Abstract. This paper gives a rigorous derivation of a new stochastic particle method for the Boltzmann equation and illustrates its numerical efficiency. Using estimates based on the local temperature of the simulation cells, any truncation error related to large velocities is avoided. Moreover, time steps between collisions are larger than in the standard direct simulation method. This fact and an efficient modelling procedure for the index distribution of the collision partners lead to a considerable reduction of computational effort in certain applications.

Contents

1. Introduction

An important field of application of Monte Carlo methods is the numerical solution of nonlinear equations of high dimension. An example is the **Boltzmann equation** in rarefied gas dynamics. In the case of monatomic gases, this equation takes the form (cf. [5])

$$\frac{\partial}{\partial t} f(t,x,v) + (v, \nabla_x) f(t,x,v) = \int_{\mathcal{R}^3} dw \int_{\mathcal{S}^2} de\, B(v,w,e) \Big[f(t,x,v^*) f(t,x,w^*) - f(t,x,v) f(t,x,w) \Big], \tag{1.1}$$

[1]Supported by the Volkswagen-Stiftung (RiP-program at Oberwolfach)

where $t \geq 0$, $x \in D \subset \mathcal{R}^3$, $v \in \mathcal{R}^3$, and appropriate initial and boundary conditions are assumed. The symbol ∇_x denotes the vector of the partial derivatives with respect to x, D is a bounded domain in the three–dimensional Euclidean space $\mathcal{R}^3$, and $(.,.)$ is the scalar product. The function B is called the collision kernel. The symbols de and dw denote the uniform surface measure on the unit sphere $\mathcal{S}^2$ and the Lebesgue measure on $\mathcal{R}^3$, respectively. The objects v^* and w^* are defined as

$$v^* = v + e\,(e, w - v)\,, \quad w^* = w + e\,(e, v - w)\,, \tag{1.2}$$

where $v, w \in \mathcal{R}^3$, $e \in \mathcal{S}^2$. They are interpreted as the post–collision velocities of two particles with the pre–collision velocities v and w. The function f describes the time evolution of the distribution of gas particles in the position and velocity space, thus depending on seven independent variables.

In real world applications (like the reentry of a space shuttle into the atmosphere) Monte Carlo methods are the most common tools for the numerical solution of the Boltzmann equation (cf. [4], [6, Ch. 10], [12], [13]). They are based on the simulation of trajectories of stochastic interacting particle systems. A system of simulation particles

$$\Big(x_i(t), v_i(t)\Big)\,, \quad i = 1, \ldots, n\,, \quad t \geq 0\,, \tag{1.3}$$

is used to approximate the behaviour of the real gas described by the Boltzmann equation (1.1). Here $x_i(t) \in D$ and $v_i(t) \in \mathcal{R}^3$ denote the position and the velocity of the i-th particle at time t, respectively. The number of particles in the system is n. The solution of Eq. (1.1) is approximated in the sense that

$$\int_D \int_{\mathcal{R}^3} \varphi(x, v)\, f(t, x, v)\, dx\, dv \quad \sim \quad \frac{1}{n} \sum_{i=1}^{n} \varphi(x_i(t), v_i(t))\,, \tag{1.4}$$

for appropriate test functions φ. Convergence results have been obtained in [1] for Nanbu's method, and in [14] for Bird's method.

The time evolution of the particle system (1.3) is defined using a splitting technique. Namely, the simulation of the free flow of the particles and the simulation of their collisions are separated on a small time interval Δt. This means that on Δt, at a first step, the free flow is simulated disregarding the possible collisions. Then, at a second step, the collisions are simulated neglecting the free flow.

During the free flow simulation step, the particles move according to their velocities, i.e.

$$x_i(t + \Delta t) = x_i(t) + \int_t^{t+\Delta t} v_i(s)\, ds\,.$$

The velocities do not change unless a particle hits the boundary. In this case, the corresponding velocity changes according to the boundary conditions.

During the collision simulation step, a partition

$$D = \bigcup_{l=1}^{l_c} D_l \tag{1.5}$$

of the spatial domain D into a finite number l_c of disjoint cells is used. There is no interaction between different cells. In each cell, collisions of the particles are simulated. The principal **modelling procedure** is as follows.

Step 1: Calculate the time step until the next collision and add it to a **time counter**. If the time counter exceeds Δt, then stop the collision simulation step.

Step 2: Generate the indices i, j of the collision partners according to some probability distribution.

Step 3: Check whether the collision is fictitious, i.e. go to step 1 with a certain probability.

Step 4: Generate the direction vector $e \in \mathcal{S}^2$ according to some probability distribution.

Step 5: Calculate the new state, i.e. replace v_i, v_j by v_i^*, v_j^* (cf. (1.2)), and go to step 1.

For an efficient numerical implementation, the following points are to be taken into account:

- the time step used in Step 1 should be easy to compute and as big as possible;
- the distribution of the indices i, j used in Step 2 should be easy to generate;
- the probability of fictitious jumps used in Step 3 should be as small as possible.

The **purpose of the paper** is to give a rigorous derivation of a new stochastic particle scheme for the Boltzmann equation and to illustrate its numerical efficiency. We assume that the **collision kernel** (cf. (1.1)) satisfies

$$\int_{\mathcal{S}^2} B(v, w, e)\, de \le c_B \, \|v - w\|^{\varepsilon}\,, \qquad \forall v, w \in \mathcal{R}^3\,, \tag{1.6}$$

for some $\varepsilon \in [0, 2)$ and some constant c_B, where $\|.\|$ denotes the Euclidean norm. This assumption assures existence of a solution to the spatially homogeneous Boltzmann equation (cf. [5, Ch. 8, § 2]). In the hard sphere case the collision kernel B corresponding to the transformation (1.2) has the form

$$B(v, w, e) = C_{hs}\, |(v - w, e)|\,, \tag{1.7}$$

for some constant C_{hs}. In this case, one easily obtains

$$\int_{\mathcal{S}^2} B(v, w, e)\, de = 2\pi\, C_{hs}\, \|v - w\| \tag{1.8}$$

so that (1.6) is fulfilled with $\varepsilon = 1$ and $c_B = 2\pi\, C_{hs}$. Note that assumption (1.6) is also fulfilled for the variable hard sphere model and the variable soft sphere model, which are commonly used in applications ([4, Ch. 2]).

The paper is organized as follows. In **Section 2** we introduce a Markov jump process related to the simulation of particle collisions. This process provides a general frame for various modifications of the DSMC (direct simulation Monte Carlo) method. In **Section 3** we present the **main results** of the paper. Using some estimates based on the local temperature of the simulation cells, we derive a scheme that avoids any truncation error related to large velocities. We show that the corresponding time steps are always larger

than in the standard DSMC method. Finally, we propose an efficient procedure for generating the corresponding index distribution. **Section 4** contains results of numerical experiments. Here we illustrate both the absence of a truncation error and the improved efficiency due to the larger time steps.

2. A Markov process related to the Boltzmann equation

Various steps of numerical simulation involved in the DSMC algorithm are related to some approximations of the Boltzmann equation (1.1). The limiting equation (as the number of simulation particles tends to infinity) corresponding to the collision simulation step has the form ([14])

$$\frac{\partial}{\partial t} f(t,x,v) = \int_D dy \int_{\mathcal{R}^3} dw \int_{\mathcal{S}^2} de\, h(x,y)\, B(v,w,e) \Big[f(t,x,v^*)\, f(t,y,w^*) - f(t,x,v)\, f(t,y,w) \Big]. \tag{2.1}$$

The function

$$h(x,y) = \sum_{l=1}^{l_c} h_l(x,y)\,, \tag{2.2}$$

where

$$h_l(x,y) = \frac{1}{|D_l|}\, \chi_{D_l}(x)\, \chi_{D_l}(y)\,, \qquad x,y \in D\,, \tag{2.3}$$

is a mollifying kernel depending on the partition (1.5). Here $|D_l|$ is the Lebesgue measure of the cell D_l, and χ denotes the indicator function.

