Asymptotic Preserving Monte Carlo Methods for the Boltzmann Equation*

Lorenzo Pareschi†
Giovanni Russo‡

Abstract

We consider some recently developed unconditionally stable numerical schemes for the Boltzmann equation, called Time Relaxed (TR) schemes. They share the important property of providing the correct fluid dynamic limit. Stability analysis of the schemes is performed, and the A-stability and L-stability of the schemes is studied. Monte Carlo methods based on TR discretizations are briefly reviewed. In particular, first and second order particle schemes are compared with a hybrid scheme, in which the equilibrium part of the distribution is described analytically.

Key words: Boltzmann equation, Implicit schemes, Monte Carlo method, Fluid dynamic limit.

1 Introduction

The Boltzmann equation gives a kinetic description of a rarefied gas, that is the gas is described in terms of a density function in phase space, $f(x, v, t)$, that gives the probability of finding a particle in a unit volume in phase space [4]. Macroscopic quantities, such as mass density, mean velocity, and temperature of the gas, can be obtained by computing the moments of the distribution function.

When a gas is near thermodynamical equilibrium, such detailed description is unnecessary, and the evolution of the moments can be described by Euler or Navier-Stokes equations. From a computational point of view, these equations are much less expensive than the full Boltzmann equation.

*This work was partially supported by TMR project "Asymptotic Methods in Kinetic Theory", Contract Number ERB FMRX CT97 0157, and partially by CNR, grant # 98.01030.CT01.

†Mathematics Department, University of Ferrara. Email: pareschi@dm.unife.it.

‡Mathematics Department, University of L'Aquila. Email: russo@univaq.it.

Far from thermodynamic equilibrium, the equation of gas dynamics do not give a satisfactory description of the physical system, and the full kinetic description is necessary.

The departure from local thermodynamic equilibrium can be measured by the so called Knudsen number, which is the ratio between the collisional mean free path of the molecules and the characteristic length of variation of macroscopic variables. When the Knudsen number is small, then the distribution function is close to the local Maxwellian. Collisions occur at a fast rate, and a kinetic treatment of the system is very expensive, because of the large ratio of time scales between macroscopic and microscopic effects. One may say that the system is numerically *stiff* [6]. Because of the stiffness, standard explicit schemes for the numerical solution of the Boltzmann equation become extremely expensive in this case.

If the Knudsen number is uniformly small in the region one is interested in, then one should use a gas dynamic description of the system. There are several circumstances, however, where the Knudsen number varies over several orders of magnitude. In such cases a standard kinetic treatment would be too expensive, and a pure gas dynamic approach inappropriate. Domain decomposition methods have been proposed for this problem, in which the computational domain is divided into an *aerodynamic region* in which the system is treated by gas dynamic equations (Euler or Navier-Stokes), and a *Boltzmann region*, where a kinetic treatment is used. Suitably boundary conditions make the two regions coupled together [3]. Such schemes can be very effective in treating large scale problems, but they are quite complicated to implement, and they require a great care by the user.

In some recently developed Monte Carlo methods a different approach is taken, with the goal of constructing simple and efficient numerical methods for the solution of the Boltzmann equation in regions with a large variation in the mean free path [12, 11].

These algorithms are based on a suitable time discretization of the Boltzmann equation, first introduced in [5], which has the following properties: for large Knudsen number it is a time discretization of the Boltzmann equation of a given order, and when the Knudsen number vanishes, the numerical solution tends to the local Maxwellian. We will call this last property *asymptotic preservation* property. In addition the time discretization preserves mass, momentum and energy. The stability property is similar to the L-stability of numerical schemes for systems of ode's [6].

Although this paper deals with the space homogeneous case, the real motivation for the development of these Monte Carlo methods is their use for spatially dependent problems. Because the distribution function approaches the local Maxwellian for small Knudsen numbers, a Monte Carlo method based on these time discretizations will naturally behave as a sort of kinetic schemes for the Euler equations in the fluid dynamic limit.

The plan of the paper is the following. In the next section we recall the time discretization and the asymptotic preserving schemes. In Section 3 we study the stability properties of these schemes and the relation to A- and L-stability. In section 4 Monte Carlo methods based on the asymptotic-preserving time discretization are reviewed. In the last section we compare the numerical results obtained by some of the Monte Carlo methods previously discussed.

2 Asymptotic preserving schemes

The space homogeneous Boltzmann equation

Let us consider the homogeneous Boltzmann equation [4]

$$\frac{\partial f}{\partial t} = \frac{1}{\varepsilon} Q(f, f), \tag{1}$$

supplemented with the initial condition

$$f(v, t=0) = f_0(v), \tag{2}$$

where $f = f(v,t)$ is a non negative function describing the time evolution of the distribution of particles which move with velocity v at time $t > 0$. The parameter $\varepsilon > 0$ is called the Knudsen number and is proportional to the mean free path between collisions. The bilinear collision operator $Q(f, f)$ describes the binary collisions of the particles and is given by

$$Q(f, f) = Q_+(f, f) - \nu[f]f, \tag{3}$$

where

$$Q_+(f, f)(v) = \int_{\mathbb{R}^3} \int_{S^2} \sigma(|v - v_*|, \omega) f(v') f(v'_*)\, d\omega\, dv_*, \tag{4}$$

represents the gain term, and

$$\nu[f](v) = \int_{\mathbb{R}^3} \int_{S^2} \sigma(|v - v_*|, \omega) f(v_*)]\, d\omega\, dv_* \tag{5}$$

is the loss term. In the above expression, ω is a unit vector of the sphere S^2, so that $d\omega$ is an element of area of the surface of the unit sphere S^2 in $\mathbb{R}^3$. Moreover (v', v'_*) represent the post-collisional velocities associated with the pre-collisional velocities (v, v_*) and the collision parameter ω; i.e.,

$$v' = \frac{1}{2}(v + v_* + |v - v_*|\omega), \quad v'_* = \frac{1}{2}(v + v_* - |v - v_*|\omega). \tag{6}$$

