not discussed...

To compute the solution of the kinetic equation for the perturbation g, a splitting procedure is performed (for both decomposition cases (10) or (12))

- Solve the transport part: $\partial_t g + \mathscr{T} g = 0$.
- Solve the source part for the Case 1 (system (10))

$$\partial_t g = -\frac{\nu g}{\varepsilon} + \Pi_M \mathscr{T} g - (I - \Pi_M)\mathscr{T} M,$$

  or for Case 2 (system (12)),

$$\partial_t g = \Pi_M \mathscr{T} g - \frac{\nu}{\varepsilon}\Big[g + \varepsilon(I - \Pi_M)(\partial_t f_1 + v \cdot \nabla_x f_1)\Big].$$

In both cases, during the second step of the above procedure particles are eliminated or respectively created depending on the shape of the solution at time n. If the amplitude of the perturbation g in one given spatial cell increases, the number of particles increases in order to better describe the departures from equilibrium. If, instead, the amplitude of the perturbation g in one given cell decreases, particles are eliminated. Thus the computational cost depends at each time step on the absolute value of the perturbation function g, no artificial conditions are imposed on the number of particles used, only the discretization of the source term determines the number of samples which are needed to describe the problem. In the following we details the algorithm described for both cases.

4.1. **Solution of the kinetic equation for the perturbation function.** We consider the solution of the equations (10a) and (12a) in a time interval $[0, \Delta t]$ by using Monte Carlo methods. This approximated solution is given by an operator splitting between transport

$$\partial_t g + \mathscr{T} g = 0,$$

and the source terms respectively for the first and the second cases

$$\partial_t g = -\frac{\nu g}{\varepsilon} + \Pi_M \mathscr{T} g - (I - \Pi_M)\mathscr{T} M,$$

$$\partial_t g = \Pi_M \mathscr{T} g - \frac{\nu}{\varepsilon}\Big[g + \varepsilon(I - \Pi_M)(\partial_t f_1 + v \cdot \nabla_x f_1)\Big].$$

In general, high order time splitting schemes can also be used. However, in this work we did not consider this possibility. In the sequel, we restrict to the one-dimensional case in x and v. Let us now introduce a space discretization of the interval $[x_{\max}, x_{\min}]$: $x_j = x_{\min} + j\Delta x,\ j = 0, \ldots, N_x$ with $\Delta x = (x_{\max} - x_{\min})/N_x$ and denote $I_j = [x_{j-1/2}, x_j + 1/2]$ with $x_{j-1/2} = x_j - \Delta x/2,\ j = 0, \ldots, N_x + 1$; we also consider the time discretization of time step Δt: $t^n = n\Delta t, n \geq 0$. The discretization of the space domain is not needed for the transport step which is solved exactly by simply pushing forward in time the particles. However, it is

necessary to solve the source part which acts locally in space by reconstruction of the moments of the perturbation g in the space cells. In a Monte Carlo method, the distribution function g is approximated by a finite set of N particles

$$g(t,x,v) = \sum_{k=1}^{N} \omega_k \delta(x - \mathrm{x}_k(t))\delta(v - \mathrm{v}_k(t)), \tag{13}$$

where $\mathrm{x}_k(t)$ represents the position, $\mathrm{v}_k(t)$ the velocity and ω_k the weight of each particle. Note that for the particles positions and velocities, we use the notation $\mathrm{x}_k(t)$ and $\mathrm{v}_k(t)$ whereas x_j and v_j are used for the fixed phase space mesh. The initial position, the initial velocity and the weight are defined as following. Let N_p be the total number of particles used to discretize the original problem (1), where the unknown is the distribution function f. Then the mass m_p of each particle is simply defined as

$$m_p = \frac{1}{N_p} \int_{x_{\min}}^{x_{\max}} \int_{\mathbb{R}} f(t=0,x,v)dvdx. \tag{14}$$

In the same way, we define N_j^0 the number of initial particles in each cell I_j (centered around x_j) is given by

$$N_j^0 = \frac{1}{m_p} \int \Big| g(t=0,x_j,v) \Big| dv. \tag{15}$$

Then, one computes the number of particles with positive weight $\omega_k = m_p$ and negative weight $\omega_k = -m_p$ by computing the integer numbers

$$N_j^{0,+} = \frac{1}{m_p} \int g^+(t=0,x_j,v)dv, \;\; N_j^{0,-} = \frac{1}{m_p} \int \Big| g^-(t=0,x_j,v) \Big| dv, \tag{16}$$

where $g^\pm = (g \pm |g|)/2$. Note that we have $N_j^{0,+} + N_j^{0,-} = N_j^0$. Formulas (15) and (16) should be interpreted through a stochastic rounding procedure defined as

$$\mathrm{Iround} = \begin{cases} \lfloor x \rfloor + 1, & \text{with probability } \; x - \lfloor x \rfloor \\ \lfloor x \rfloor, & \text{with probability } \; 1 - x + \lfloor x \rfloor \end{cases} \tag{17}$$

with $\lfloor x \rfloor$ the integer part of x. Once the initial number of samples is fixed, positions are uniformly assigned over the cells centered in x_j. Successively, fixing a grid in velocity space of size Δv and the boundaries of the velocity space $v_{\min}$ and $v_{\max}$, we approximate $g(t=0,x_j,\cdot)$ by a piecewise constant function on each velocity cell of size Δv. Then, in each cell x_j, we have to sample N_j^0 particles in the velocity direction. To do so, the velocities are assigned by inverting the following equation

$$G^\pm(t=0,x_j,\mathrm{v}_k^\pm(0)) = \xi_k,$$

where $\mathrm{v}_k^\pm$ are the particles velocity for g with a weight $\omega_k = \pm m_p$, and where ξ_k is a random number in the interval $[0,1]$; finally $G^\pm(t=0,x_j,v) = \int_{-\infty}^{v} |g^\pm(t=0,x_j,y)|dy$.

We discuss now the two steps of the splitting. Suppose that g^n, M^n and f_1^n are known as well as x_k^n and v_k^n for $k = 1,\dots,N_g^n$, where $N_g^n = \sum_{j=0}^{N_x} N_j^n$ and $N_j^n = N_j^{n,+} + N_j^{n,-}$ with $N_j^{n,\pm} = \frac{1}{m_p} \int |g^\pm(t^n,x_j,v)|dv$. During the transport step, we solve the characteristic equations between t^n and t^{n+1}

$$\dot{\mathrm{x}}_k(t) = \mathrm{v}_k(t), \;\; \mathrm{x}_k(t^n) = \mathrm{x}_k^n, \tag{18}$$

which means that particles are pushed forward in time. Any type of time integrator can be employed to solve this step, we consider the standard first order Euler solver

which, knowing x_k^n, v_k^n reads $\mathrm{x}_k^\star = \mathrm{x}_k^n + \Delta t\, \mathrm{v}_k^n$. Obviously, more accurate solvers can be used, we refer to [20] for an overview of time integration techniques. The second part of the splitting is discretized by an implicit-explicit scheme which permits to deal with the stiff terms. In the first case (10), it reads

$$\tilde{g}^{n+1} = g^\star - \frac{\Delta t \nu}{\varepsilon} g^{n+1} + \Delta t \Big[\Pi_M \mathscr{T} g^n - (I - \Pi_M) \mathscr{T} M^n \Big],$$

which gives

$$\tilde{g}^{n+1} = q\, g^\star + (1-q)\mathscr{P}, \quad \text{with } q = \frac{\varepsilon/\nu}{\varepsilon/\nu + \Delta t}. \tag{19}$$

In the second case (12), it reads

$$\tilde{g}^{n+1} = g^\star - \frac{\Delta t \nu}{\varepsilon} g^{n+1} + \Delta t \Big[\Pi_M \mathscr{T} g^n - \varepsilon (I - \Pi_M)(\partial_t f_1 + v \cdot \nabla_x f_1) \Big],$$

which gives

$$\tilde{g}^{n+1} = q\, g^\star + (1-q)\mathscr{P}_1, \quad \text{with } q = \frac{\varepsilon/\nu}{\varepsilon/\nu + \Delta t}. \tag{20}$$

