BURNETT SPECTRAL METHOD FOR HIGH-SPEED RAREFIED GAS FLOWS[*]

ZHICHENG HU[†] AND ZHENNING CAI[‡]

Abstract. We introduce a numerical solver for the spatially inhomogeneous Boltzmann equation using the Burnett spectral method. The modeling and discretization of the collision operator are based on the previous work [Z. Cai, Y. Fan, and Y. Wang, *Comput. & Fluids*, 200 (2020), 104456], which is the hybridization of the BGK operator for higher moments and the quadratic collision operator for lower moments. To ensure the preservation of the equilibrium state, we introduce an additional term to the discrete collision operator, which equals zero when the number of degrees of freedom tends to infinity. Compared with the previous work [Z. Hu, Z. Cai, and Y. Wang, *SIAM J. Sci. Comput.*, 42 (2020), pp. B105–B134], the computational cost is reduced by one order. Numerical experiments such as shock structure calculation and Fourier flows are carried out to show the efficiency and accuracy of our numerical method.

Key words. Boltzmann equation, Burnett spectral method, steady-state preserving

AMS subject classification. 76P05

DOI. 10.1137/19M1294010

1. Introduction. In rarefied gas dynamics, models based on continuum hypothesis such as Euler equations and Navier–Stokes equations cannot provide accurate prediction of the flow structure. To correctly describe the motion of fluids, one needs to employ gas kinetic theory to capture the rarefaction effects. However, switching from continuum models to kinetic models may greatly increase the computational difficulty, since the kinetic theory uses the distribution function to describe the velocity distribution of gas molecules, which doubles the dimensions of the unknown function in the differential equations. Thus, solving kinetic models deterministically has long been considered intractable, and the direct simulation of Monte Carlo (DSMC) method has been playing an important role in the simulation of rarefied gases [6, 7].

Nowadays, due to the fast improvement of the CPU performance, researchers have tried to solve the six-dimensional kinetic equations deterministically [26, 15, 17]. At the same time, a number of new ideas have been proposed to accelerate the numerical solver [18, 37, 16]. In particular, for the Boltzmann equation, which has an additional difficulty due to its highly complicated binary collision term, significant progress on fast algorithms has been made in recent years [36, 2, 20, 21, 3, 27]. These works have shown great promise for practical applications of these deterministic Boltzmann solvers in the near future. Our work also contributes to this research field.

[†]Department of Mathematics, College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (huzhicheng@nuaa.edu.cn).

[‡]Department of Mathematics, National University of Singapore, Level 4, Block S17, 10 Lower Kent Ridge Road, Singapore 119076 (matcz@nus.edu.sg).

In this paper, we are going to introduce a new numerical solver for the spatially inhomogeneous Boltzmann equation.

Currently, the most popular numerical solver for the binary collision term of the Boltzmann equation is likely to be the Fourier spectral method [29, 8] and its variations [22, 28, 19]. However, for general gas molecules, the computational complexity for evaluating the collision term is quadratic in the number of degrees of freedom in the velocity space. Thus, when a large number of Fourier modes are needed to resolve the distribution function (e.g., when the distribution function is discontinuous), this still introduces a large computational cost. In [11, 35, 10], the authors have been trying to reduce the computational cost by restricting the evaluation of the expensive collision term only to a small number of degrees of freedom, so that the computation of the collision term may still be affordable even for a relatively large number of degrees of freedom in the velocity space. To achieve such a goal, Hermite/Burnett polynomials are used instead of Fourier basis functions, so that the idea of BGK-type modeling can be applied. Such a strategy has been verified in the spatially inhomogeneous problems [12, 25], which shows it is a competitive approach to give satisfactory results, although the price to pay, for the expensive part of the collision operator, is a higher computational complexity than the Fourier spectral method with the same number of degrees of freedom. In [10], it has been observed in the numerical solution of the spatially homogeneous Boltzmann equation that the implementation using Burnett polynomials (orthogonal polynomials based on spherical coordinates) is much faster than Hermite polynomials (orthogonal polynomials based on Cartesian coordinates), despite their mathematical equivalence. This work can be considered a test of its performance with the presence of spatial variables.

For the special implementation of the collision term [35, 10], the transition from the spatially homogeneous case to the spatially inhomogeneous case is not as straightforward as the Fourier spectral method. As discussed in [25], the major reason is that the implementation in [35, 10] requires the parameters in the weight function must be, respectively, the local mean velocity and the local temperature of the distribution function, which may not hold for all spatial locations. If these parameters are different from the local mean velocity and the local temperature, then the direct application of the discrete collision term of [10] can only automatically conserve mass, momentum, and energy, but it does not preserve the equilibrium state without additional numerically tricks. In this work, we propose a novel remedy of such a problem, which is more efficient and robust in comparison to the approach proposed in [25]. The method will be tested in several examples in one spatial dimension. The results show both good efficiency and high accuracy.