The kernel σ is a nonnegative function which characterizes the details of the binary interactions. In the case of inverse k-th power forces between particles the kernel has the form

$$\sigma(|v - v_*|, \theta) = b_\alpha(\theta)|v - v_*|^\alpha, \tag{7}$$

where $\alpha = (k-5)/(k-1)$. For numerical purposes, a widely used model is the Variable Hard Sphere(VHS) model [2], corresponding to take $b_\alpha(\theta) = C_\alpha$ where C_α is a positive constant. The case $\alpha = 0$ is referred to as Maxwellian gas whereas the case $\alpha = 1$ yields the Hard Sphere gas.

During the evolution process, mass, momentum and energy are conserved; i.e.,

$$\int_{I\!R^3} Q(f,f)\phi(v)\,dv = 0, \quad \phi(v) = 1, v, v^2, \tag{8}$$

The equilibrium distribution function satisfies the equation $Q(f,f) = 0$, and has the form of a locally Maxwellian distribution

$$M(\rho, u, T)(v) = \frac{\rho}{(2\pi T)^{3/2}} \exp(-\frac{|u-v|^2}{2T}), \tag{9}$$

where ρ, u, T are the density, mean velocity and temperature of the gas.

Time relaxed (TR) schemes

As proposed in [5], a general idea for deriving unconditionally stable and asymptotic preserving numerical schemes for a nonlinear equation like (1) is to replace high order terms of a suitable well-posed power series expansion by the local equilibrium. This class of schemes has been called *Time Relaxed* (TR) schemes.

Here we will briefly recall the schemes presented in [5] and some generalization used in [12].

Derivation

Let us consider a differential system of the type

$$\frac{\partial f}{\partial t} = \frac{1}{\varepsilon}\left[P(f,f) - \mu f\right] \tag{10}$$

with the initial condition

$$f(0) = f_0 \tag{11}$$

where $\mu \geq 0$ is a constant and P a bilinear operator.

Then the function f can be expressed by the following formal expansion (Wild sum) [15, 5]

$$f(t) = e^{-\mu t/\varepsilon} \sum_{k=0}^{\infty} \left(1 - e^{-\mu t/\varepsilon}\right)^k f_k, \tag{12}$$

where the functions f_k are given by the recurrent formula

$$f_{k+1} = \frac{1}{k+1} \sum_{h=0}^{k} \frac{1}{\mu} P(f_h, f_{k-h}), \quad k = 0, 1, \ldots \tag{13}$$

From this representation, a class of numerical schemes can be naturally derived. Let us introduce the scaled time $\tau = 1 - \exp(-\mu t/\varepsilon)$.

In [5] the following class of numerical schemes, based on a suitable truncation for $m \geq 1$ of (12), has been constructed

$$f^{n+1}(v) = e^{-\mu\Delta t/\varepsilon} \sum_{k=0}^{m} \left(1 - e^{-\mu\Delta t/\varepsilon}\right)^k f_k^n(v) + \left(1 - e^{-\mu\Delta t/\varepsilon}\right)^{m+1} M(v), \qquad (14)$$

where $f^n = f(n\Delta t)$ and Δt is a small time interval. The quantity M (referred as the local Maxwellian associated to f) is the asymptotic stationary solution of the equation.

It can be shown that the schemes obtained in this way are of order m in time. Furthermore, they satisfy the following properties:

i) *conservation:*

If $P(f, g)$ is a non negative bilinear operator such that there exist some functions $\phi(v)$ with the following property

$$\int_{R^3} P(f, f)\phi(v)\, dv = \mu \int_{R^3} f\phi(v)\, dv, \qquad (15)$$

and the initial condition f^0 is a non negative function, then f^{n+1} is nonnegative for any $\mu\Delta t/\varepsilon$, and satisfies

$$\int_{R^3} f^{n+1}\phi(v)\, dv = \int_{R^3} f^n\phi(v)\, dv; \qquad (16)$$

ii) *asymptotic preservation:*

For any $m \geq 1$, we have

$$\lim_{\mu\Delta t/\varepsilon \to \infty} f^{n+1} = M(v). \qquad (17)$$

When applied to the Boltzmann equation, the first property guarantees exact conservation of mass, momentum, and energy, for the semi-discrete scheme, while the second property guarantees the correct fluid dynamic limit.

Generalization

The approach we have seen in the previous sections can be generalized using different weight functions to combine the influence of the high order coefficients appearing in the Wild sum (12). In general such schemes can be written as

$$f^{n+1} = \sum_{k=0}^{m} A_k f_k + A_{m+1} M, \qquad (18)$$

where the coefficients f_k are given by Eq. (13). The weights $A_k = A_k(\tau)$ are nonnegative function that satisfy the following properties

i) *consistency:*

$$\lim_{\tau\to 0} A_1(\tau)/\tau = 1, \quad \lim_{\tau\to 0} A_k(\tau)/\tau = 0, \quad k = 2,\dots,m+1 \tag{19}$$

ii) *conservation:*

$$\sum_{k=0}^{m+1} A_k = 1 \quad \tau \in [0,1], \tag{20}$$

iii) *asymptotic preserving:*

$$\lim_{\tau\to 1} A_k(\tau) = 0, \quad k = 0,\dots,m \tag{21}$$

Such properties are a generalization of those defined in [12] for first and second order schemes. It is possible to show that a scheme that satisfies these properties will satisfy properties (16) and (17).