In the above equations, $g^\star$ represents the solution after the transport step and the precise expression for

$$\mathscr{P}(x,v) := \varepsilon/\nu \left[\Pi_M \mathscr{T} g - (I - \Pi_M) \mathscr{T} M \right], \tag{21}$$

$$\mathscr{P}_1(x,v) := \varepsilon/\nu \left[\Pi_M \mathscr{T} g - \varepsilon (I - \Pi_M)(\partial_t f_1 + v \cdot \nabla_x f_1) \right], \tag{22}$$

are detailed in Appendix A. Equations (19)-(20) are convex combinations of two terms for each choice of the time step Δt. The first term $g^\star$ is a sink, the second one $\mathscr{P}$ or $\mathscr{P}_1$ is a source. From a Monte Carlo point of view, these equations state that with probability $(1-q) = \frac{\Delta t}{\varepsilon/\nu + \Delta t}$, particles are discarded and with a probability $q = \frac{\varepsilon/\nu}{\varepsilon/\nu + \Delta t}$, particles are created from a sampling of the sources $\mathscr{P}(x,v)$ or $\mathscr{P}_1(x,v)$. More precisely, after the transport step, the functions $\mathscr{P}(x_j,v)$ and $\mathscr{P}_1(x_j,v)$ are computed in each cell I_j. After the pushing step (18), we first need to evaluate the cell to which each particle belongs to. Then, the algorithm to approximate (19) or (20) is the following

- Discard $(1-q)N_j^{n,\pm}$ particles from cell I_j.
- Keep $qN_j^{n,\pm}$ particles from cell I_j.
- Create $(1-q)M_j^\pm$ new particles in cell I_j, where

$$M_j^\pm = \frac{1}{m_p} \int \Big| \mathscr{P}^\pm(x_j,v) \Big| dv$$

  by sampling $\mathscr{P}^+(x_j,v)$ for positive weights particles and $\mathscr{P}^-(x_j,v)$ for negative weight particles (this is done as explained previously). This gives the total new number of particles in the cell I_j: $N_j^{n+1,\pm} = qN_j^{n,\pm} + (1-q)M_j^{n,\pm}$.

The second case is considered by simply changing $\mathscr{P}$ by $\mathscr{P}_1$ given by (22). As for the function g, we defined $\mathscr{P}^\pm(x_j,v)$ by $\mathscr{P}^\pm(x_j,v) = \frac{1}{2}[\mathscr{P}^\pm(x_j,v) \pm \mathscr{P}^\pm(x_j,v)]$ (and similarly for $\mathscr{P}_1^\pm(x_j,v)$). The positions for particles which do not disappear remain unchanged $\mathrm{x}_k^{n+1} = \mathrm{x}_k^\star$ while the new positions x_k^{n+1} for the new samples are computed uniformly over the space cell. Concerning the velocities, they remain the same if the particle is kept $\tilde{\mathrm{v}}_k^{n+1} = \mathrm{v}_k^\star$ while their values depend on $\mathscr{P}[v,M]$ and $\mathscr{P}_1[v,f_1]$ in the other case. Let us observe that we use $\tilde{\mathrm{v}}$ and $\tilde{g}$ to indicate that one last passage is necessary to define the new velocities and consequently the new g which is discussed in section 4.3.

In the asymptotic regime $\varepsilon \to 0$, we have $q = \mathcal{O}(\varepsilon)$ and $M_j^{\pm} = \mathcal{O}(\varepsilon)$ since $\mathscr{P} = \mathcal{O}(\varepsilon)$ and $\mathscr{P}_1[g, f_1] = \mathcal{O}(\varepsilon)$ thanks to (21) and (22). As a consequence, the number of particles in the cell I_j used to sample g satisfies

$$N_j^{n+1} = N_j^{n+1,+} + N_j^{n+1,-}, \text{ with } N_j^{n+1,\pm} = qN_j^{n,\pm} + (1-q)M_j^{n,\pm} = \mathcal{O}(\varepsilon),$$

which means that the number of particles diminishes automatically as $\varepsilon \to 0$.

Note that a similar approach has been developed in [32] where the micro part of the micro-macro decomposition is approximated by a hybrid Monte Carlo type method with negative particles (see [33]) are used.

4.2. **Numerical discretization of the macroscopic equations.** In this section we discuss the discretization of the moment equations (10b)

$$\partial_t U + \nabla_x \cdot F(U) + \nabla_x \cdot \langle vmg \rangle = 0,$$

and (12b)

$$\partial_t U + \nabla_x \cdot F(U) + \varepsilon \mathscr{D} U + \nabla_x \cdot \langle vmg \rangle = 0.$$

The first system is composed by an equilibrium part which corresponds to the classical compressible Euler equations and by a non equilibrium part which takes into account departures from equilibrium. The second system is composed by the same equilibrium part which gives the compressible Euler fluxes and by a diffusion part which gives the Navier-Stokes fluxes. The time discretizations are the following

$$U^{n+1} = U^n - \Delta t \nabla_x \cdot F(U^n) - \nabla_x \cdot \langle vmg^n \rangle, \tag{23}$$

and

$$U^{n+1} = U^n - \Delta t \nabla_x \cdot F(U^n) - \varepsilon \mathscr{D} U^n - \nabla_x \cdot \langle vmg^n \rangle. \tag{24}$$

For the space discretization of the Euler fluxes $F(U)$ in both systems, we consider second order MUSCL central schemes. They read

$$\psi_{j+1/2}(U^n) = \frac{1}{2}(F(U_j^n) + F(U_{j+1}^n)) - \frac{1}{2}\alpha(U_{j+1}^n - U_j^n) + \frac{1}{4}(\sigma_j^{n,+} - \sigma_{j+1}^{n,-}) \tag{25}$$

where

$$\sigma_i^{n,\pm} = \left(F(U_{j+1}^n) \pm \alpha U_{j+1}^n - F(U_j^n) \mp \alpha U_j^n\right) \varphi(\chi_j^{n,\pm}) \tag{26}$$

with φ a modified slope limiter, α equal to the larger eigenvalue of the Euler system and $\chi_j^{n,\pm}$ the ratio of consecutive discrete gradients of the fluxes. The slope limiters proposed here intend to avoid the increase of the statistical fluctuations which may occur by using high order reconstructions, in particular when the perturbation g is large. In such a case, it is better to smooth out the fluctuations by passing from a second order scheme to a first order one. Thus, the limiters are

$$\varphi(\chi_j^{n,\pm}) = \varphi_c(\chi_j^{n,\pm})\beta(\tilde{\chi}_j^{n,\pm}),$$

where $\varphi_c(\chi_j^{n,\pm})$ is the usual slope limiter (we used the Van Leer one), and $\beta(\tilde{\chi}_j^{n,\pm})$ is the unit step function which depends on $\tilde{\chi}_j^{n,\pm}$, the ratio of the discrete gradients of the equilibrium and non equilibrium fluxes. In details, when the non equilibrium fluxes exceed the compressible Euler fluxes of a certain size, then, automatically, the discrete second order fluxes pass to discrete first order ones. In the numerical test section, we fixed this threshold to 0.25.

We now discuss how to discretize the non equilibrium terms $\nabla_x \cdot \langle vmg^n \rangle$ and the Navier-Stokes term $\varepsilon \mathscr{D} U^n$. The Navier-Stokes term is discretized by employing the same second order MUSCL scheme used for the Euler fluxes where the numerical diffusion α is taken equal to zero. Concerning the perturbation terms, we introduce

a filter to eliminate eventual fluctuations. Indeed, we choose the weighted moving average method, which consists of a convolution of the pointwise space values of g^n with a fixed weighting function. Thus, given the function g^n known on the spatial mesh, we define

$$\bar{g}_j^n = \frac{1}{2K+1} \sum_{k=-K}^{k=K} \gamma_k g_{j-k}^n, \qquad \sum_{k=-K}^{k=K} \gamma_k = 1. \tag{27}$$

The smoothing permits to reduce the fluctuations but it also causes a degradation of the accuracy in the solution, thus only a weak filter is used for the numerical tests which employs only three cells. Once the filter is imposed, a first order discretization is employed for the non equilibrium term without numerical diffusion, *i.e.* $\alpha = 0$. It reads

$$\Psi_{j+1/2}(\langle vm\bar{g}^n\rangle) = \frac{1}{2}(\langle vm\bar{g}_j^n\rangle + \langle vm\bar{g}_{j+1}^n\rangle). \tag{28}$$

4.3. **Projection step.** In this section we discuss the projection of the function g onto $\mathscr{N}(Q)$ given by (7). Indeed, due to the use of a Monte Carlo method, even if the moments of g vanish at the beginning of the iteration, we can not ensure that $\langle mg\rangle = 0$ in each cell around x_j at the end of the iteration. Hence, to ensure this property, a projection step is performed at the end of the time step. This projection step is done locally on the cells, for this reason we only consider one fixed spatial cell, at a fixed time step, the same procedure is repeated for all cells and at each time step (we then omit the superscript ^{n} in the sequel).

In order to proceed, we suppose that the solution $\tilde{g}$ is known at a given time step and everywhere by particles with weight ω_k, velocity $\tilde{\mathrm{v}}_k$ and position x_k, for $k = 1, \ldots, N_g$ (with $N_g = \sum_{j=0}^{N_x} N_j$ and $N_j = N_j^+ + N_j^-$ where $N_j^\pm$ are defined by (16) forgetting the time dependency), after using for one time step the scheme presented in section 4.1 from t^{n-1} to t^n. Now, what we should ensure

$$\sum_{k=1}^{N_j} m(\mathrm{v}_k)\omega_k = 0 \tag{29}$$

in each cell centered around x_j. In the sequel, we discuss how to restore the three conservations.