The Markov process related to the collision simulation step in one spatial cell D_l for the mollified Boltzmann equation (2.1)-(2.3) has the **infinitesimal generator**

$$\mathcal{A}(\Phi)(\bar{z}) = \frac{1}{2} \sum_{1 \le i \ne j \le n} \int_{\mathcal{S}^2} q(\bar{z},i,j,e) \Big[\Phi(J(\bar{z},i,j,e)) - \Phi(\bar{z}) \Big]\, de\,, \tag{2.4}$$

where

$$\bar{z} = \Big((x_1,v_1), \ldots, (x_n,v_n) \Big) \in \mathcal{Z} = \Big(D \times \mathcal{R}^3 \Big)^n \tag{2.5}$$

and Φ is an appropriate test function. The jump transformation is

$$[J(\bar{z},i,j,e)]_k = \begin{cases} (x_k,v_k)\,, \text{ if } & k \ne i,j\,, \\ (x_i,v_i^*)\,, \text{ if } & k = i\,, \\ (x_j,v_j^*)\,, \text{ if } & k = j\,, \end{cases} \tag{2.6}$$

where

$$v_i^* = v_i + e\,(e, v_j - v_i)\,, \quad v_j^* = v_j + e\,(e, v_i - v_j)\,, \tag{2.7}$$

and the intensity function is

$$q(\bar{z}, i, j, e) = \frac{1}{n}\, h_l(x_i, x_j)\, B(v_i, v_j, e)\,. \tag{2.8}$$

The generator (2.4) can be rewritten in the usual form of a jump process generator ([7, Ch. 4, § 2])

$$\mathcal{A}(\Phi)(\bar{z}) = \int_{\mathcal{Z}} \Big[\Phi(\bar{\zeta}) - \Phi(\bar{z})\Big]\, \mathcal{Q}(\bar{z}, d\bar{\zeta})\,, \tag{2.9}$$

where

$$\mathcal{Q}(\bar{z}, d\bar{\zeta}) = \frac{1}{2} \sum_{1 \le i \ne j \le n} \int_{\mathcal{S}^2} \delta_{J(\bar{z},i,j,e)}(d\bar{\zeta})\, q(\bar{z}, i, j, e)\, de \tag{2.10}$$

and δ denotes the Dirac measure.

The generator (2.9) does not change if one replaces $\mathcal{Q}$ by

$$\hat{\mathcal{Q}}(\bar{z}, d\bar{\zeta}) = \tag{2.11}$$
$$\frac{1}{2} \sum_{i,j=1}^{n} \left\{ \int_{\mathcal{S}^2} \delta_{J(\bar{z},i,j,e)}(d\bar{\zeta})\, q(\bar{z}, i, j, e)\, de + \delta_{\bar{z}}(d\bar{\zeta}) \Big[\hat{q}(\bar{z}, i, j) - \int_{\mathcal{S}^2} q(\bar{z}, i, j, e)\, de\,\Big] \right\},$$

where $\hat{q}$ is a function such that (cf. (2.8))

$$\int_{\mathcal{S}^2} q(\bar{z}, i, j, e)\, de = \frac{1}{n}\, h_l(x_i, x_j) \int_{\mathcal{S}^2} B(v_i, v_j, e)\, de \le \hat{q}(\bar{z}, i, j)\,. \tag{2.12}$$

Note that $J(\bar{z}, i, i, e) = \bar{z}$, according to (2.6), (2.7).

The **pathwise behaviour of the Markov process** (1.3) with the infinitesimal generator (2.9)-(2.11) is described as follows.

Coming to a state $\bar{z}$ (cf. (2.5)), the process stays there for a **random waiting time** $\hat{\tau}(\bar{z})$, which has an exponential distribution with the parameter

$$\hat{\pi}(\bar{z}) = \int_{\mathcal{Z}} \hat{\mathcal{Q}}(\bar{z}, d\bar{\zeta}) = \frac{1}{2} \sum_{i,j=1}^{n} \hat{q}(\bar{z}, i, j)\,, \tag{2.13}$$

i.e.

$$\mathrm{Prob}\,\{\hat{\tau}(\bar{z}) \ge t\} = \exp(-\hat{\pi}(\bar{z})\, t)\,.$$

Note that the expectation of the random waiting time is $\hat{\pi}(\bar{z})^{-1}$. If this value is sufficiently small, it can be used as a deterministic approximation to the random time step, i.e.

$$\hat{\tau}(\bar{z}) = \hat{\pi}(\bar{z})^{-1}\,. \tag{2.14}$$

After the time $\hat{\tau}(\bar{z})$, the process jumps into a state $\bar{\zeta}$, which is distributed according to the **jump distribution**

$$\hat{\pi}(\bar{z})^{-1}\, \hat{\mathcal{Q}}(\bar{z}, d\bar{\zeta})\,.$$

This distribution represents a superposition of simpler distributions (cf. (2.11)),

$$\hat{\pi}(\bar{z})^{-1}\,\hat{\mathcal{Q}}(\bar{z},d\bar{\zeta}) \;=\; \sum_{i,j=1}^{n} \frac{\hat{q}(\bar{z},i,j)}{2\,\hat{\pi}(\bar{z})} \left\{ \frac{\int_{\mathcal{S}^2} q(\bar{z},i,j,e)\,de}{\hat{q}(\bar{z},i,j)} \int_{\mathcal{S}^2} \delta_{J(\bar{z},i,j,e)}(d\bar{\zeta})\, \frac{q(\bar{z},i,j,e)}{\int_{\mathcal{S}^2} q(\bar{z},i,j,e)\,de}\,de \right.$$
$$\left. + \,\delta_{\bar{z}}(d\bar{\zeta}) \left[1 - \frac{\int_{\mathcal{S}^2} q(\bar{z},i,j,e)\,de}{\hat{q}(\bar{z},i,j)} \right] \right\} .$$

Consequently, first the **distribution of the indices** $i,j \in \{1,2,\ldots n\}$ is determined by the probabilities

$$\frac{\hat{q}(\bar{z},i,j)}{2\,\hat{\pi}(\bar{z})} = \frac{\hat{q}(\bar{z},i,j)}{\sum_{i,j=1}^{n} \hat{q}(\bar{z},i,j)} \,. \tag{2.15}$$

Given i and j , the new state is $\bar{\zeta} = \bar{z}$ with probability (cf. (2.8))

$$1 - \frac{\int_{\mathcal{S}^2} q(\bar{z},i,j,e)\,de}{\hat{q}(\bar{z},i,j)} = 1 - \frac{h_l(x_i,x_j)\int_{\mathcal{S}^2} B(v_i,v_j,e)\,de}{n\,\hat{q}(\bar{z},i,j)} \,. \tag{2.16}$$

Expression (2.16) is therefore called **probability of a fictitious jump** . Otherwise, i.e. with the remaining probability, the **new state** is

$$\bar{\zeta} = J(\bar{z},i,j,e)\,,$$

where the **distribution of the direction vector** $e \in \mathcal{S}^2$ is (cf. (2.8))

$$\frac{q(\bar{z},i,j,e)}{\int_{\mathcal{S}^2} q(\bar{z},i,j,e)\,de} = \frac{B(v_i,v_j,e)}{\int_{\mathcal{S}^2} B(v_i,v_j,e)\,de} \,. \tag{2.17}$$

The distribution of the Markov process and therefore its convergence properties do not depend on the function $\hat{q}$ in (2.12). However, the choice of this function is of importance for numerical purposes, since it provides different ways of generating trajectories of the process. We illustrate this by two examples.