A choice of functions which satisfies the previous requirements is given by

$$A_k = (1-\tau)\tau^k,\ k = 0,\dots,m, \quad A_{k+1} = \tau^{m+1}, \tag{22}$$

which correspond to the scheme (14).

However, other choices are possible and it is an open problem the determination of the *optimal* set of functions A_k that satisfies the previous requirements and guarantees the most accurate approximation. Some preliminary results in this direction will be shown in the next section.

3 Analysis of the TR schemes

Linear stability

In this section we perform a stability analysis of the TR schemes. In particular, we show that the first and second order schemes used in [5] are L-stable.

Let us apply the TR scheme to the test equation

$$\left\{ \begin{array}{rcl} y' &=& \lambda y \\ y(0) &=& 1 \end{array} \right. \tag{23}$$

with $\lambda \in C^-$. The equation can be written in the form

$$y' = P(y,y) - \mu y$$

with $P(f,g) = \frac{1}{2}(\mu+\lambda)(f+g)$, and μ is a positive constant such that $|\lambda| < \mu$. The numerical solution after one time step is given by

$$y^1 = \sum_{k=0}^{m} A_k y_k \tag{24}$$

The $m+1$ term vanishes because the equilibrium solution is zero. The term y_k can be explicitly computed by applying the formula (13). A simple calculation yields

$$y_k = \frac{1}{k!}\prod_{j=1}^{k}(j+a) = \frac{\Gamma(k+1+a)}{\Gamma(k+1)\,\Gamma(1+a)}, \quad k=0,\ldots,m \tag{25}$$

where $a=\lambda/\mu$. The numerical solution after one time step, called the *function of absolute stability*, is therefore a function of a real parameter τ and a complex parameter a

$$R_m(\tau,a) = \sum_{k=0}^{m} A_k(\tau) y_k(a),$$

where $\tau = 1-\exp(-\mu\Delta t)$ is the relaxed time, and a is a complex number $a \in C^-, |a| \leq 1$. Here we give definitions of A- and L-stability which are the natural generalization of the standard definition [6] to the case where the function of absolute stability depends on an additional real parameter.

Definition 3.1 (A-stability) *We say that the scheme is A-stable if*

$$|R(\tau,a)| \leq 1 \quad \forall \tau \in [0,1],\ a \in C^-, |a| \leq 1$$

Definition 3.2 (L-Stability) *We say that the scheme is L-stable if it is A-stable and*

$$\lim_{\tau\to 1} R(\tau,a) = 0, \quad \forall a \in C^-, |a| \leq 1$$

Definition 3.3 ($A\alpha$-stability) *We say that the scheme is $A\alpha$-stable if*

$$|R(\tau,a)| \leq 1 \quad \forall \tau \in [0,1],\ a = -\rho\exp(i\theta),\ \rho \leq 1, \theta \in [-\alpha,\alpha]$$

Definition 3.4 (Lα-Stability) *We say that the scheme is $L\alpha$-stable if it is $A\alpha$-stable and*

$$\lim_{\tau\to 1} R(\tau,a) = 0 \quad \forall a = -\rho\exp(i\theta),\ \rho \leq 1, \theta \in [-\alpha,\alpha]$$

It is clear that, because of the asymptotic preserving properties, if a TR scheme is A-stable it is also L-stable.

Here we prove that the first and second order schemes (14) are in fact A-stable (and therefore L-stable).

We conjecture that the A-stability property is shared by schemes (14) of all order, but we have not been able to prove it.

First order schemes Let us consider the first order scheme given by (22, 24) for $m = 1$

$$y^1 = (1 - \tau)y_0 + \tau(1 - \tau)y_1. \tag{26}$$

The function of absolute stability in this case is given by

$$R_1(\tau, a) = (1 - \tau)[1 + \tau(1 + a)].$$

We prove the following

Theorem 3.1 *The scheme (26) is A-stable according to Def. 3.1, i.e.*

$$|R_1(\tau, a)| \leq 1, \forall \tau \in [0, 1], \ a \in C^-, |a| \leq 1.$$

Proof

Let $a = -\rho \exp(i\theta)$, $\theta \in [-\pi/2, \pi/2]$, $\rho \in [0, 1]$, *and let*

$$F_1(\tau, \rho, \theta) = 1 - |R_1(\tau, -\rho e^{i\theta})|^2.$$

Direct computation shows that

$$F_1 = A + B\cos\theta$$

with $A = \tau^2[(1 - \tau^2) + 1 - \rho^2(1 - \tau)^2]$, $B = 2\rho\tau(1 + \tau)(1 - \tau)^2$

Because of the range of the variables, it is $F_1 \geq A \geq 0$, *and therefore* $|R_1(\tau, a)| \leq 1$.