4.3.1. *Conservation of mass.* The loss of mass conservation is due to the second part of the splitting, namely (19) and (20). In the sequel, we detail a way to ensure that the local mass conservation is ensured, *i.e.* $\langle g\rangle = 0$ in each spatial cell.

If $\sum_{k=1}^{N_j} \omega_k \neq 0$ in one given cell centered around x_j, it means that $N_j^+ \neq N_j^-$. Thus, in order to restore equilibrium between positive and negative weight samples, we can choose to eliminate some samples or to create new ones. The best choice is the following: if $N_j^+ > N_j^-$ we eliminate $(N_j^+ - N_j^-)/2$ samples with positive weight and we create $(N_j^+ - N_j^-)/2$ samples with negative weight. If conversely $N_j^- > N_j^+$ we eliminate $(N_j^- - N_j^+)/2$ samples with negative weight and we create $(N_j^- - N_j^+)/2$ samples with positive weight. The elimination is done uniformly while for a creation of a new particle we proceed as following. First, we search in the set of the particles which have been discarded by computing the solution of (19) or (20), if there are enough particles. In this case, samples are taken with uniform probability from the positive or negative discarded samples. If conversely there are not enough particles in the discarded set, new samples are created by replicating

existing ones. When a new particle is created, its position is chosen uniformly in the cell grid. Finally, let observe that $|N_j^- + N_j^+|$ can be an odd number, in this case the last particle we need to restore equilibrium is taken with the same probability from the positive or negative weight sets.

4.3.2. *Conservation of momentum and energy.* In order to ensure conservation of momentum and energy, we propose a modification of the renormalization used in [13]. The strategy is composed of two steps: in the first one, we adjust the average momentum by simply translating the samples, while in the second one, we adjust the energy by using two different scalings for the positive and negative samples. In details, the transformation is the following. Suppose to be known $\tilde{\mathrm{v}}_k^{n+1}$, the velocities after the transport and the relaxation steps described previously; suppose also that the mass conservation process described in the previous paragraph in each spatial cell has been done, then the momentum conservation is ensured thanks to the following transformation

$$\bar{\mathrm{v}}_k^{n+1} = \tilde{\mathrm{v}}_k^{n+1} - \left(\sum_{k=1}^{N_j} \omega_k \tilde{\mathrm{v}}_k^{n+1} \right) / N_j, \tag{30}$$

which is done for all particles which belongs to I_j (the cell centered around x_j). Then, after computing the energy associated with the positive and the negative samples as

$$E^{\pm} = \sum_{k=1}^{N_j^{\pm}} \frac{1}{2} \omega_k (\bar{\mathrm{v}}_k^{n+1})^2 \tag{31}$$

the following transformation is imposed

$$\mathrm{v}_k^{n+1} = c^{\pm} \bar{\mathrm{v}}_k^{n+1} \quad \forall k = 1, \cdots, N_j^{\pm}, \tag{32}$$

where c^+ and c^- are defined by

$$c^+ = \sqrt{\frac{E^+ + (E^- - E^+)/2}{E^+}}, \quad c^- = \sqrt{\frac{E^+ + (E^+ - E^-)/2}{E^-}}. \tag{33}$$

This is enough to ensure that the energy associated to the perturbation g is zero in each cell at each instant of time. We conclude this section by briefly describing the complete algorithm.

4.4. **Algorithm.** We first give the algorithm in the case of decomposition (10). Suppose that two grids of width Δx and a time step Δt have been fixed. Suppose also that the moments U^n and the function g^n are known in each point of the spatial grid at time n as well as particles number N_g^n, positions x_k, velocities v_k and weights ω_k. Then the algorithm for passing from time t^n to time t^{n+1} is the following:

1. Push forward the particles by solving (18).
2. Compute, in each cell of the spatial grid the quantities $(I - \Pi_M)(\mathscr{T} M)$ and $\Pi_M(\mathscr{T} g)$, and then sample them.
3. Eliminate particles from the distribution g and create new particles sampled from $\mathscr{P}(v)$ using (19).
4. Create new positive or negative samples to guarantee mass conservation (see 4.3).
5. Modify the velocities v_k to ensure $\langle mg \rangle = 0$ in each cell (see 4.3).
6. Compute $\nabla_x \cdot \langle vmg \rangle$ using (27) and (28).

7. Advance the macroscopic moments equation using (23).

Let us now consider Suppose now the case of decomposition (12). As before, two grids of width Δx and a time step Δt are fixed. Suppose also that the moments U^n and the functions f_1^n are known in each cell of the spatial grid at time n as well as particle number N_g^n, positions x_k, velocities v_k and weights ω_k. Then the algorithm for the decomposition (12) for passing from time t^n to time t^{n+1} is the following

1. Push forward the particles by solving (18).
2. Compute, in each cell of the spatial grid the quantities $(I-\Pi_M)(\partial_t f_1+v\cdot\nabla_x f_1)$ and $\Pi_M(\mathscr{T} g)$, and then sample them.
3. Eliminate particles from the the distribution g and create new particles sampled from $\mathscr{P}_1(v)$ using (20).
4. Create new positive or negative samples to guarantee mass conservation (see 4.3).
5. Modify the velocities v_k to ensure $\langle mg\rangle = 0$ (see 4.3).
6. Compute $\nabla_x \cdot \langle vmg\rangle$ using (27) and (28).
7. Advance the macroscopic moments equation using (24).

5. **Numerical results.** In this section we discuss several numerical results obtained with the schemes presented. From now on, the Asymptotic Preserving and Time Diminishing methods described in the Algorithm 4.4 will be called respectively APTD for the decomposition (10) and APTDNS (Navier-Stokes) for decomposition (12). The two approaches will be compared with a classical Monte Carlo method for (1) which we call MC. In order to have a reference solution we also use a deterministic second order in space scheme based on a Discrete Velocity Model formulation (DVM scheme) of (1) and a finite volume method for the hydrodynamic model (4) which we call Euler scheme.

For all tests, the number of cells in the physical space will be the same for the five cited methods, while for the APTD, the APTDNS and the MC methods the mass of a test particle will be the same. As clear from the description of the APTD and APTDNS methods in the previous Sections, given the mass of a particle for the MC method m_p, the mass of a particle for the APTD or the APTDNS methods will be either m_p or $-m_p$ according to the sign of g.

This Section is divided into three parts dedicated to test our method against a smooth solution and against some classical problems arising in gas dynamics.

5.1. **Test 1: Unsteady Shock.** In this test, a gas with constant density ($\rho = 1$) and temperature $T = 5$ is pushed against a wall at a speed $u = -1$ with perfectly reflecting boundary conditions. In the fluid limit such an initialization produces a shock wave which moves from the left to the right of the domain, after the shock, the speed of the flow being equal to zero. When collisions are not sufficient to reach the equilibrium state at each instant of time, the shock wave is smoothed out and a continuous transition from the right to the left state of the macroscopic quantities is realized which is faster as the fluid limit is approached.

The numerical parameters are chosen as follows. The number of cells in the physical space is equal to $N_x = 300$, the final time is set equal to $t_{\mathrm{f}} = 0.067$ while the test has been run for $\varepsilon = 10^{-2}, 10^{-3}, 10^{-4}$. The number of particles used in the Monte Carlo method is at the beginning $N = 10^6$ and it grows with time since the density is physically non conserved in the domain for this test case. For the APTD and APTDNS methods, the mass of each particle is the same as the mass of the

particles in the original MC method ($m_p \simeq 10^{-7}$). Concerning the DVM and the Euler schemes, a second order MUSCL scheme has been used in both cases with a number of points in velocity space for the DVM equal to 120.

In Figure 1, the density, the mean velocity and the temperature for respectively the APTDNS, on the left and the MC scheme, on the right, are reported as a function of x for a rarefied regime which corresponds to $\varepsilon = 10^{-2}$. In Figure 2, the same macroscopic quantities are reported for $\varepsilon = 10^{-3}$ which corresponds to the intermediate regime. Finally in Figure 3, these quantites are reported for a regime close to the fluid limit, $\varepsilon = 10^{-4}$. The results show as expected that the reduction of the variance is more important when the interactions between the molecules become more important. For the three different interaction scales chosen, an important reduction of the statistical noise with respect to the classical Monte Carlo scheme is observed. In Figure 4 (left), we can observe that the number of particles effectively employed in the APTD method in comparison with the number of particles used by Monte Carlo. On the middle part of Figure 4, we report the ratio between these two numbers where we clearly see that, as expected, this number decreases as the equilibrium is approached. In particular we observe that, in the worst studied scenario ($\varepsilon = 10^{-2}$), the number of particles effectively used is less than 1.5% of the number of samples of the corresponding MC method, while this ratio goes to 0.01% for $\varepsilon = 10^{-4}$. This gives precise indications concerning the computational cost of the method, since most of the computations are due to the kinetic part of the model, the computations due to the macroscopic part of the solution being negligible. In Figure 4 (right), we finally report the ratio between the number of particles used by the APTD and the APTDNS schemes as a function of time for the different values of the Knudsen number we considered. We see that this ratio is always less than 1 and in particular we see that in the case of small Knudsen number, the number of particles used in the APTDNS is around 25% less than the number of particles of the APTD scheme.