Example 2.1 (The "direct simulation" scheme) *A trivial choice of the function $\hat{q}$ in (2.12) is*

$$\hat{q}(\bar{z},i,j) = \frac{1}{n}\,h_l(x_i,x_j) \int_{\mathcal{S}^2} B(v_i,v_j,e)\,de \,. \tag{2.18}$$

In the hard sphere case (1.7) one obtains (cf. (1.8))

$$\hat{q}(\bar{z},i,j) = \frac{2\pi\,C_{hs}}{n}\,h_l(x_i,x_j)\,\|v_i - v_j\| \,.$$

The time step (2.14), (2.13) takes the form

$$\hat{\tau}(\bar{z}) = \frac{n\,|D_l|}{\pi\,C_{hs}\,\sum_{i,j\,:\,x_i,x_j\in D_l} \|v_i - v_j\|} \,. \tag{2.19}$$

The distribution (2.15) of i, j is

$$\frac{\chi_{D_l}(x_i)\,\chi_{D_l}(x_j)\,\|v_i - v_j\|}{\sum_{i,j\,:\,x_i,x_j\in D_l}\|v_i - v_j\|}\,, \tag{2.20}$$

i.e. the pairs of particles are chosen from the set

$$\{i, j = 1, 2, \ldots, n \;:\; x_i, x_j \in D_l\} \tag{2.21}$$

with probabilities proportional to their relative velocities. The probability (2.16) of a fictitious jump is zero (cf. (2.18)) as well as the probability of the event $i = j$. Finally, the distribution (2.17) of e is (cf. (1.7), (1.8))

$$\frac{|(v_i - v_j, e)|}{2\pi\,\|v_i - v_j\|}\,.$$

The modelling procedure for the direct simulation scheme is extremely simple. However, the numerical application may face serious problems if the number of particles n is large. In general, in the calculation of the waiting time parameter (2.19) or the probabilities (2.20) one has quadratic effort with respect to n . Therefore it is important to look for an appropriate choice of the function $\hat{q}$ in (2.12), which may lead to a substantial gain in the efficiency of the modelling of the process.

Example 2.2 (Bird's "no time counter" scheme) *Using assumption (1.6) one chooses the function $\hat{q}$ in (2.12) in the form*

$$\hat{q}(\bar{z}, i, j) = \frac{c_B}{n}\,h_l(x_i, x_j)\,U^{\varepsilon}_{max}\,, \tag{2.22}$$

where

$$U_{max} = U_{max}(\bar{z}) = \max_{i,j\,:\,x_i,x_j\in D_l}\|v_i - v_j\|\,. \tag{2.23}$$

The time step (2.14), (2.13) is

$$\hat{\tau}(\bar{z}) = \frac{2\,n\,|D_l|}{c_B\,n_l^2\,U^{\varepsilon}_{max}}\,, \tag{2.24}$$

where

$$n_l = n_l(\bar{z}) = \sum_{i=1}^{n}\chi_{D_l}(x_i) \tag{2.25}$$

denotes the number of particles in the cell D_l . The indices i, j are generated uniformly on the set (2.21) according to (2.15). The probability of the event $i = j$ is $1/n_l$. The probability (2.16) of a fictitious jump is

$$1 - \frac{\int_{S^2} B(v_i, v_j, e)\,de}{c_B\,U^{\varepsilon}_{max}}\,. \tag{2.26}$$

The distribution of e is (2.17).

The value of $U_{max}(\bar{z})$ may change after each collision. Its calculation takes a quadratic effort with respect to n_l. Therefore, the following approximate procedure is used.

Let σ_k, $k = 1, 2, \ldots$, denote the time moments at which a (possibly fictitious) collision takes place, and i_k, j_k the corresponding indices of collision partners. Then the **accumulated maximal norm of the relative velocities** *of collision partners up to time t is defined as*

$$\tilde{U}_{max}(t) = \max\left(\tilde{U}_{max}(0), \max_{k\,:\,\sigma_k \le t} \|v_{i_k} - v_{j_k}\|\right). \tag{2.27}$$

The value $\tilde{U}_{max}(t)$ is used in (2.24) and (2.26) instead of U_{max} from (2.23) during the collision simulation step on the time interval $[t, t+\Delta t]$ (cf. Section 1). The starting value $\tilde{U}_{max}(0)$ is a reasonable but low quantity (cf. [4, p. 443]).

The method described in Example 2.2 has been developed by G. A. Bird since the sixties (cf. [2], [3]). We also refer to [11], [9], [8], [10], and the detailed discussion of the historical development in [4, § 11.1].

Since the function (2.22) with $\tilde{U}_{max}$ does not always satisfy (2.12) at the beginning of the collision simulation step, there will be a certain additional error in this procedure. This error will vanish when $\tilde{U}_{max}$ increases and adapts to the system. On the other hand, $\tilde{U}_{max}$ remembers all events from the past. Therefore it may become too large later on and create redundant fictitious collisions. These effects will be illustrated by numerical examples in Section 4. In the next section we derive a method that avoids both problems.

3. The temperature time counter scheme

3.1. Derivation of a temperature based time counter

In the cell D_l we consider the **local mean velocity**

$$V = V(\bar{z}) = \frac{1}{n_l(\bar{z})} \sum_{i\,:\,x_i \in D_l} v_i \tag{3.1}$$

and the **local temperature**

$$T = T(\bar{z}) = \frac{1}{3\,n_l(\bar{z})} \sum_{i\,:\,x_i \in D_l} \|v_i - V(\bar{z})\|^2 = \frac{1}{3}\left[\frac{1}{n_l(\bar{z})} \sum_{i\,:\,x_i \in D_l} \|v_i\|^2 - \|V(\bar{z})\|^2\right], \tag{3.2}$$

where $\bar{z}$ is as in (2.5) and n_l is defined in (2.25). These quantities are preserved during the collision simulation step. Using (3.1) and (3.2) one obtains

$$\begin{aligned} \sum_{j\,:\,x_j \in D_l} \|v_i - v_j\|^2 &= \sum_{j\,:\,x_j \in D_l} \left[\|v_i - V\|^2 - 2\,(v_i - V, v_j - V) + \|v_j - V\|^2\right] \\ &= \|v_i - V\|^2\, n_l + 3\,T\, n_l \end{aligned} \tag{3.3}$$

and

$$\sum_{i,j\,:\,x_i, x_j \in D_l} \|v_i - v_j\|^2 = 6\,T\,n_l^2\,. \tag{3.4}$$

Let $\alpha, \beta > 0$ be such that

$$x^{\varepsilon} \le \alpha\, x^2 + \beta\,, \qquad \forall x \ge 0\,. \tag{3.5}$$

Then, according to (1.6), we obtain

$$\int_{S^2} B(v,w,e)\, de \le c_B\, T^{\frac{\varepsilon}{2}} \left(\frac{\|v-w\|}{\sqrt{T}} \right)^{\varepsilon} \le c_B\, T^{\frac{\varepsilon}{2}} \left[\alpha\, \frac{\|v-w\|^2}{T} + \beta \right]$$

and choose the function $\hat{q}$ in (2.12) as

$$\hat{q}(\bar{z}, i, j) = \frac{c_B}{n}\, h_l(x_i, x_j)\, T^{\frac{\varepsilon}{2}} \left[\alpha\, \frac{\|v_i - v_j\|^2}{T} + \beta \right]. \tag{3.6}$$

According to (3.4), the time step (2.14), (2.13) takes the form

$$\hat{\tau}(\bar{z}) = \frac{2\, n\, |D_l|}{c_B\, T^{\frac{\varepsilon}{2}} \sum_{i,j\,:\, x_i, x_j \in D_l} \left[\alpha\, \frac{\|v_i - v_j\|^2}{T} + \beta \right]} = \frac{2\, n\, |D_l|}{c_B\, T^{\frac{\varepsilon}{2}}\, n_l^2\, [6\,\alpha + \beta]}\,. \tag{3.7}$$

In order to increase the time step (3.7), we minimize the expression $6\,\alpha + \beta$ with respect to α, β satisfying (3.5).