■

Second order schemes Now we consider the second order scheme given by (22, 24) with $m = 2$,

$$y^1 = (1 - \tau)[y_0 + \tau y_1 + \tau^2 y_2]. \tag{27}$$

The function of absolute stability in this case is given by

$$R_2(\tau, a) = (1 - \tau)[1 + \tau(1 + a) + \frac{1}{2}\tau^2(1 + a)(2 + a)].$$

Let $\Omega_a \equiv \{z \in C : \Re z \leq 0, |z| \leq 1\}$. We observe that $R_2(\tau, a)$ is analytic in a, and therefore, for any fixed value of the parameter $\tau \in [0, 1]$, the function takes the maximum absolute value on the border Γ_a of Ω_a. Using this property, we can prove the following

Theorem 3.2 *The scheme (27) is A-stable according to Def. 3.1, i.e.*

$$|R_2(\tau, a)| \leq 1, \forall \tau \in [0, 1], a \in \Omega_a \equiv \{z \in C : \Re z \leq 0, |z| \leq 1\}.$$

Proof

Let $a = -\rho\exp(i\theta)$. *Then* $\Gamma_a = \partial\Omega_a = \Gamma_1 \cup \Gamma_2$, *where* $\Gamma_1 = \{z \in C : z = iy, y \in [-1,1]\}$, *and* $\Gamma_2 = \{z \in C : z = -e^{i\theta}, \theta \in [-\pi/2, \pi/2]\}$.

Direct calculation shows that

$$1 - |R_2(\tau, a)|^2 = \frac{1}{4}\tau F_2,$$

with

$$\begin{aligned} F_2(\tau, -\rho e^{i\theta}) &= -8\tau\rho^2(1+\tau)(1-\tau^3)\cos^2\theta \\ &+ 2\rho(3\tau+2)(1-\tau)^2(\rho^2\tau^2 + 2\tau^2 + 2\tau + 2)\cos\theta \\ &+ \tau^2[8 - 4\tau^3 - (5\tau+8)(1-\tau)^2\rho^2 - \tau(1-\tau)^2\rho^4] \end{aligned}$$

The restriction of F_2 *to* Γ_1 *is*

$$F_2(\tau, iy) = (A\xi^2 + B\xi + C),$$

where

$$\xi \equiv y^2,\ A = -\tau^3(1-\tau)^2,\ B = -\tau^2(5\tau+8)(1-\tau)^2,\ C = (8 - 4\tau^3)\tau^2.$$

It is $F_2(\tau, 0) = \tau^2(8 - 4\tau^3) \geq 0$, $F_2(\tau, i) = \tau^3(5 + 2\tau - 5\tau^2) \geq 0$,*and* $A \leq 0$, *therefore* $F_2(\tau, iy) \geq 0, \forall y \in [-1,1]$.

The restriction of F_2 *to* Γ_2 *is*

$$F_2(\tau, e^{i\theta}) = \tilde{F}(\tau, x) \equiv \tilde{A}x^2 + \tilde{B}x + \tilde{C},$$

where

$$x = \cos\theta, \quad \tilde{A} = -8\tau(1+\tau)(1-\tau)^3,$$

$$\tilde{B} = 2(3\tau+2)(1-\tau^2)(3\tau^2 + 2\tau + 2), \quad \tilde{C} = 2\tau^3(5 + 2\tau - 5\tau^2).$$

It is $\tilde{F}(\tau, 0) \geq 0$, $\tilde{F}(\tau, 1) = 4 - 2\tau > 0$, *and* $A \leq 0$, *therefore* $\tilde{F}(\tau, x) \geq 0, \forall \tau \in [0,1]$, $\theta \in [-\pi/2, \pi/2]$.

■

Accuracy consideration

As we said in the previous section, the choice of the weights (14) is not necessarily the best for numerical computations. In fact, numerical experiments [12] suggest that other choices give better results. Let us consider the family of schemes defined by the following coefficients

$$A_k = (1-\tau)\tau^k,\ k = 0, m-1, \quad A_m = (1 - \tau^{r+1})\tau^m, \quad A_{m+1} = \tau^{m+r+1}, \tag{28}$$

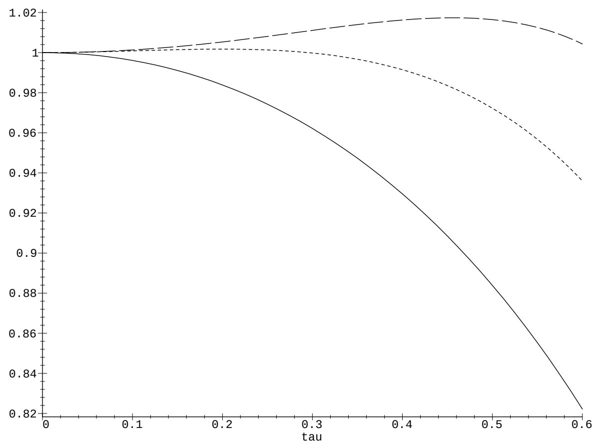

Figure 1: $R(\tau, -1/2)/\exp(-\mu\Delta t/2)$ vs τ. First order schemes. Continuous line: $r = 0$, dotted line: $r = 1$, dashed line: $r = 2$

where $r = 0, 1, \ldots$. Schemes (14) are a particular member of this family, and correspond to the choice $r = 0$. Schemes with $r \geq 1$ are more accurate than the schemes corresponding to $r = 0$ for small time steps. This can be seen by computing the leading error term in the Taylor expansion of the difference between the numerical solution and the exact solution. The constant in front the term Δt^{m+1} is in fact smaller for $r \geq 1$ than for $r = 0$.

A good choice seems $r = 1$. This is suggested by the comparison between the exact solution and the function $R_m(\tau, a)$ for real values of $a \in [-1, 0]$.

In Fig. (1) we show the ratio $R_1(\tau, -1/2)/e^{-\mu t/2}$, as a function of the scaled time τ, for the first order schemes, corresponding to $r = 0, 1, 2$. In Fig. (2) the same ratio are shown, for second order schemes. It is evident that for $r = 1$ the ratio is very close to 1, even for considerable values of the parameter τ.