5.2. **Test 2: Sod problem.** The initial data consists in a Riemann problem with discontinuity located at $x = L/2$ with $L = 1$ with Dirichlet boundary conditions and $\rho_\ell = 1$, $u_\ell = 0$, $T_\ell = 5$ on the left and $\rho_r = 0.125$, $u_r = 0$, $T_r = 4$ on the right. The number of cells in physical space is equal to $N_x = 300$, the final time is set equal to $t_{\mathrm{f}} = 0.05$ while the test has been run for different values of $\varepsilon = 10^{-2}, 10^{-3}, 10^{-4}$ as before. The number of particles used in the Monte Carlo method is at the beginning $N = 10^6$ and it remains in average constant with time. For the APTD and APTDNS methods, the mass of each particle is the same as the mass of the particles in the original MC method, *i.e.* $m_p \simeq 10^{-7}$. Concerning the DVM and the Euler schemes, a second order MUSCL scheme has been used in both cases with a number of points in velocity space for the DVM equal to 120.

In Figures 5, 6 and 7, the density, the mean velocity and the temperature for respectively the APTDNS, on the left and the MC scheme, on the right, are reported for the same regimes considered in the first test: $\varepsilon = 10^{-2}, 10^{-3}, \varepsilon = 10^{-4}$. The results show as before that the reduction of the statistical error is more important when the interactions grow. For all regimes considered, our schemes give less noisy results compares to Monte Carlo. Finally, Figure 8 (left part) shows the number of particles used by the APTD method in comparison with the number of particles used by the Monte Carlo one. In the middle of Figure 8, the ratio between these two numbers is reported while on the right part, the ratio between the number of

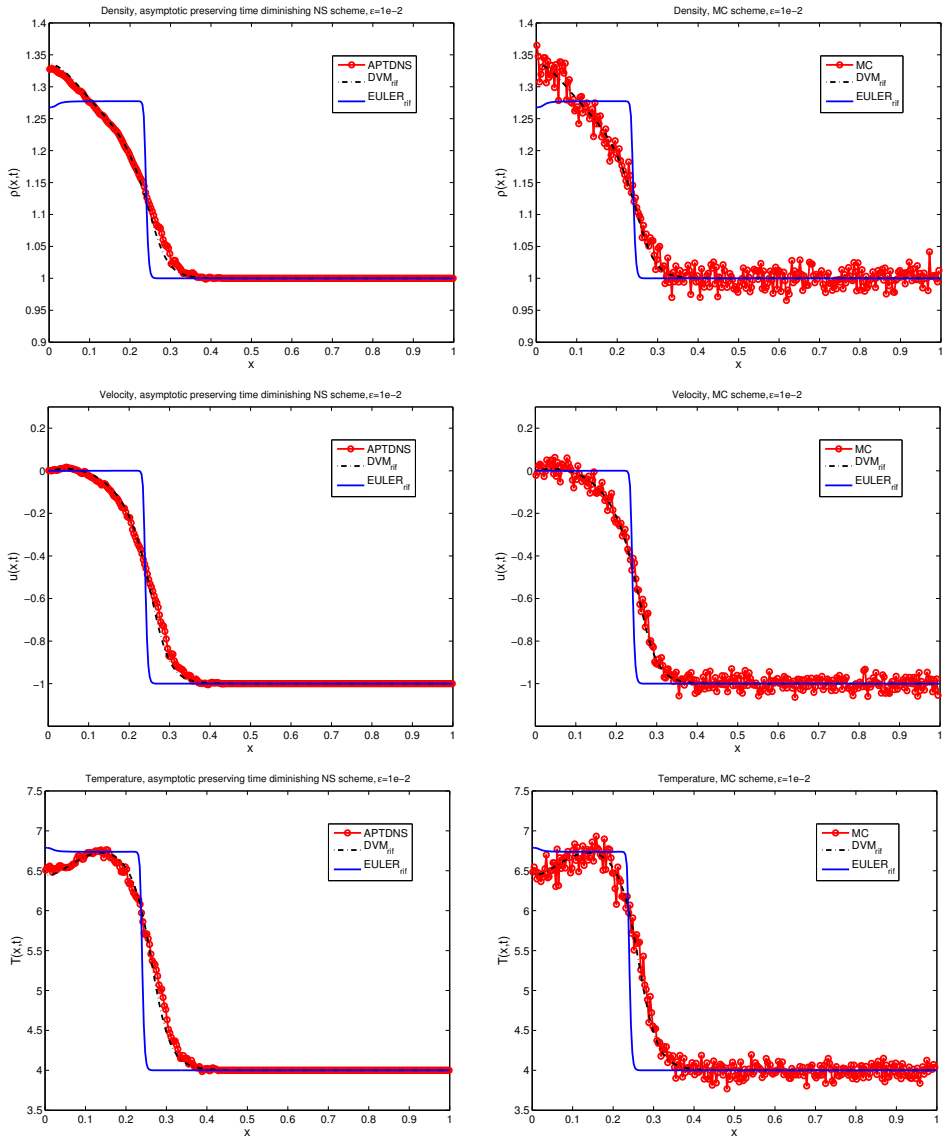

FIGURE 1. Density, velocity and temperature profiles (from top to bottom) for: left Asymptotic Preserving Time Diminishing Navier-Stokes method, right Monte Carlo method, $\varepsilon = 10^{-2}$. The dotted line is a reference solution computed with a deterministic DVM method while the continuous line is a reference solution for the compressible Euler equations. Unsteady Shock test.

particles used by the APTD and the APTDNS schemes is shown. The results can be resumed by saying that the APTD methods produce less noisy solutions with less particles and this number decreases when the solution becomes closer to the fluid regime. In particular, the APTD method which employs the second decomposition