Lemma 3.1 *The expression $6\,\alpha + \beta$ takes its minimum from among the α, β satisfying (3.5) for*

$$\alpha = \alpha(\varepsilon) = 6^{\frac{\varepsilon}{2}-1}\, \frac{\varepsilon}{2} \tag{3.8}$$

and

$$\beta = \beta(\varepsilon) = 6^{\frac{\varepsilon}{2}} \left(1 - \frac{\varepsilon}{2} \right). \tag{3.9}$$

Proof. The function

$$\varphi(x) = \alpha\, x^2 - x^{\varepsilon} + \beta\,, \quad x \ge 0\,, \quad \alpha, \beta > 0\,, \quad \varepsilon \in (0,2)\,,$$

takes its minimum at some point x_0 satisfying

$$\varphi'(x_0) = 2\,\alpha\, x_0 - \varepsilon\, x_0^{\varepsilon - 1} = 0\,,$$

i.e.

$$x_0 = \left(\frac{\varepsilon}{2\,\alpha} \right)^{\frac{1}{2-\varepsilon}} \quad \text{or} \quad \alpha = \frac{\varepsilon\, x_0^{\varepsilon-2}}{2}\,. \tag{3.10}$$

The minimum is non-negative if

$$\beta \ge x_0^{\varepsilon} - \alpha\, x_0^2 = x_0^{\varepsilon} \left(1 - \frac{\varepsilon}{2} \right). \tag{3.11}$$

In order to minimize the expression $6\,\alpha + \beta$, we consider the function

$$\psi(x_0) = 6\, \frac{\varepsilon\, x_0^{\varepsilon-2}}{2} + x_0^{\varepsilon} \left(1 - \frac{\varepsilon}{2} \right).$$

Condition

$$\psi'(x_*) = 3\,(\varepsilon-2)\,\varepsilon\,x_*^{\varepsilon-3} + \varepsilon\,x_*^{\varepsilon-1}\left(1-\frac{\varepsilon}{2}\right) = 0$$

implies $x_* = \sqrt{6}$. Now (3.8), (3.9) follow from (3.10) and (3.11). ∎

With the optimal choice (3.8), (3.9) of the parameters α, β in the function (3.6) the **time step** (3.7) takes the form

$$\hat{\tau}(\bar{z}) = \frac{2\,n\,|D_l|}{c_B\,n_l^2\,(6\,T)^{\frac{\varepsilon}{2}}}\,. \tag{3.12}$$

The **distribution of the indices** i, j corresponding to (3.6) is (cf. (2.15), (3.4))

$$p_{i,j} = \frac{\alpha(\varepsilon)\,\frac{\|v_i-v_j\|^2}{T} + \beta(\varepsilon)}{n_l^2\,[6\,\alpha(\varepsilon)+\beta(\varepsilon)]} = \frac{\frac{\varepsilon}{12}\,\frac{\|v_i-v_j\|^2}{T} + 1 - \frac{\varepsilon}{2}}{n_l^2} \tag{3.13}$$

being concentrated on the set (2.21). The probability of the event $i=j$ is

$$\frac{\beta(\varepsilon)}{n_l\,[6\,\alpha(\varepsilon)+\beta(\varepsilon)]} = \frac{1}{n_l}\left(1-\frac{\varepsilon}{2}\right).$$

The **probability of a fictitious jump** corresponding to (3.6) is according to (2.16)

$$1 - \frac{\int_{S^2} B(v_i, v_j, e)\,de}{c_B\,T^{\frac{\varepsilon}{2}}\left[\alpha(\varepsilon)\,\frac{\|v_i-v_j\|^2}{T} + \beta(\varepsilon)\right]} = 1 - \frac{\int_{S^2} B(v_i, v_j, e)\,de}{c_B\,(6\,T)^{\frac{\varepsilon}{2}}\left[\frac{\varepsilon}{12}\,\frac{\|v_i-v_j\|^2}{T} + 1 - \frac{\varepsilon}{2}\right]}\,. \tag{3.14}$$

The **distribution of the direction vector** e is (2.17) independently of (3.6).

3.2. Comparison with the standard DSMC method

It follows from (3.4) that (cf. (2.23))

$$U_{max} \geq \sqrt{6\,T}\,. \tag{3.15}$$

Thus, the time step (3.12) of the temperature time counter scheme is always larger than the time step (2.24) of the standard DSMC method. These time steps may differ by several orders of magnitude, as the following example shows.

Example 3.2 *Let the particle system approximate the measure*

$$g\,\delta_{-c\,\frac{1-g}{g}}(dv) + (1-g)\,\delta_c(dv)\,, \qquad c \in \mathcal{R}^3\,, \quad g \in (0,1)\,,$$

i.e.

$$v_i = -c\,\frac{1-g}{g}\,, \quad i = 1, \ldots, [g\,n_l]\,, \qquad v_i = c\,, \quad i = [g\,n_l]+1, \ldots, n_l\,,$$

where $[.]$ *denotes the integer part. We have (cf. (3.1), (3.2))*

$$V = 0\,, \qquad 3\,T = \frac{1}{n_l}\left[\|c\|^2\,\frac{(1-g)^2}{g^2}\,g\,n_l + \|c\|^2\,n_l\,(1-g)\right] = \|c\|^2\,\frac{1-g}{g}$$

and

$$U_{max} = \max_{i,j}\|v_i - v_j\| = \|c\|\,\frac{1}{g}$$

so that

$$\frac{U_{max}}{\sqrt{6\,T}} = \frac{1}{\sqrt{2\,g\,(1-g)}}\,. \tag{3.16}$$

In the case $g = \frac{1}{2}$ *identity (3.16) takes the form* $\frac{U_{max}}{\sqrt{6\,T}} = \sqrt{2}$ *being relatively close to the lower bounds given by (3.15). If* $g \sim 0$ *or* $g \sim 1$, *the expression (3.16) is arbitrarily large.*

However, the distribution of the indices (3.13) is much more complicated than the uniform distribution related to the time counter (2.24). An efficient modelling procedure for the distribution (3.13) will be studied in the next subsection.

3.3. Modelling of the index distribution

Using (3.3) we represent the probabilities (3.13) in the form $p_{i,j} = p_i\,p_{j|i}\,,$ where the probability of i is

$$p_i = \sum_j p_{i,j} = \frac{\frac{\varepsilon}{12}\,\frac{\|v_i - V\|^2}{T} + 1 - \frac{\varepsilon}{4}}{n_l} \tag{3.17}$$

and the probability of j given i takes the form

$$p_{j|i} = \frac{p_{i,j}}{p_i} = \frac{\frac{\varepsilon}{12}\,\frac{\|v_i - v_j\|^2}{T} + 1 - \frac{\varepsilon}{2}}{n_l\left[\frac{\varepsilon}{12}\,\frac{\|v_i - V\|^2}{T} + 1 - \frac{\varepsilon}{4}\right]}\,. \tag{3.18}$$

Both distributions (3.17) and (3.18) are generated using the **acceptance-rejection technique**. This technique based on the following lemma is useful in many modifications of the DSMC method.