We remark here that for this more general schemes, A-stability is not guaranteed, as can be explicitly checked for the first order scheme with $r = 1$. It can be shown, however, that this scheme is A-α stable, with $\alpha \approx \pi/2 - 0.05$.

4 Time Relaxed Monte Carlo (TRMC) methods

In this section we describe the *Time Relaxed Monte Carlo* (TRMC) methods for the evolution of the density function f.

Monte Carlo methods are naturally derived from the probabilistic interpretation of

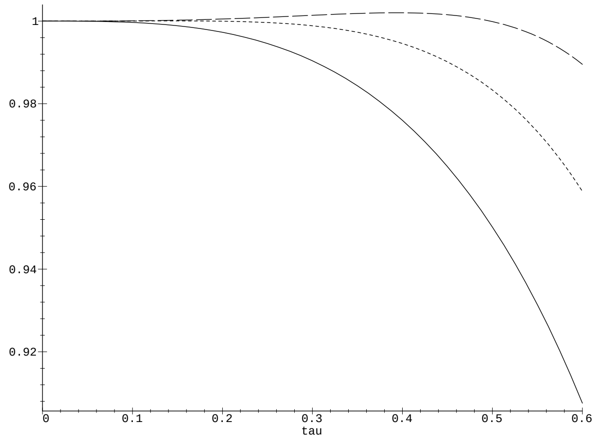

Figure 2: $R(\tau, -1/2)/\exp(-\mu\Delta t/2)$ vs τ. Second order schemes. Continuous line: $r = 0$, dotted line: $r = 1$, dashed line: $r = 2$

time discrete schemes. Since we deal only with the space homogeneous case, we shall make the simplifying assumption that the function $f^n(v)$ is a probability distribution, i.e.

$$\int_{R^3} f^n(v)\, dv = 1.$$

Now we consider a popular DSMC scheme, which can be considered a prototype of MC schemes for the Boltzmann equation. The scheme is based on an idea by Nanbu, and a subsequent modification by Babovsky, who improved the original scheme by making it conservative and with a complexity $\mathcal{O}(N)$ per time step (see [9], [1].) Here we show that such scheme can be written in a form sililar to the one used for the development of Wild sums and TR schemes.

Let us rewrite the Boltzmann equation (1) in the form

$$\frac{\partial f}{\partial t} = P(f,f) - \mu f, \tag{29}$$

where μ is a positive constant such that

$$\mu \geq \nu(v) \equiv \int_{\mathbb{R}^3} \int_{S^2} \sigma(|v - v_1|, \omega) f(v_*)\, d\omega\, dv_*,$$

and $P(f,f) = Q_+(f,f) + (\mu - \nu(v))f$. Explicit Euler scheme applied to Eq. (29) is given by

$$f^{n+1} = (1 - \mu\Delta t) f^n + \mu\Delta t f_1 \tag{30}$$

where $f_1 = P(f^n, f^n)/\mu$. Direct Simulation Monte Carlo is obtained from the probabilistic interpretation of Eq. (30).

A particle extracted from f^n:

- does not collide with probability $1 - \mu\Delta t$
- undergoes a *dummy collision* with another particle sampled from f^n with probability $\mu\Delta t$ (i.e. it is sampled from the function f_1).

Here

$$f_1 = \left(1 - \frac{\nu(v)}{\mu}\right) f^n + \frac{Q_+(f^n, f^n)}{\mu}$$

The probabilistic interpretation of the above equation is obvious. The fraction

$$1 - \frac{\int_{R^3} \nu(v)\, dv}{\mu}$$

represents the average probability that a dummy collision is rejected.

Note that the probability interpretation of the scheme is valid as long as $\mu\Delta t \leq 1$. This corresponds to the stability restriction of the explicit Euler scheme.

Before we describe TR schemes, a remark is in order. We observe that in actual implementations, one choses a given number of *dummy collision pairs*, equal to $\mu N\Delta t/2$. Such pairs are sampled uniformly among all possible pairs. The parameter μ is obtained by an estimate σ_m of the maximum cross section between all particle pairs. Then dummy collisions are performed, and each of them is accepted with a probability proportional to the ratio $\sigma(v_i, v_j)/\sigma_m$. If we do not had to compute σ_m, then the computational cost of such Monte Carlo method would be independend on the time step, since its cost would be proportional to the number of collisions, which is basically independent on the time step, for small time steps. This is the case, for example, of Maxwell molecules. However, if we compute σ_m at every time step, then the cost of the computation increases when we decrease the time step, since the estimate of σ_m has a cost $\mathcal{O}(N)$ and not $\mathcal{O}(N\Delta t)$.

First order TR schemes are written as

$$f^{n+1} = A_0 f^n + A_1 f_1 + A_2 \tag{31}$$

The probabilistic interpretation of the above equation is the following.

A particle extracted from f^n

- does not collide with probability A_0,
- collides with another particle extracted from f^n with probability A_1,
- is replaced by a particle sampled from a Maxwellian with probability A_2.

In this formulation the probabilistic interpretation holds uniformly in $\mu\Delta t$, at variance with standard DSMC, which requires $\mu\Delta t < 1$. Furthermore, as $\mu\Delta t \to \infty$, the distribution at time $n+1$ is sampled from a Maxwellian. In this limit, the density f^{n+1} relaxes immediately to its equilibrium distribution. In a space non homogeneous case, this would be equivalent to the particle method for Euler equations proposed by Pullin [13].