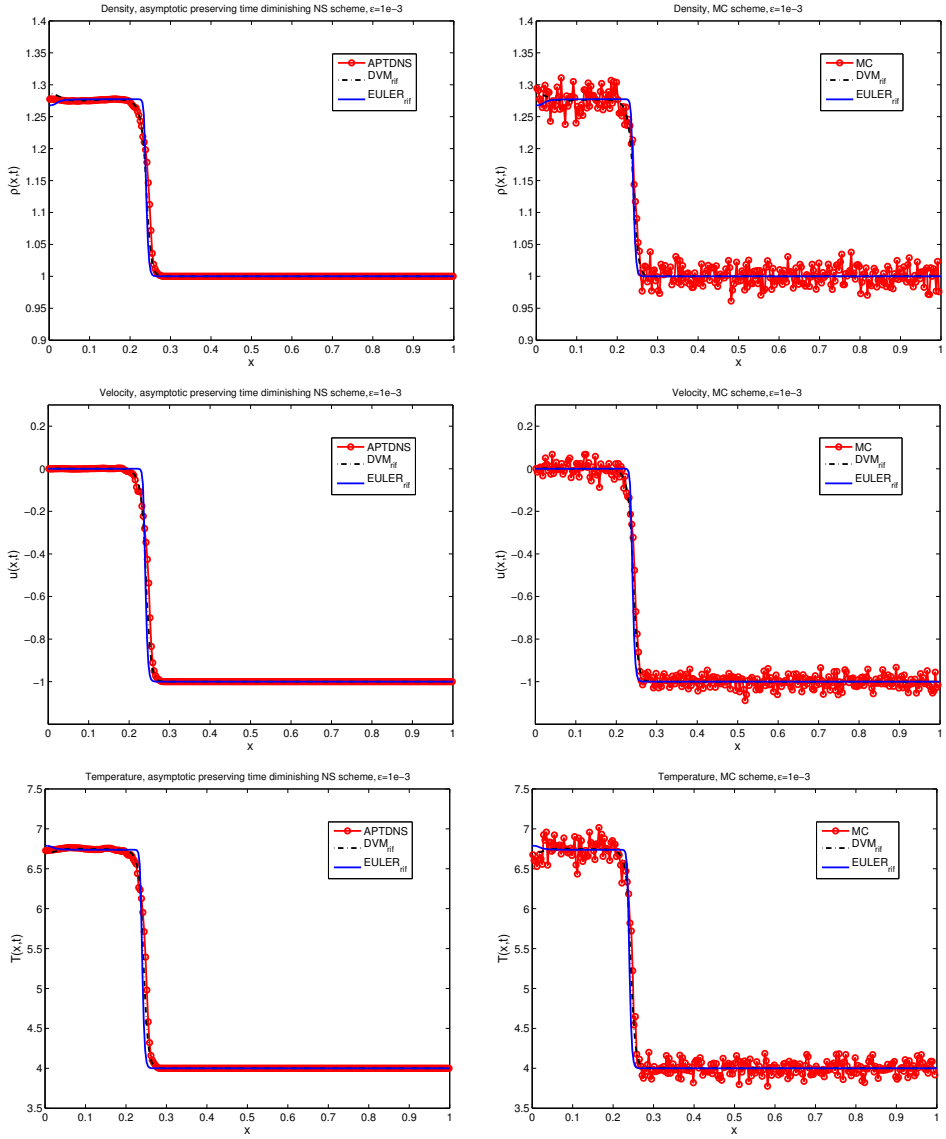

FIGURE 2. Density, velocity and temperature profiles (from top to bottom) for: left Asymptotic Preserving Time Diminishing Navier-Stokes method, right Monte Carlo method, $\varepsilon = 10^{-3}$. The dotted line is a reference solution computed with a deterministic DVM method while the continuous line is a reference solution for the compressible Euler equations. Unsteady Shock test.

(12) is able to produce solutions with less particles than the APTD method since the Navier-Stokes terms are computed in a deterministic way.

5.3. **Test 3: smooth solution problem.** In this last test, we consider an initial smooth solution as a Maxwellian with density $\rho = 1 + 0.5\sin(kx)$ (with $k = 2\pi/L$),

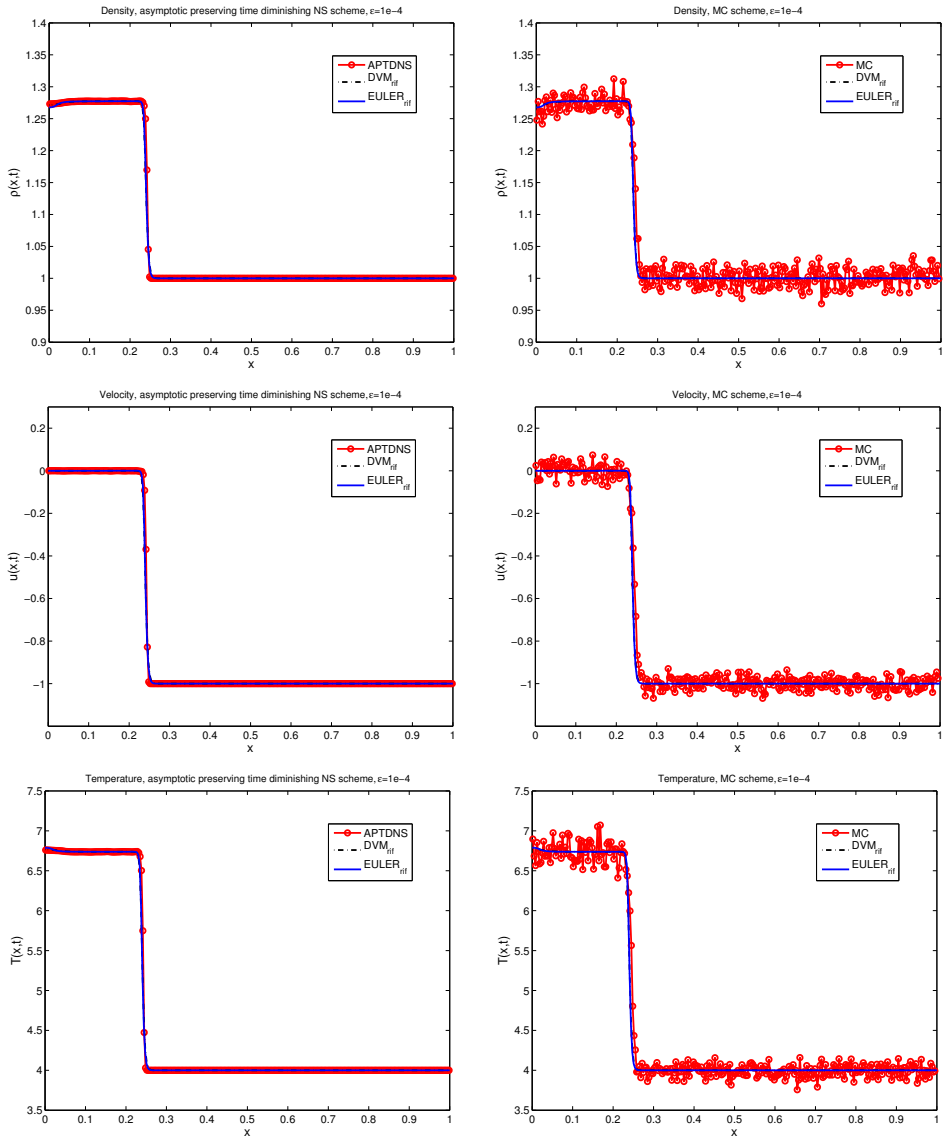

FIGURE 3. Density, velocity and temperature profiles (from top to bottom) for: left Asymptotic Preserving Time Diminishing Navier-Stokes method, right Monte Carlo method, $\varepsilon = 10^{-4}$. The dotted line is a reference solution computed with a deterministic DVM method while the continuous line is a reference solution for the compressible Euler equations. Unsteady Shock test.

mean velocity $u = 0.5$ and temperature $T = 4$ with $L = 1$ the length of the periodic domain. The final time is fixed to $t_{\mathrm{f}} = 0.05$ while the number of cells in space is $N_x = 300$. We consider three different values of the Knudsen number $\varepsilon = 10^{-2}, 10^{-3}, 10^{-4}$ and for each of these values, we compute the solution by

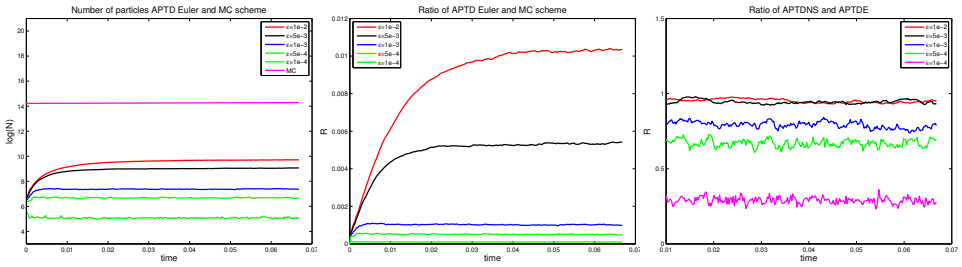

FIGURE 4. Left: time evolution of the number of effective particles (in semi-logarithmic scale) used in the Asymptotic Preserving Time Diminishing (Euler) method and in the Monte Carlo method for different values of the Knudsen number ($\varepsilon = 10^{-2}, 5\cdot 10^{-3}, 10^{-3}, 5\cdot 10^{-4}, 10^{-4}$). Middle: time evolution of the ratio of the number of particles used for APTD versus the number of particles for the corresponding MC simulation for different values of ε. Right: time evolution of the ratio of the number of particles used for APTD versus the number of particles for APTDNS for different values of ε. Unsteady Shock test.

using the APTDNS method and the classical MC scheme using different numbers of particles $N_p = 7.5\cdot 10^4, 1.5\cdot 10^5, 3\cdot 10^5, 6\cdot 10^5$. Let observe that the number of particles effectively used for the APTDNS method depends as in the previous test cases on the size of g. The obtained results show that the ratio between the number of particles of the APTDNS scheme and the MC scheme for this problem depends on the Knudsen number values and this ratio is very close to the previous results; for this reason, we do not report the results for this test. Instead, in Figure 9, we report the L^1 norm of the error for APTD (left column) and MC (right column) for the different Knudsen numbers. These have been obtained by measuring the differences between the macroscopic quantities computed by the APTDNS and MC schemes with respect to a reference solution. The reference solution has been obtained by a second order in space deterministic scheme based on discrete velocity model version of the original kinetic equations. The results show that we have a reduction of the error compared to classical MC schemes for all the situations considered. In particular, the reduction of the error becomes larger as the fluid limit is approached.