Lemma 3.3 *Consider a measurable space* (X, μ) *and two functions* f *and* F *on* X *satisfying the* **majorant condition**

$$0 \le f(x) \le F(x)\,, \quad \forall x \in X\,. \tag{3.19}$$

Assume that

$$\int_X f(x)\,\mu(dx) > 0 \quad \textit{and} \quad \int_X F(x)\,\mu(dx) < \infty\,.$$

Let a random variable ξ be defined by the following procedure:

1. *Generate a random variable η with the probability density*

$$P(x) = \frac{F(x)}{\int_X F(x)\,\mu(dx)}\,. \tag{3.20}$$

2. *Generate independently a random variable u uniformly distributed on $[0,1]$.*
3. *If the* **acceptance condition**

$$u \le \frac{f(\eta)}{F(\eta)} \tag{3.21}$$

is satisfied, then $\xi = \eta$ and stop.

4. *If (3.21) is not fulfilled, then go to 1.*

Then the random variable ξ has the probability density

$$p(x) = \frac{f(x)}{\int_X f(x)\,\mu(dx)}\,.$$

We apply the idea of **ordering the particles** with respect to the absolute values of their normalized velocities $\frac{v_i - V}{\sqrt{T}}$. Note that the definitions (3.1) and (3.2) immediately imply that

$$\frac{1}{n_l}\sum_{i\,:\,x_i\in D_l}\frac{v_i - V}{\sqrt{T}} = 0 \quad \text{and} \quad \frac{1}{3\,n_l}\sum_{i\,:\,x_i\in D_l}\left\|\frac{v_i - V}{\sqrt{T}}\right\|^2 = 1\,.$$

Consider some values

$$0 < b_1 < \ldots < b_K\,, \qquad K \ge 1\,, \tag{3.22}$$

where

$$\frac{\|v_i - V\|}{\sqrt{T}} \le b_K = b_K(\bar z)\,, \quad \forall i\,. \tag{3.23}$$

Define

$$\hat b(v_i) = \min\left\{b_k\,,\ k = 1,\ldots,K\ :\ \frac{\|v_i - V\|}{\sqrt{T}} \le b_k\right\}$$

and note that

$$\frac{\|v_i - V\|}{\sqrt{T}} \le \hat b(v_i)\,, \quad \forall i\,. \tag{3.24}$$

The function $\hat b$ taking values $b_1,\ldots,b_K$ provides a certain non-global majorant for the normalized velocities. The normalized velocities are divided into groups according to their individual majorants. Let

$$\gamma_k = \gamma_k(\bar z) = \#\{i\ :\ \hat b(v_i) = b_k\}\,, \qquad k = 1,\ldots,K\,,$$

denote the number of normalized velocities with the individual majorant b_k.

The **distribution of the first index** i (cf. (3.17), (3.19), (3.24)) is generated using the acceptance–rejection technique with

$$X = \{i = 1, 2, \ldots, n \,:\, x_i \in D_l\}\,, \tag{3.25}$$

$$f_i = \frac{\varepsilon}{12}\,\frac{\|v_i - V\|^2}{T} + 1 - \frac{\varepsilon}{4} \tag{3.26}$$

and

$$F_i = \frac{\varepsilon}{12}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{4}\,. \tag{3.27}$$

In order to generate (3.20) we calculate

$$\sum_j F_j = \sum_{k=1}^{K} \sum_{j\,:\,\hat{b}(v_j)=b_k} F_j = \sum_{k=1}^{K} \gamma_k \left[\frac{\varepsilon}{12}\,b_k^2 + 1 - \frac{\varepsilon}{4}\right] = \frac{\varepsilon}{12}\,\sum_{k=1}^{K} \gamma_k\, b_k^2 + n_l \left[1 - \frac{\varepsilon}{4}\right],$$

and

$$\frac{F_i}{\sum_j F_j} = \frac{\frac{\varepsilon}{12}\,\sum_{k=1}^{K} \gamma_k\, b_k^2}{\frac{\varepsilon}{12}\,\sum_{k=1}^{K} \gamma_k\, b_k^2 + n_l\,(1 - \frac{\varepsilon}{4})}\,\frac{\hat{b}(v_i)^2}{\sum_{k=1}^{K} \gamma_k\, b_k^2} + \frac{n_l\,(1 - \frac{\varepsilon}{4})}{\frac{\varepsilon}{12}\,\sum_{k=1}^{K} \gamma_k\, b_k^2 + n_l\,(1 - \frac{\varepsilon}{4})}\,\frac{1}{n_l}\,.$$

According to this representation the index i is distributed **uniformly** with probability

$$\frac{n_l\,(1 - \frac{\varepsilon}{4})}{\frac{\varepsilon}{12}\,\sum_{k=1}^{K} \gamma_k\, b_k^2 + n_l\,(1 - \frac{\varepsilon}{4})}\,.$$

With probability

$$1 - \frac{n_l\,(1 - \frac{\varepsilon}{4})}{\frac{\varepsilon}{12}\,\sum_{k=1}^{K} \gamma_k\, b_k^2 + n_l\,(1 - \frac{\varepsilon}{4})}\,,$$

the index i is distributed according to the distribution

$$\frac{\hat{b}(v_i)^2}{\sum_{k=1}^{K} \gamma_k\, b_k^2}\,. \tag{3.28}$$

Thus, **first** a number of a group of indices is chosen according to the probabilities

$$\frac{\gamma_k\, b_k^2}{\sum_{\mu=1}^{K} \gamma_\mu\, b_\mu^2}\,, \qquad k = 1, \ldots, K\,, \tag{3.29}$$

and **then** the index i is chosen **uniformly** in the group $\{i \,:\, \hat{b}(v_i) = b_k\}$. Finally, the index i is **accepted** with probability (cf. (3.21), (3.26), (3.27))

$$\frac{\frac{\varepsilon}{12}\,\frac{\|v_i - V\|^2}{T} + 1 - \frac{\varepsilon}{4}}{\frac{\varepsilon}{12}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{4}}\,. \tag{3.30}$$

The **distribution of the second index** j (cf. (3.18), (3.19)) is generated using the acceptance–rejection technique with (3.25),

$$f_{j|i} = \frac{\varepsilon}{12}\,\frac{\|v_i - v_j\|^2}{T} + 1 - \frac{\varepsilon}{2} \tag{3.31}$$

and

$$F_{j|i} = \frac{\varepsilon}{6}\left[\hat{b}(v_i)^2 + \hat{b}(v_j)^2\right] + 1 - \frac{\varepsilon}{2}\,. \tag{3.32}$$

Note that

$$\frac{\|v_i - v_j\|}{\sqrt{T}} \le \frac{\|v_i - V\|}{\sqrt{T}} + \frac{\|v_j - V\|}{\sqrt{T}} \le \left[\hat{b}(v_i) + \hat{b}(v_j)\right],$$

according to (3.24), and therefore

$$\frac{\varepsilon}{12}\,\frac{\|v_i - v_j\|^2}{T} + 1 - \frac{\varepsilon}{2} \le \frac{\varepsilon}{6}\left[\hat{b}(v_i)^2 + \hat{b}(v_j)^2\right] + 1 - \frac{\varepsilon}{2}\,.$$

One obtains

$$\sum_j F_{j|i} = n_l\left[\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2}\right] + \frac{\varepsilon}{6}\sum_{k=1}^{K}\gamma_k\, b_k^2$$

and the representation (cf. (3.20))

$$\begin{aligned}\frac{F_{j|i}}{\sum_\mu F_{\mu|i}} \;=\; & \frac{n_l\,(\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2})}{n_l\,(\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2}) + \frac{\varepsilon}{6}\,\sum_{k=1}^{K}\gamma_k\, b_k^2}\;\frac{1}{n_l} + \\ & \frac{\frac{\varepsilon}{6}\,\sum_{k=1}^{K}\gamma_k\, b_k^2}{n_l\,(\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2}) + \frac{\varepsilon}{6}\,\sum_{k=1}^{K}\gamma_k\, b_k^2}\;\frac{\hat{b}(v_j)^2}{\sum_{k=1}^{K}\gamma_k\, b_k^2}\,.\end{aligned}$$