A second order Monte Carlo scheme is obtained by the TR scheme

$$f^{n+1} = A_0 f^n + A_1 f_1 + A_2 f_2 + A_3 M, \tag{32}$$

with

$$f_1 = \frac{P(f^n, f^n)}{\mu}, \quad f_2 = \frac{P(f^n, f_1)}{\mu}.$$

The probabilistic interpretation of scheme (32) is the following. Given N particles distributed according to f^n:

- NA_0 particles do not collide,
- NA_1 are sampled from f_1, as in the first order scheme,
- NA_2 are sampled from f_2, i.e. $NA_2/2$ particles sampled from f^n will undergo dummy collisions with $NA_2/2$ particles sampled from f_1,
- NA_3 particles are sampled from a Maxwellian.

The Monte Carlo schemes described above are conservative in the mean. It is possible to make them exactly conservative [12]. This goal is achieved by

- selecting collision pairs uniformly, rather than individual particles,
- using a suitable algorithm for sampling a set of particles with prescribed momentum and energy from a Maxwellian [14].

We remark here that the TR scheme is a direct consequence of the time discretization of the Boltzmann equation. In this respect it is not crucial to select a particular Monte Carlo scheme in order to perform the comparison. Here we considered a commonly used Monte Carlo scheme. If a different technique is used to model the collision, then the same technique could be used in conjunction with TR disretization.

The hybrid distribution: TRMCH scheme The variance of the Monte Carlo solution can be considerably reduced if one represents the equilibrium part of the distribution function analytically by a Maxwellian. Such idea has been recently presented in the paper [11]. One makes the *ansatz* that the distribution function is the sum of a continuous and a discrete part,

$$f^n(v) = (1-\beta^n) g^n(v) + \beta^n M(v).$$

By substituting this expression in the first order scheme (31), one obtains the evolution equations for β^n and g^n

$$\beta^{n+1} = A_0\beta^n + A_1(\beta^n)^2 + A_2 \tag{33}$$
$$g^{n+1} = p_1 g^n + p_2\left[q_1\frac{P(g^n,g^n)}{\mu} + q_2\frac{P(g^n,M)}{\mu}\right], \tag{34}$$

where

$$p_1 = \frac{A_0}{A_0 + A_1(1+\beta^n)}, \quad p_2 = \frac{A_1(1+\beta_n)}{A_0 + A_1(1+\beta_n)}, \tag{35}$$
$$q_1 = \frac{1-\beta^n}{1+\beta^n}, \quad q_2 = \frac{2\beta^n}{1+\beta^n}. \tag{36}$$

Note that $p_1, p_2 \geq 0$, $p_1 + p_2 = 1$, $q_1, q_2 \geq 0$, $q_1 + q_2 = 1$.

The probabilistic interpretation of the evolution equation for g^n is similar to the one for the purely particle scheme. Note that with this approach the number of particles is a function of time (it is $N^n \approx N^0(1-\beta^n)$).

The advantage of this approach with respect to the purely particle method is twofold: the variance of the distribution is reduced because a fraction of the distribution function is represented analytically, and the scheme is more efficient, because the number of particles is reduced.

5 Numerical Examples

In this section we test the different TRMC methods presented in the previous section by comparing them with the standard DSMC. Since we are interested in checking the accuracy of the schemes we consider two initial value problems, respectively for the Kac equation and for the Boltzmann equation for Maxwell molecules, for which we have an exact solution. To simplify notations we denote by TRMC1, TRMC2 and TRMCH respectively the first order and second order TRMC methods and the hybrid TRMC method. In our tests we use the family of schemes characterized by (28) with $r = 1.$ and perform a single run, with a number of particles sufficiently large to control the effects of the fluctuations. We express the results as a function of the scaled time variable t/ϵ which we denote again by t.

First we compare the Monte Carlo solutions with the exact solution of the Kac equation [8]

$$f(v,t) = \frac{1}{2}\left[\frac{3}{2}(1-C(t))\sqrt{C(t)} + (3C(t)-1)C(t)^{3/2}v^2\right]e^{-C(t)v^2},$$

where

$$C(t) = \left[3 - 2e^{-\sqrt{\pi}t/16}\right]^{-1}.$$

We reconstruct the density function on a grid, by convolving the particle distribution by a suitable mollifier [11, 12].

All the simulations have been performed for $t \in [0, 6]$ by starting with $N = 5 \times 10^4$ particles.

In Fig. 3 we report the L^2-norm of the error obtained with the different schemes using the same time step $\Delta t = 0.5$.

It is evident that all the TRMC methods provide a more accurate representation of the solution. In particular, by comparing the solution obtained with DSMC and TRMC1 one can see the improvement given by the first order TR scheme with respect to the explicit Euler time discretization. It is remarkable that the TRMCH method, as a consequence of its deterministic evolution of the equilibrium part, yields the most accurate solution.

Fig. 4 shows the L^2-norm of the error for the same test where the time step of the TRMC schemes are respectively two times (TRMC1), three times (TRMC2) and four times (TRMCH) larger then the one used by DSMC. Note that in spite of the different time steps the accuracy of the different schemes is similar.

Next we consider the numerical solution of the two-dimensional homogeneous Boltzmann equation for Maxwell molecules.

An exact solution for this equation is given by

$$f(v,t) = \frac{1}{2\pi C(t)}\left[1 - \frac{1}{C(t)}(1 - C(t))\left(1 - \frac{v^2}{2C(t)}\right)\right] e^{-v^2/2C(t)},$$

where $C(t) = 1 - (1/2)e^{-t/8}$.