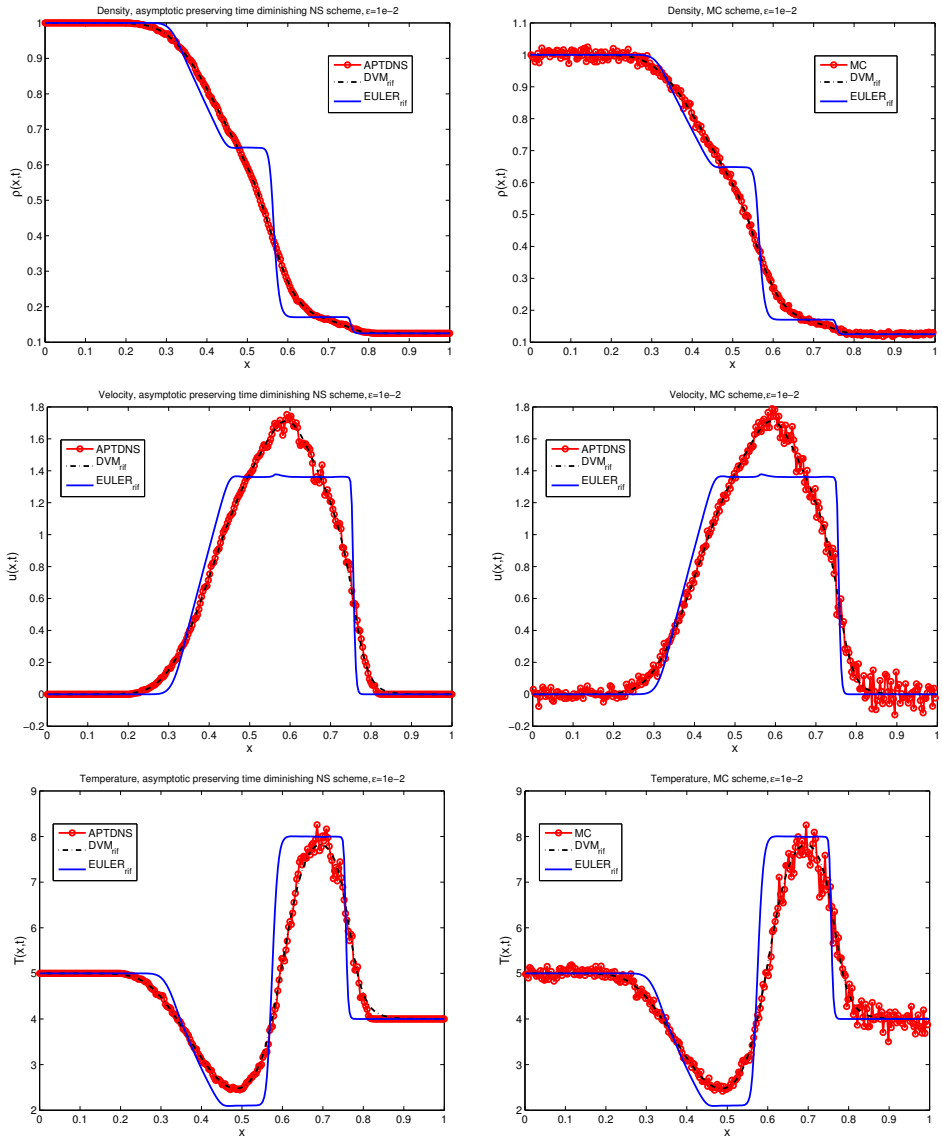

FIGURE 5. Density, velocity and temperature profiles (from top to bottom) for: left Asymptotic Preserving Time Diminishing Navier-Stokes method, right Monte Carlo method, $\varepsilon = 10^{-2}$. The dotted line is a reference solution computed with a deterministic DVM method while the continuous line is a reference solution for the compressible Euler equations. Sod test.

6. **Conclusion.** In this work, we have presented a new numerical method for the BGK equation which enjoys the following properties: (*i*) its statistical noise is smaller than the one of standard Monte Carlo methods; (*ii*) it is asymptotically stable with respect to the Knudsen number; (*iii*) its computational cost as well

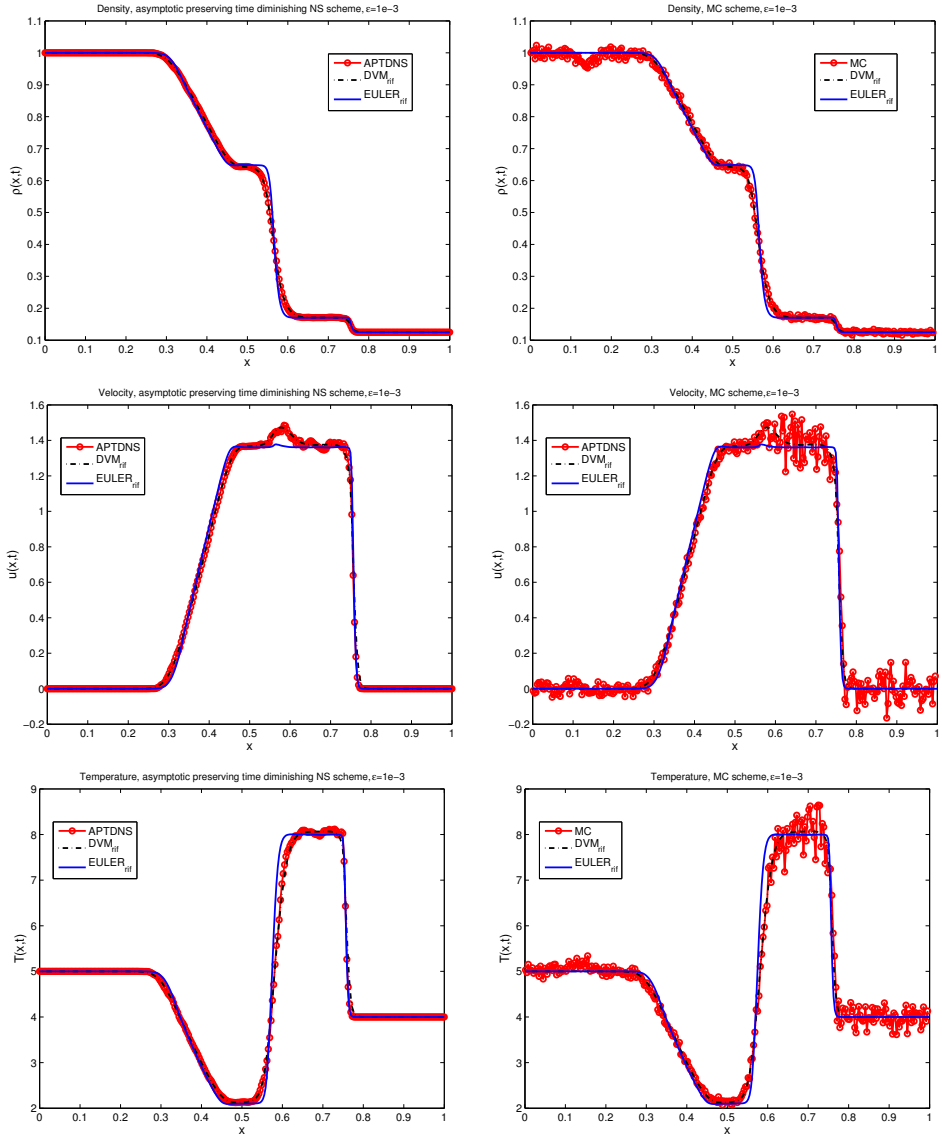

FIGURE 6. Density, velocity and temperature profiles (from top to bottom) for: left Asymptotic Preserving Time Diminishing Navier-Stokes method, right Monte Carlo method, $\varepsilon = 10^{-3}$. The dotted line is a reference solution computed with a deterministic DVM method while the continuous line is a reference solution for the compressible Euler equations. Sod test.

as its variance diminish as the equilibrium is approached; (*iv*) no artificial criteria is required. The method is based on a micro-macro decomposition (Euler-kinetic or Navier-Stokes-kinetic) for which the macro part is solved using a finite volume method whereas the micro part uses a Monte-Carlo method. This enables to derive