Thus, with probability

$$\frac{n_l\,(\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2})}{n_l\,(\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2}) + \frac{\varepsilon}{6}\,\sum_{k=1}^{K}\gamma_k\, b_k^2}\,,$$

the index j is distributed **uniformly**. With probability

$$1 - \frac{n_l\,(\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2})}{n_l\,(\frac{\varepsilon}{6}\,\hat{b}(v_i)^2 + 1 - \frac{\varepsilon}{2}) + \frac{\varepsilon}{6}\,\sum_{k=1}^{K}\gamma_k\, b_k^2}\,,$$

the index j is distributed according to (3.28), i.e. **first** the number of the group is chosen according to (3.29) and **then** the index j is generated **uniformly** in the corresponding group. Finally, the index j is **accepted** with probability (cf. (3.21), (3.31), (3.32))

$$\frac{\frac{\varepsilon}{12}\,\frac{\|v_i - v_j\|^2}{T} + 1 - \frac{\varepsilon}{2}}{\frac{\varepsilon}{6}\left[\hat{b}(v_i)^2 + \hat{b}(v_j)^2\right] + 1 - \frac{\varepsilon}{2}}\,. \tag{3.33}$$

Remark 3.4 *In our test calculations we define b_K from the initial configuration of the system as (cf. (3.23))*

$$b_K = \max\left\{ \frac{\|v_i - V\|}{\sqrt{T}} \ : \ i = 1, \ldots, n \,,\, x_i \in D_l \right\},$$

and put

$$b_k = k\,\frac{b_K}{K}\,, \quad k = 1, \ldots, K\,.$$

The value of b_K and the values of γ_k, $k = 1, \ldots, K$, are updated during the simulation. Note that, if some γ_k equals zero, then the corresponding group is simply not chosen (cf. (3.29)).

4. Numerical experiments

4.1. Test example

The initial distribution of our test example is a weighted mixture of two Maxwellians, namely

$$f_0(v) = g\,\frac{1}{(2\pi\,T_1)^{3/2}}\exp\left(-\frac{\|v - V_1\|^2}{2\,T_1}\right) + (1-g)\frac{1}{(2\pi\,T_2)^{3/2}}\exp\left(-\frac{\|v - V_2\|^2}{2\,T_2}\right). \qquad (4.1)$$

This form is chosen in order to illustrate the effect described in Example 3.2. The function $f(t,v)$ is the solution of the spatially homogeneous Boltzmann equation with the initial condition $f(0,v) = f_0(v)$. It relaxes to the final equilibrium Maxwell distribution

L : 6.7

$$f_\infty(v) \ = \ \frac{1}{(2\pi\,T)^{3/2}}\exp\left(-\frac{\|v - V\|^2}{2\,T}\right)$$

with the parameters

$$V = g\,V_1 + (1-g)\,V_2 \qquad (4.2)$$

and

$$T = g\,T_1 + (1-g)\,T_2 + \frac{1}{3}\,g\,(1-g)\,\|V_1 - V_2\|^2\,. \qquad (4.3)$$

We consider the problem of calculating the time evolution of the second moments

$$m_{i,j}(t) = \int_{\mathcal{R}^3} v_i\,v_j\,f(t,v)\,dv\,, \quad i,j = 1,2,3\,, \qquad (4.4)$$

and the third moments

$$r_i(t) = \int_{\mathcal{R}^3} v_i\,\|v\|^2\,f(t,v)\,dv\,, \quad i = 1,2,3\,. \qquad (4.5)$$

The initial values of these functionals are given by

$$m_{i,j}(0) = \Big(g\,T_1 + (1-g)\,T_2\Big)\delta_{i,j} + g\,V_{1,i}\,V_{1,j} + (1-g)\,V_{2,i}\,V_{2,j}\,, \tag{4.6}$$

where $\delta_{i,j}$ denotes the Kronecker symbol, and

$$r_i(0) = g\Big(5T_1 + \|V_1\|^2\Big)V_{1,i} + (1-g)\Big(5T_2 + \|V_2\|^2\Big)V_{2,i}\,. \tag{4.7}$$

Our **first test case** is

$$g = 0.001\,, \quad V_1 = (-999, 0, 0)\,, \quad V_2 = (1, 0, 0)\,, \quad T_1 = 1\,, \quad T_2 = 1\,. \tag{4.8}$$

For this special choice we obtain from (4.2), (4.3)

$$V = 0\,, \qquad T = 334 \tag{4.9}$$

so that

$$\lim_{t\to\infty} m_{11}(t) = 334\,, \qquad \lim_{t\to\infty} r_1(t) = 0\,. \tag{4.10}$$

The initial values (4.6), (4.7) take the form

$$m_{11}(0) = 1000\,, \qquad r_1(0) = -997004\,. \tag{4.11}$$

A particle system approximating the distribution (4.1), (4.8) consists of many particles ($\sim$ 99.9%) with small velocities ($\sim V_2$) and of a few particles ($\sim$ 0.1%) with large velocities ($\sim V_1$).

Our **second test case** is the relaxation problem with equal weights, i.e.

$$g = 0.5\,, \quad V_1 = (-31.6, 0, 0)\,, \quad V_2 = (31.6, 0, 0)\,, \quad T_1 = 1\,, \quad T_2 = 1\,, \tag{4.12}$$

so that (4.9) holds.

We apply two methods:

- the standard DSMC method (cf. Example 2.2) with the accumulated maximal norm of the relative velocities (2.27);
- the TTC method (described in Section 3) with different numbers of groups (cf. (3.22)).

4.2. Statistical notions

First we introduce some definitions and notations that are helpful for the understanding of stochastic numerical procedures.

Functionals of the form (cf. (4.4), (4.5))

$$F(t) = \int_{\mathcal{R}^3} \varphi(v)\, f(t,v)\, dv \tag{4.13}$$

are approximated by the random variable (cf. (1.4))

$$\xi^{(n)}(t) = \frac{1}{n}\sum_{i=1}^{n}\varphi(v_i(t))\,, \tag{4.14}$$

where $(v_1(t),\ldots,v_n(t))$ are the velocities of the particle system. In order to estimate and to reduce the random fluctuations of the estimator (4.14), a number N of independent ensembles of particles is generated. The corresponding values of the random variable are denoted by $\xi_1^{(n)}(t),\ldots,\xi_N^{(n)}(t)$. The **empirical mean value** of the random variable (4.14), i.e.

$$\eta_1^{(n,N)}(t) = \frac{1}{N}\sum_{j=1}^{N}\xi_j^{(n)}(t)\,, \tag{4.15}$$

is then used as an approximation to the functional (4.13). The error of this approximation is $|\eta_1^{(n,N)}(t) - F(t)|$ consisting of the following two components.