The density distribution is obtained by reconstructing the function on a regular grid of spacing $\Delta v = 0.25$ by the "weighted area rule" [7] and the simulations have been performed for $t \in [0, 16]$ by starting with $N = 10^5$ particles.

In Fig. 5, 6 we show the time evolution of the fourth order moment of the solution. The results confirm the gain of accuracy and the reduction of fluctuations of the TRMC methods with respect to the DSMC method for larger time steps.

We conclude this section with some additional considerations on the efficiency of TRMC methods.

First, in spite of the fact that the time step for TRMC methods can be chosen larger than that of DSMC, the number of dummy collision is higher for DSMC. The reason is that this number is proportional to $\mu\Delta t$ for DSMC, and it is proportional to $1-\exp(-\mu\Delta t)$ for TRMC. This is a further advantage of TRMC methods with respect to DSMC.

Second, we emphasize that TRMC2 takes into account re-collisions of particles during the same time step. Thus, although the direct combination of the previous schemes with a solver for the transport step gives a scheme which is still first order, we expect that the overall accuracy of the solution will be improved also in space non homogeneous problems. Furthermore, it is well known that a suitable combination of collision and convection steps may lead to second order accurate schemes both in space and time steps (an application to deterministic methods for the Boltzmann equation is given for example, in [10]). In a future work, we hope to construct a second order TRMC scheme for non homogeneous Boltzmann equation that works uniformly with respect to the Knudsen number.

A similar effect is more evident in the homogeneous test problems we have presented here, where the accuracy of the virtually first order TRMCH is higher than the accuracy of the second order TRMC2. In fact, the reduction of fluctuation given by the deterministic evolution of the Maxwellian part is stronger than the reduction of fluctuation given by the re-collision of particles.

References

[1] Hans Babovsky. On a simulation scheme for the Boltzmann equation. *Mathematical Methods in the Applied Sciences*, 8:223–233, 1986.

[2] G.A. Bird. *Molecular Gas Dynamics.* Oxford University Press, London, 1976.

[3] J.F Bourgat, P. LeTallec, B. Perthame, and Y. Qiu. Coupling Boltzmann and Euler equations without overlapping. In *Domain Decomposition.* AMS, 1992.

[4] C. Cercignani. *The Boltzmann Equation and its Applications.* Springer-Verlag, 1988.

[5] E. Gabetta, L. Pareschi, and G. Toscani. Relaxation schemes for nonlinear kinetic equations. *SIAM Journal on Numerical Analysis*, 34:2168–2194, 1997.

[6] E. Hairer and G. Wanner. *Solving ordinary differential equations, Vol.2 Stiff and differential-algebraic problems.* Springer-Verlag, New York, 1987.

[7] R.W. Hockney and J.W. Eastwood. *Computer simulation using particles.* Mc-Graw Hill International Book Co., 1981.

[8] M. Kac. Probability and related topics in physical sciences. In *Lectures in Applied Mathematics.* Interscience Publishers London-New York, 1957.

[9] K. Nanbu. Theoretical basis of the direct simulation Monte Carlo method. In V. Boffi and C. Cercignani, editors, *Proceedings of the 15th International Symposium on Rarefied Gas Dynamics*, pages 369–383, 1986.

[10] T. Ohwada. Higher order approximation methods for the boltzmann equation. *Journal of Computational Physics*, 139:1–14, 1998.

[11] L. Pareschi and R.E. Caflisch. Implicit monte carlo methods for rarefied gas dynamics I: the space homogeneous case. *Journal of Computational Physics*, to appear.

[12] L. Pareschi and G. Russo. Time relaxed monte carlo method for the boltzmann equation. *preprint.*

[13] D.I. Pullin. Direct simulation methods for compressible inviscid ideal gas flow. *Journal of Computational Physics*, 1978.

[14] D.I. Pullin. Generation of normal variates with given sample. *J. Statist. Comput. Simul.*, 9:303–309, 1979.

[15] E. Wild. On Boltzmann's equation in the kinetic theory of gases. *Proc. Camb. Phil. Soc.*, 47:602–609, 1951.

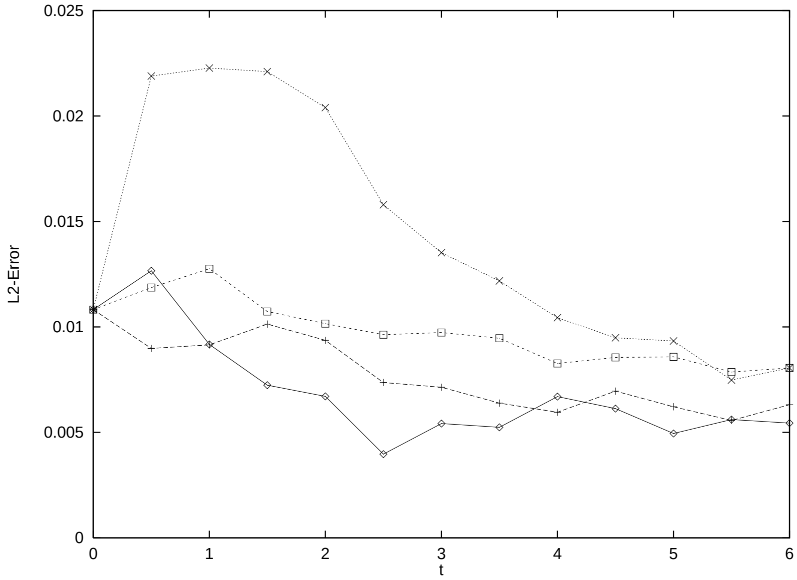

Figure 3: Kac equation : L^2 error vs time with $\Delta t = 0.5$ for DSMC ($\times$), TRMC1 ($\Box$), TRMC2 ($+$) and TRMCH ($\diamond$).