The rest of this paper is organized as follows. Section 2 is a review of the Boltzmann equation and Burnett's expansion of the distribution function. Our numerical method will be introduced in detail in section 3. Numerical tests will be given in section 4, and we conclude this paper with a brief summary in section 5.

2. Boltzmann equation and the Burnett spectral method. To better describe our numerical algorithm, we would like to first clarify the mathematical model to be solved and the framework of the numerical method. Some relevant previous research works will also be reviewed in this section.

2.1. Boltzmann equation. The governing equation for the rarefied gas flow is the Boltzmann equation:

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} f = Q[f, f],$$

where $\boldsymbol{x} = (x, y, z)^T \in \mathbb{R}^3$ is the spatial variable, $\boldsymbol{v} \in (v_x, v_y, v_z)^T \in \mathbb{R}^3$ is the velocity variable, and $f(t, \boldsymbol{x}, \boldsymbol{v})$ is the distribution function. The right-hand side $Q[f, f]$ is the collision term, which will be detailed later. In this paper, we restrict ourselves to one spatial dimension, so that the Boltzmann equation can be written as

$$(1) \qquad \frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial x} = Q[f, f].$$

With the initial condition

$$(2) \qquad f(0, x, \boldsymbol{v}) = f_0(x, \boldsymbol{v}),$$

(1) holds for any $t > 0$, $x \in I \subset \mathbb{R}$, and $\boldsymbol{v} \in \mathbb{R}^3$, where the interval $I = (a, b)$ can either be finite or infinite. When $a > -\infty$, we consider Maxwell's wall boundary condition at $x = a$. The solid wall at $x = a$ may have a velocity $\boldsymbol{u}_a^W(t)$. In this paper, we assume that the x-component of $\boldsymbol{u}_a^W(t)$ is zero so that the computational domain does not change. Maxwell's wall boundary condition assumes that among all the particles that hit the wall, some particles undergo specular reflection, and others undergo diffusive reflection. The proportion of the particles that undergo diffusive reflection is called the accommodation coefficient χ_a. For a solid wall with temperature $T_a^W(t)$ at time t, the boundary condition can be formulated as

$$(3) \qquad f(t, a, \boldsymbol{v}) = \chi_a f_a^W(t, \boldsymbol{v}) + (1 - \chi_a) f(t, a, \overline{\boldsymbol{v}}) \qquad \text{for} \quad v_x > 0,$$

where $\overline{\boldsymbol{v}} = (-v_x, v_y, v_z)^T$ and

$$(4) \qquad f_a^W(t, \boldsymbol{v}) = \frac{n_a^W(t)}{[2\pi \theta_a^W(t)]^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}_a^W(t)|^2}{2\theta_a^W(t)}\right).$$

In (4), the quantity $\theta_a^W(t)$ is the temperature of the gas represented in the unit of specific energy, defined by

$$\theta_a^W(t) = k_B T_a^W(t)/\mathfrak{m},$$

where k_B is the Boltzmann constant, and $\mathfrak{m}$ is the mass of a single molecule. The quantity $n_a^W(t)$ is chosen such that the "no mass flux" boundary condition is satisfied. Its precise expression is

$$(5) \qquad n_a^W(t) = \sqrt{\frac{2\pi}{\theta_a^W(t)}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x f(t, a, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z.$$

Similarly, if $b < +\infty$ and the solid wall at $x = b$ has velocity $\boldsymbol{u}_b^W(t)$, temperature $\theta_b^W(t)$, and accommodation coefficient χ_b, then the boundary condition at $x = b$ is

$$f(t, b, \boldsymbol{v}) = \frac{\chi_b n_b^W(t)}{[2\pi \theta_b^W(t)]^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}_b^W(t)|^2}{2\theta_b^W(t)}\right) + (1 - \chi_b) f(t, b, \overline{\boldsymbol{v}}) \qquad \text{for} \quad v_x < 0,$$

where $n_b^W(t)$ is given by

$$n_b^W(t) = \sqrt{\frac{2\pi}{\theta_b^W(t)}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} v_x f(t, b, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z.$$

Here we remind readers that the boundary conditions need to be prescribed only for a half of the distribution function which describes the particles moving into the domain I.