The **systematic error** is the difference between the mathematical expectation of the random variable (4.14) and the exact value of the functional, i.e.

$$e_{sys}^{(n)}(t) = E\xi^{(n)}(t) - F(t)\,.$$

The **statistical error** is the difference between the empirical mean value and the expected value of the random variable, i.e.

$$e_{stat}^{(n,N)}(t) = \eta_1^{(n,N)}(t) - E\xi^{(n)}(t)\,.$$

A **confidence interval** for the expectation of the random variable $\xi^{(n)}(t)$ is obtained as

$$I_p = \left[\eta_1^{(n,N)}(t) - \lambda_p\sqrt{\frac{\mathrm{Var}\,\xi^{(n)}(t)}{N}}\,,\ \eta_1^{(n,N)}(t) + \lambda_p\sqrt{\frac{\mathrm{Var}\,\xi^{(n)}(t)}{N}}\right], \tag{4.16}$$

where

$$\mathrm{Var}\,\xi^{(n)}(t) := E\left[\xi^{(n)}(t) - E\xi^{(n)}(t)\right]^2 = E\left[\xi^{(n)}(t)\right]^2 - \left[E\xi^{(n)}(t)\right]^2 \tag{4.17}$$

is the **variance** of the random variable (4.14), and $p \in (0,1)$ is the **confidence level**. This means that

$$\mathrm{Prob}\left\{E\xi^{(n)}(t) \notin I_p\right\} = \mathrm{Prob}\left\{|e_{stat}^{(n,N)}(t)| \geq \lambda_p\sqrt{\frac{\mathrm{Var}\,\xi^{(n)}(t)}{N}}\right\} \sim p\,.$$

Thus, the value

$$c^{(n,N)}(t) = \lambda_p\sqrt{\frac{\mathrm{Var}\,\xi^{(n)}(t)}{N}}$$

is a probabilistic upper bound for the statistical error.

In the calculations we use a confidence level of $p = 0.999$ and $\lambda_p = 3.2$. The variance (4.17) is approximated by the corresponding empirical value, i.e.

$$\mathrm{Var}\,\xi^{(n)}(t) \sim \eta_2^{(n,N)}(t) - \left[\eta_1^{(n,N)}(t)\right]^2,$$

where

$$\eta_2^{(n,N)}(t) = \frac{1}{N}\sum_{j=1}^{N}\left[\xi_j^{(n)}(t)\right]^2$$

is the **empirical second moment** of the random variable (4.14).

4.3. Results of computations

First we study the time evolution of the accumulated maximal norm of the relative velocities $\tilde{U}_{max}$ (cf. (2.27)) for the standard DSMC method in the test case (4.8). The behaviour of the quantity $\tilde{U}_{max}(t)$ on the time interval $[0, 0.025]$ is displayed in **Figure 1** for different particle numbers. The lines correspond to $n = 100000$, $n = 10000$ and $n = 1000$ from top to bottom.

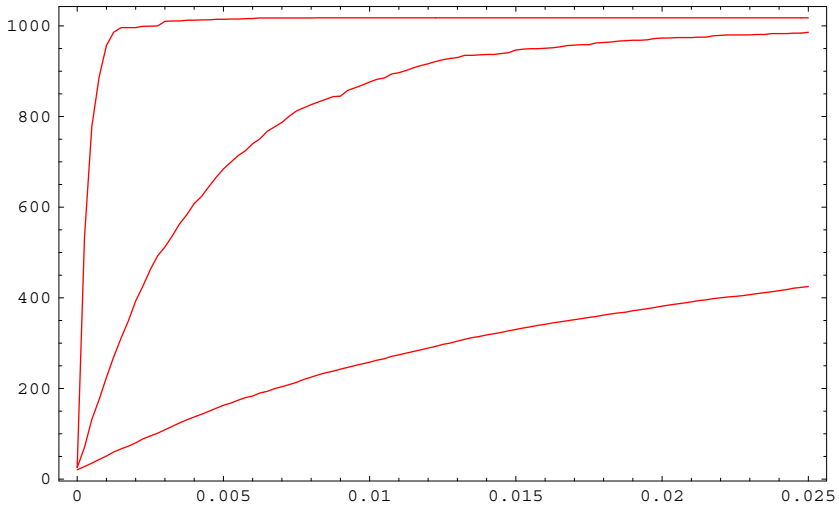

Figure 1: Accumulated maximal relative velocity norm

On a certain starting interval $\tilde{U}_{max}$ is much less than the correct value 1000. Thus, some additional systematic error is expected for the standard DSMC method if the number particles is not large enough. However, this interval becomes smaller when the number of particles increases.

We illustrate this error for the second moment $m_{11}(t)$ (cf. (4.4), (4.10), (4.11)) in **Figure 2**. The solid lines (from top to bottom) correspond to 1000, 10000 and 100000 particles for the standard DSMC method. The dashed line represents the numerical solution obtained using the TTC method with only 1000 particles. Note that the number of shells in the TTC method does not influence the systematic error. The results for the DSMC method with 1000000 particles are identical with the dashed line.

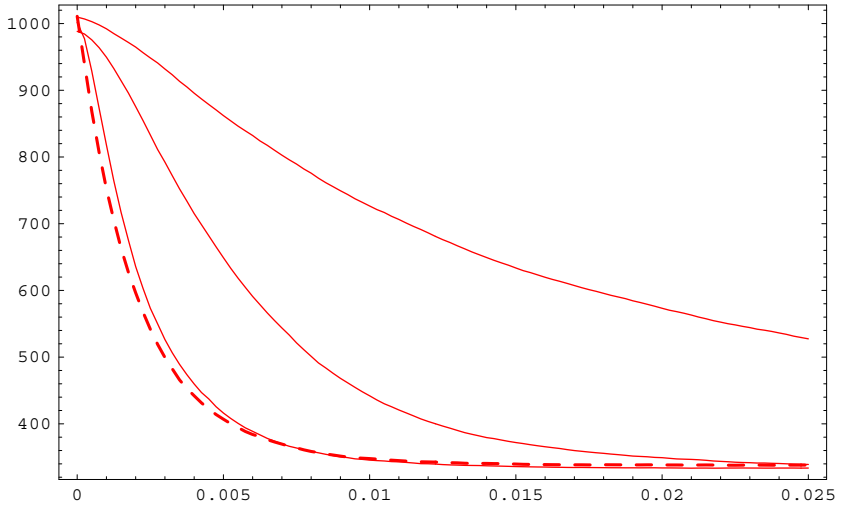

Figure 2: Time relaxation of the second moment $m_{11}(t)$

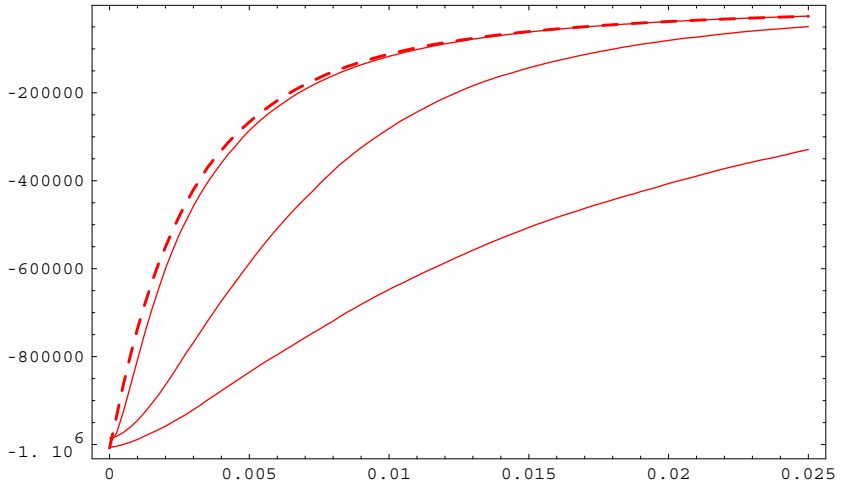

Figure 3: Time relaxation of the third moment $r_1(t)$

In **Figure 3** a similar behaviour is observed for the third moment $r_1(t)$ (cf. (4.5), (4.10), (4.11)). For 100000 particles the standard DSMC method delivers a result close to the dashed line obtained using the TTC method with 1000 particles. All computations have been done using $N = 100$ independent ensembles (cf. (4.15)) for 100000 particles, $N = 1000$ ensembles for 10000 particles and $N = 10000$ repetitions if the number of particles was 1000. The confidence intervals (4.16) are of negligible size in this case.