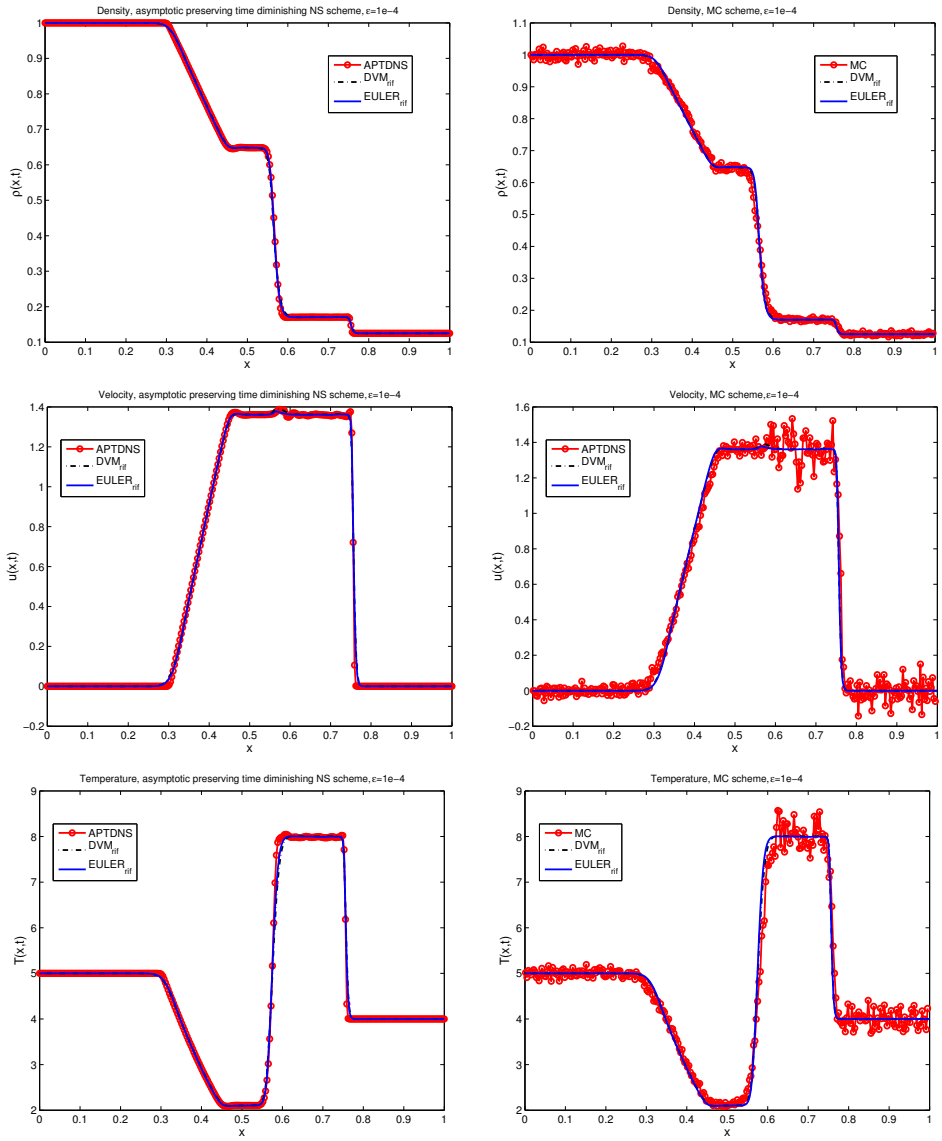

FIGURE 7. Density, velocity and temperature profiles (from top to bottom) for: left Asymptotic Preserving Time Diminishing Navier-Stokes method, right Monte Carlo method, $\varepsilon = 10^{-4}$. The dotted line is a reference solution computed with a deterministic DVM method while the continuous line is a reference solution for the compressible Euler equations. Sod test.

a low variance Monte Carlo based method for which additionally, the number of particles used to sample the micro unknown diminishes automatically as the fluid regime is approached. The numerical results illustrate the efficiency of the proposed method compared to the standard Monte Carlo approach.

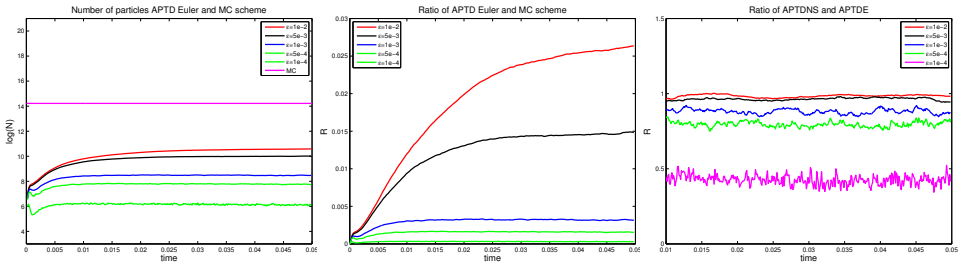

FIGURE 8. Left: time evolution of the number of effective particles (in semi-logarithmic scale) used in the Asymptotic Preserving Time Diminishing (Euler) method and in the Monte Carlo method for different values of the Knudsen number ($\varepsilon = 10^{-2}, 5\cdot 10^{-3}, 10^{-3}, 5\cdot 10^{-4}, 10^{-4}$). Middle: time evolution of the ratio of the number of particles used for APTD versus the number of particles for the corresponding MC simulation for different values of ε. Right: time evolution of the ratio of the number of particles used for APTD versus the number of particles for APTDNS for different values of ε. Sod test.

Appendix A. **Details of the equations 1D space - 1D velocity.** We use the notations $\mathscr{T} g = v\partial_x g$ and Π_M which is given by (8). In this part, we detail the calculations of $\Pi_M(\mathscr{T} g)$ and $(I - \Pi_M)(\mathscr{T} M)$.

Let us first compute $\Pi_M(v\partial_x g)$

$$\begin{aligned}
\Pi_M\left(v\partial_x g\right) &= \frac{M}{\rho}\left[\langle v\partial_x g\rangle + \frac{(v-u)\langle (v-u)v\partial_x g\rangle}{T}\right.\\
&\left. + \left(\frac{|v-u|^2}{2T} - \frac{1}{2}\right)\left\langle\left(\frac{|v-u|^2}{T} - 1\right) v\partial_x g\right\rangle\right]\\
&= \frac{M}{\rho}\left(\frac{|v-u|^2}{2T} - \frac{1}{2}\right)\left\langle\left(\frac{|v-u|^2}{T} - 1\right) v\partial_x g\right\rangle \qquad (34)\\
&= \frac{M}{\rho}\left(\frac{|v-u|^2}{2T} - \frac{1}{2}\right)\frac{1}{T}\left\langle v^3\partial_x g\right\rangle\\
&= \frac{M}{\rho}\left(\frac{|v-u|^2}{2T^2} - \frac{1}{2T}\right)\partial_x\left\langle v^3 g\right\rangle.
\end{aligned}$$

We can rewrite $\Pi_M \mathscr{T} g$ as a polynomial of degree 2 in the v variable times a Maxwellian

$$\Pi_M \mathscr{T} g(x,v) = M(x,v)\sum_{\ell=0}^{2} a_\ell(x) v^\ell, \qquad (35)$$

where a_ℓ depends on x through the moments ρ, u, T and the spatial derivatives of $\langle v^3 g\rangle$ (see the expression of (34)).

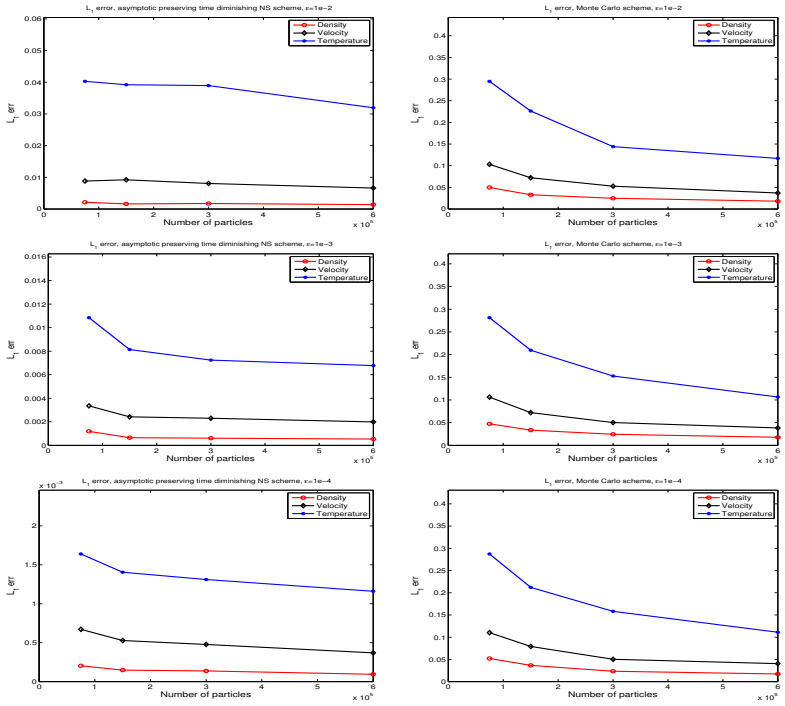

FIGURE 9. Error (L^1 norm) for the density, the mean velocity and the temperature for the Asymptotic Preserving Time Diminishing NS method (left column), and for the Monte Carlo method (right column) for different values of ε (from top to bottom, $\varepsilon = 10^{-2}, 10^{-3}, 10^{-4}$).