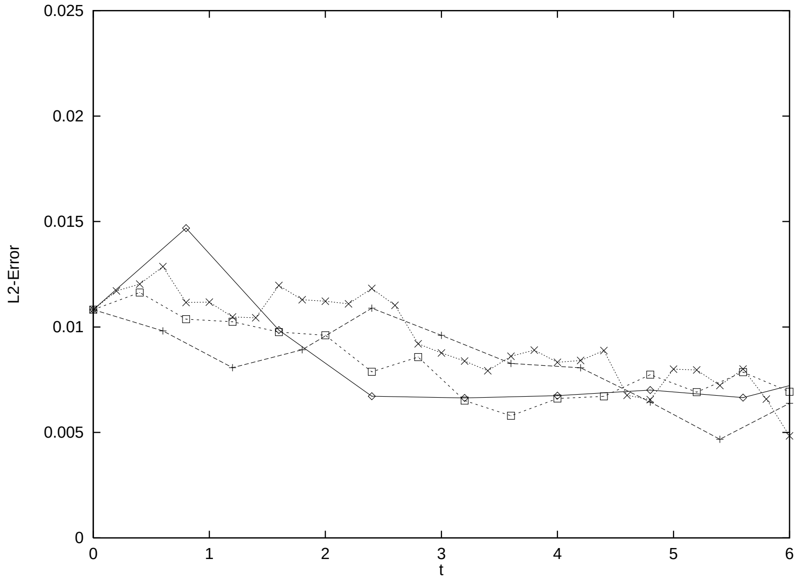

Figure 4: Kac equation : L^2 error vs time with $\Delta t = 0.2$ for DSMC ($\times$), $\Delta t = 0.4$ for TRMC1 ($\Box$), $\Delta t = 0.6$ for TRMC2 ($+$) and $\Delta t = 0.8$ for TRMCH ($\diamond$).

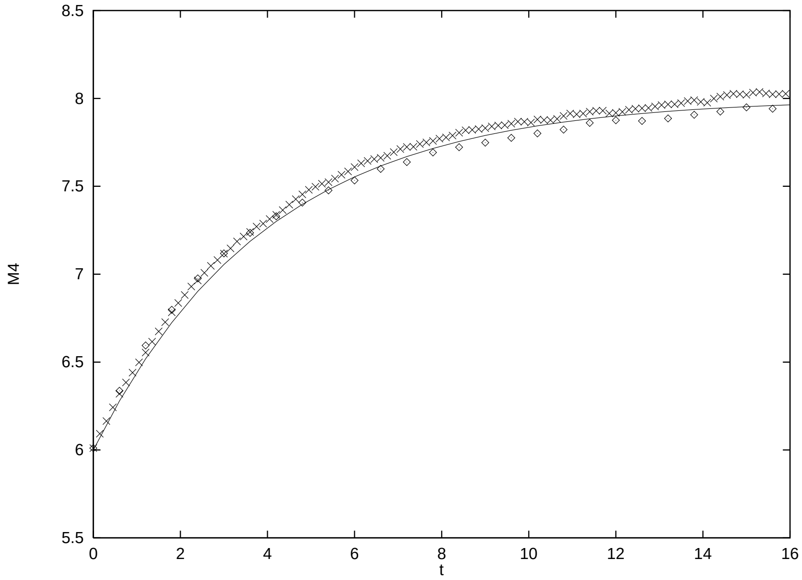

Figure 5: Maxwell molecules : Fourth order moment vs time with $\Delta t = 0.15$ for DSMC ($\times$) and $\Delta t = 0.6$ for TRMCH ($\diamond$). The line represents the exact solution.

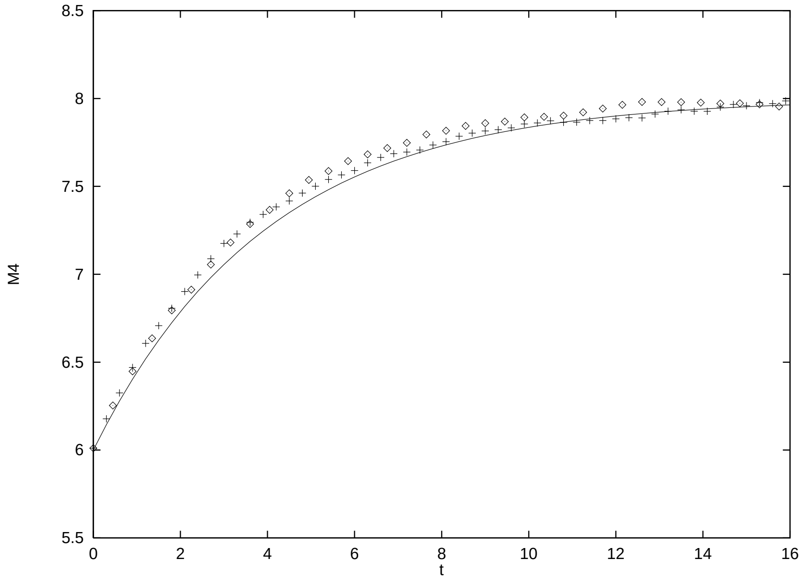

Figure 6: Maxwell molecules : Fourth order moment vs time with $\Delta t = 0.3$ for TRMC1 ($\times$) and $\Delta t = 0.45$ for TRMC2 ($\diamond$). The line represents the exact solution.