Next we study the efficiency of the temperature time counter scheme compared with the standard DSMC method. We consider the test case (4.8) on the time interval $[0, 0.1]$ with a particle number $n = 100000$.

The information about the computational time for these methods is contained in the second column of **Table 1**. We assume the computational time needed by the DSMC method as 100% so that the quantities in the table represent the percentage of time needed using TTC with different numbers of shells corresponding to (3.22). The other columns show the numbers of collisions, fictitious collisions (according to (3.14)) and rejections (according to (3.30), (3.33)), respectively (measured in multiples of 10^4).

Table 1

Method	CPU	Collisions	Fict. collisions	Rejections
DSMC	100	500	486	0
TTC/1	3156	23.8	9.20	26300
TTC/2	834	23.8	9.24	6570
TTC/4	246	23.8	9.31	1630
TTC/8	101	23.8	9.30	409
TTC/16	65.5	23.8	9.31	115
TTC/32	56.8	23.8	9.22	44.1
TTC/64	54.7	23.8	9.28	23.2
TTC/128	54.7	23.8	9.29	16.4
TTC/256	55.3	23.8	9.21	13.8

The gain factor in efficiency is about 2, when the number of shells is between 32 and 256. This reduction is due to a decrease in the number of rejections of the selected candidates for the collisions corresponding to the probabilities (3.30) and (3.33) (as shown in the last column). If the number of rejections in the TTC method reaches the number of fictitious collisions in the standard DSMC method (see line TTC/8) the efficiency of both methods is the same. In the case of optimal efficiency (see line TTC/64) only two attempts in the acceptance-rejection technique are necessary to generate one sample. The number of rejections decreases further when the number of shells grows, but this positive effect gradually disappears due to the additional effort required for sorting particles.

Finally, we study the efficiency of both methods in the test case (4.12). The purpose of this test case is to show that the TTC scheme performs comparably in situations that are actually well solved by the standard DSMC method. The results are displayed in **Table 2**. The number of rejections in the temperature time counter scheme is reduced using the shell technique. However, starting from $K = 32$ it reaches a stable value so that no further gain in efficiency can be expected. Note that the number of real collisions if about $20.3 * 10^4$ in both methods.

Table 2

Method	CPU	Collisions	Fict. collisions	Rejections
DSMC	100	51.2	30.9	0
TTC/1	142	22.4	2.16	140
TTC/2	140	22.4	2.16	129
TTC/4	134	22.4	2.16	100
TTC/8	128	22.4	2.17	61.5
TTC/16	121	22.4	2.19	40.1
TTC/32	121	22.4	2.16	34.8
TTC/64	124	22.4	2.18	34.7
TTC/128	129	22.4	2.18	34.0
TTC/256	141	22.4	2.17	33.6

5. Concluding remarks

The temperature time counter scheme introduced in this paper and the standard DSMC (direct simulation Monte Carlo) method have been considered in the unifying framework provided by a Markov process related to the Boltzmann equation. These algorithms contain three main components – the time steps, the distribution of the collision partners, and the probability of fictitious collisions.

The times steps and the probability of fictitious collisions are closely related to each other. If the time step is small then there are many attempts to collide and the probability of fictitious collisions must be large, since the number of real collisions should remain constant. In the experiment represented in Table 1 the number of real collision is about $14 * 10^4$. The relationship (3.16) between the time steps of both methods is

$$\frac{U_{max}}{\sqrt{6\,T}} = \frac{1000}{\sqrt{6 * 334}} \sim 22.3$$

in the first test case (4.8), (4.9). Correspondingly, the number of collisions is about 22 times larger in the standard DSMC method (cf. third column of Table 1). This is the "reserve" that one can use to generate the more difficult index distribution. If the modelling procedure is good enough, some gain in efficiency remains.

There is an interplay between the time steps and the distribution of the collision partners. In the standard DSMC scheme there are small time steps, but a very simple distribution of the collision partners. In the TTC method the time steps are much larger. However, more effort has to be spent in order to generate the distribution of the collision partners. This effort can be influenced (decreases) by an appropriate application of the acceptance-rejection technique, which was done in this paper by introducing the shells. It should be mentioned that the relatively complicated index distribution of the TTC scheme can be implemented easily on a computer. There are further opportunities to increase the efficiency of modelling the index distribution. The shell parameters b_k , $k = 1, \ldots, K-1$, in (3.22) may be adapted to the system during the simulation. The majorant (3.32) for the function (3.31) may be improved. To this end, one could use some ordering with respect to the orientation in the velocity space, to take into account the term $-2\,(v_i - V, v_j - V)$

in the expression

$$\|v_i - v_j\|^2 = \|v_i - V\|^2 - 2\,(v_i - V, v_j - V) + \|v_j - V\|^2\,. \tag{5.1}$$

The particular initial distribution (4.1), (4.8) was chosen to illustrate the advantages of the temperature time counter scheme as compared with the standard DSMC method. This example provides some "extreme" situation, when a relatively small number of particles has a strong influence on the macroscopic quantities. However, the second test (4.12) case shows that both methods perform comparably in "normal" situations.

A remarkable feature of the TTC scheme is the fixed time step during the collision simulation (as in the no time counter scheme in Example 2.2). Furthermore, it has two properties that are favourable compared with the standard DSMC method. The first is the absence of an additional systematic error due to the wrong accumulated maximal relative velocity bound (2.27). This gives better stability with respect to short time perturbations. The second is a better long term performance in cases far from Maxwellian. These features are of importance in applications, when a low number of molecules of high energy influences the behaviour of the gas.

References

[1] H. Babovsky and R. Illner, *A convergence proof for Nanbu's simulation method for the full Boltzmann equation*, SIAM J. Numer. Anal., 26 (1989), pp. 45–65.

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

[3] ——, *Perception of numerical methods in rarefied gas dynamics*, Progr. Astronaut. Aeronaut., 118 (1989), pp. 211–226.

[4] ——, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Clarendon Press, Oxford, 1994.

[5] C. Cercignani, *The Boltzmann Equation and its Applications*, Springer, New York, 1988.

[6] C. Cercignani, R. Illner, and M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Springer, New York, 1994.

[7] S. N. Ethier and T. G. Kurtz, *Markov Processes, Characterization and Convergence*, Wiley, New York, 1986.

[8] M. S. Ivanov and S. V. Rogazinskij, *Comparative analysis of algorithms of the direct statistical modeling method in rarefied gas dynamics*, Zh. Vychisl. Mat. i Mat. Fiz., 28 (1988), pp. 1058–1070. In Russian.

[9] K. Koura, *Null–collision technique in the direct–simulation Monte Carlo method*, Phys. Fluids, 29 (1986), pp. 3509–3511.

[10] A. V. Lukshin and S. N. Smirnov, *An efficient stochastic algorithm for solving the Boltzmann equation*, Zh. Vychisl. Mat. i Mat. Fiz., 29 (1989), pp. 118–124. In Russian.

[11] K. Nanbu, *Interrelations between various direct simulation methods for solving the Boltzmann equation*, J. Phys. Soc. Jpn., 52 (1983), pp. 3382–3388.

[12] H. Neunzert and J. Struckmeier, *Particle methods for the Boltzmann equation*, in Acta Numerica 1995, Cambridge University Press, Cambridge, 1995, pp. 417–457.

[13] *Rarefied Gas Dynamics (J. Harvey and G. Lord, eds.)*, vol. 1 and 2, Oxford University Press, Oxford, 1995.

[14] W. Wagner, *A convergence proof for Bird's direct simulation Monte Carlo method for the Boltzmann equation*, J. Statist. Phys., 66 (1992), pp. 1011–1044.