Then, we compute $\Pi_M(v\partial_x M)$:

$$\Pi_M\left(v\partial_x M\right) = \frac{M}{\rho}\left[\langle v\partial_x M\rangle + \frac{(v-u)\langle (v-u)v\partial_x M\rangle}{T} + \left(\frac{|v-u|^2}{2T} - \frac{1}{2}\right)\left\langle\left(\frac{|v-u|^2}{T} - 1\right)v\partial_x M\right\rangle\right].$$

We have:

$$\langle v\partial_x M\rangle = \partial_x\langle vM\rangle = \partial_x(\rho u),$$

$$\langle (v-u)v\partial_x M\rangle = \partial_x\int v^2 M\ dv - u\partial_x\int vM\ dv = \partial_x\left(\rho u^2 + \rho T\right) - u\partial_x(\rho u).$$

$$\left\langle\frac{|v-u|^2}{T}v\partial_x M\right\rangle = \frac{1}{T}\int(v^3 - 2uv^2 + u^2 v)\partial_x M\ dv = \frac{1}{T}\left(\partial_x\int v^3 M\ dv - 2u\partial_x\int v^2 M\ dv + u^2\partial_x(\rho u)\right),$$

where $\int v^3 M\ dv = 3\rho T u + \rho u^3$. So, we have:

$$\begin{aligned}\Pi_M\left(v\partial_x M\right) &= \frac{M}{\rho}\Big\{\partial_x(\rho u) + \frac{(v-u)}{T}\left[\partial_x\left(\rho u^2+\rho T\right) - u\partial_x(\rho u)\right]\\ &+ \left(\frac{|v-u|^2}{2T^2}-\frac{1}{2T}\right)\left[-u\partial_x\left(\rho u^2+\rho T\right)\right.\\ &+ \left.\left(\rho u^2+\rho T\right)\partial_x u + \left(u^2+2T\right)\partial_x(\rho u) + 2\rho u\partial_x T\right]\\ &- \left(\frac{|v-u|^2}{2T}-\frac{1}{2}\right)\partial_x(\rho u)\Big\}.\end{aligned} \tag{36}$$

Finally, $v\partial_x M$ is

$$v\partial_x M = v\left[\frac{\partial_x\rho}{\rho} + \frac{v-u}{T}\partial_x u + \left(\frac{(v-u)^2}{2T^2}-\frac{1}{2T}\right)\partial_x T\right] M. \tag{37}$$

Using (36) and (37), we then have the expression of $(I-\Pi_M)\mathscr{T}M$. We can rewrite $(I-\Pi_M)\mathscr{T}M$ as a polynomial of degree 3 in the v variable times a Maxwellian

$$(I-\Pi_M)\mathscr{T}M(x,v) = M(x,v)\sum_{\ell=0}^{3} b_\ell(x)v^\ell, \tag{38}$$

where b_ℓ depends on x through the moments ρ, u, T and their spatial derivatives $\partial_x\rho, \partial_x u, \partial_x T$ (see the expression of (36) and (37)).

REFERENCES

[1] H. Babovsky, *On a simulation scheme for the Boltzmann equation*, Math. Methods Appl. Sci., **8** (1986), pp. 223–233.

[2] C. Bardos, F. Golse, D. Levermore, *Fluid dynamic limits of kinetic equations I. Formal derivation*, J. Statist. Phys., **63** (1991), pp. 323–344.

[3] G. A. Bird, Molecular gas dynamics and direct simulation of gas flows, Clarendon Press, Oxford 1994.

[4] M. Bennoune, M. Lemou, L. Mieussens, *Uniformly stable numerical schemes for the Boltzmann equation preserving the compressible Navier–Stokes asymptotics*, J. Comput. Phys., **227** (2008), pp. 3781–3803.

[5] J. Burt, I. Boyd, *A hybrid particle approach for continuum and rarefied flow simulation*, J. Comput. Phys., **228** (2009), pp. 460–475.

[6] S. Brunner, E. Valeo, J. A. Krommes, *Collisional delta-f scheme with evolving background for transport time scale simulations*, Phys. of Plasmas, **12** (1999).

[7] S. Brunner, E. Valeo, J. A. Krommes, *Linear delta-f simulations of nonlocal electron heat transport*, Phys. of Plasmas, **7** (2000).

[8] R. Caflisch, *Monte Carlo and Quasi-Monte Carlo Methods*, Acta Numerica (1998), pp. 1–49.

[9] C. Cercignani, The Boltzmann equation and its applications, Springer-Verlag, New York, 1988.

[10] A. Crestetto, N. Crouseilles, M. Lemou, *Kinetic/fluid micro-macro numerical schemes for Vlasov-Poisson-BGK equation using particles*, Kin. Rel. Models, **5** (2012), pp. 787-816.

[11] N. Crouseilles, M. Lemou, *An asymptotic preserving scheme based on a micro-macro decomposition for collisional Vlasov equations: diffusion and high-field scaling limits*, Kin. Rel. Models, **4** (2011), pp. 441–477.

[12] P. Degond, G. Dimarco, L. Pareschi, *The moment guided Monte Carlo method*, International Journal for Numerical Methods in Fluids, **67** (2011), pp. 189–213,.

[13] P. Degond, G. Dimarco, *Fluid simulations with localized Boltzmann upscaling by direct simulation Monte-Carlo*, J. Comput. Phys., **231** (2012), pp. 2414–2437.

[14] P. Degond, S. Jin, L. Mieussens, *A smooth transition model between kinetic and hydrodynamic equations*, J. Comput. Phys., **209** (2005), pp. 665–694.

[15] G. Dimarco, *The hybrid moment guided Monte Carlo method for the Boltzmann equation*, Kin. Rel. Models, **6** (2013), pp. 291–315.

[16] G. Dimarco, L. Pareschi, *Hybrid multiscale methods II. Kinetic equations*, SIAM MMS, **6** (2007), pp.1169–1197.

[17] R. Caflisch, C. Wang, G. Dimarco, B. Cohen, A. Dimits, *A hybrid method for accelerated simulation of Coulomb collisions in a plasma*, SIAM MMS, **7** (2008), pp. 865–887.

[18] G. Dimarco, L. Pareschi, *A fluid solver independent hybrid method for multiscale kinetic equations*, SIAM J. Sci. Comput., **32** (2010), pp. 603–634.

[19] G. Dimarco, L. Pareschi, *Asymptotic preserving implicit-explicit Runge-Kutta methods for non linear kinetic equations*, SIAM J. Num. Anal., **32** (2013), pp. 1064–1087.

[20] G. Dimarco, L. Pareschi, *Numerical methods for kinetic equations*, Acta Numerica, **32** (2014), pp. 369–520.

[21] D. B. Hash and H. A. Hassan, *Assessment of schemes for coupling Monte Carlo and Navier-Stokes solution methods*, J. Thermophys. Heat Transf., **10** (1996), pp. 242–249.

[22] T. Homolle, N. Hadjiconstantinou, *A low-variance deviational simulation Monte Carlo for the Boltzmann equation.* J. Comp. Phys., **226** (2007), pp. 2341–2358.

[23] T. Homolle, N. Hadjiconstantinou, *Low-variance deviational simulation Monte Carlo.* Phys. Fluids, **19** (2007), 041701.

[24] S. Jin, *Efficient Asymptotic-Preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput., **21** (1999), pp. 441–454.

[25] M. Lemou, *Relaxed micro-macro schemes for kinetic equations*, Comptes Rendus Mathématique, **348** (2010), pp. 455–460.

[26] M. Lemou, L. Mieussens, *A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit*, SIAM J. Sci. Comp., **31** (2008), pp. 334–368.

[27] P. LeTallec and F. Mallinger, *Coupling Boltzmann and Navier-Stokes by half fluxes* J. Comput. Phys., **136** (1997), pp. 51–67.

[28] S. Liu, Monte Carlo strategies in scientific computing, Springer, 2004.
[29] K. Nanbu, *Direct simulation scheme derived from the Boltzmann equation*, J. Phys. Soc. Japan, **49** (1980), pp. 2042–2049.
[30] G. A. Radtke, J-P. M. Peraud, N. Hadjiconstantinou, *On efficient simulations of multiscale kinetic transport*, Phil. Trans. Royal Soc. A, **371** (2013), 2012182.
[31] S. Tiwari, A. Klar, S. Hardt, *A particle-particle hybrid method for kinetic and continuum equations*, J. Comput. Phys., **228** (2009), pp. 7109–7124.
[32] B. Yan, *A hybrid method with deviational particles for spatial inhomogeneous plasma*, J. Comput. Phys., **309** (2016), pp. 18–36.
[33] B. Yan, R. Caflisch, *A Monte Carlo method with negative particles for Coulomb collisions*, J. Comput. Phys., **298** (2015), pp. 711–740.

Acknowledgments. This work was supported by the french ANR project MOONRISE ANR-14-CE23-0007-01. N. Crouseilles and M. Lemou are supported by the Enabling Research EUROFusion project CfP-WP14-ER-01/IPP-03.

E-mail address: nicolas.crouseilles@inria.fr
E-mail address: giacomo.dimarco@unife.it
E-mail address: mohammed.lemou@univ-rennes1.fr