The collision term $Q[f,f]$ is the most complicated part in the Boltzmann equation, whose general form is

(6)
$$Q[f,f](t,x,\boldsymbol{v}) = \int_{\mathbb{R}^3} \int_{\mathbb{S}^2_+} B(\boldsymbol{v}-\boldsymbol{v}_1,\boldsymbol{n})[f(t,x,\boldsymbol{v}'_1)f(t,x,\boldsymbol{v}')-f(t,x,\boldsymbol{v}_1)f(t,x,\boldsymbol{v})]\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{v}_1.$$

Here $\boldsymbol{v}'_1$ and $\boldsymbol{v}'$ are postcollisional velocities,

$$\boldsymbol{v}'_1 = \boldsymbol{v}_1 - [(\boldsymbol{v}_1-\boldsymbol{v})\cdot\boldsymbol{n}]\boldsymbol{n}, \qquad \boldsymbol{v}' = \boldsymbol{v} - [(\boldsymbol{v}-\boldsymbol{v}_1)\cdot\boldsymbol{n}]\boldsymbol{n},$$

and $B(\cdot,\cdot)$ is the collision kernel determined by the potential function between gas molecules. It can be seen that for any function $g(\boldsymbol{v})$ defined on the velocity space, the collision term $Q[g,g](\boldsymbol{v})$ can be defined similarly to (6) with t and x removed. For any distribution function $g(\boldsymbol{v})$, the corresponding collision term satisfies the following conservation property,

$$\int_{\mathbb{R}^3} \begin{pmatrix} 1 \\ \boldsymbol{v} \\ \frac{1}{2}|\boldsymbol{v}|^2 \end{pmatrix} Q[g,g](\boldsymbol{v})\,\mathrm{d}\boldsymbol{v} = 0,$$

which indicates the local conservation of mass, momentum, and energy. Another important property is

$$Q[\mathcal{M},\mathcal{M}] = 0 \iff \mathcal{M}(\boldsymbol{v}) = \exp(\alpha + \boldsymbol{\beta}\cdot\boldsymbol{v} + \gamma|\boldsymbol{v}|^2) \text{ and } \gamma < 0.$$

It shows that the manifold of local equilibrium states are formed by isotropic Gaussian distribution functions, which are also called Maxwellians.

In this paper, we will mainly focus on the specific collision models induced by the inverse power laws, in which the force between two molecules is always repulsive, and the magnitude is proportional to a certain negative power of the distance between them ($F = \kappa r^{-\eta}$ with F, r, η, and κ being the force, distance, power index, and intensity constant, respectively). By choosing a reference velocity $\overline{\boldsymbol{u}}$ and reference temperature $\overline{T}$, it has been demonstrated in [25] that the collision term for inverse power laws can be nondimensionalized by

(7)
$$Q[g,g](\overline{\boldsymbol{u}}+\sqrt{\overline{\theta}}\boldsymbol{v}^*) = \frac{\overline{\rho}^2}{\mathfrak{m}^2\overline{\theta}^{\frac{3}{2}}} \left(\frac{2\kappa}{\mathfrak{m}}\right)^{\frac{2}{\eta-1}} \overline{\theta}^{\frac{\eta-5}{2(\eta-1)}} Q^*[g^*,g^*](\boldsymbol{v}^*),$$

where $\overline{\theta} = k_B\overline{T}/\mathfrak{m}$, and the density of the gas $\overline{\rho}$ as well as the nondimensionalized distribution function g^* are given by

(8)
$$\overline{\rho} = \mathfrak{m}\int_{\mathbb{R}^3} g(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}, \qquad g^*(\boldsymbol{v}^*) = \frac{\mathfrak{m}\overline{\theta}^{3/2}}{\overline{\rho}}g(\overline{\boldsymbol{u}}+\sqrt{\overline{\theta}}\boldsymbol{v}^*).$$

In (7), the dimensionless collision operator Q^* is independent of κ and $\mathfrak{m}$. The only parameter in Q^* is the power index η. This formula can be further simplified by introducing the reference viscosity coefficient (see (36) in the appendix for the definition of $A_2(\eta)$),

(9)
$$\overline{\mu} = \frac{5\mathfrak{m}(\overline{\theta}/\pi)^{1/2}(2\mathfrak{m}\overline{\theta}/\kappa)^{2/(\eta-1)}}{8A_2(\eta)\Gamma(4-2(\eta-1))},$$

by which we find that

$$Q[g,g](\overline{\boldsymbol{u}}+\sqrt{\overline{\theta}}\boldsymbol{v}^*) = \frac{\overline{\rho}^2}{\mathfrak{m}\overline{\mu}\sqrt{\overline{\theta}}} \hat{Q}[g^*,g^*](\boldsymbol{v}^*),$$

where $\hat{Q}$ is the new dimensionless collision operator given by

$$\hat{Q}[g^*, g^*] = \frac{5}{4^{\frac{3}{2} - \frac{2}{\eta-1}} \sqrt{\pi} A_2(\eta) \Gamma(4 - 2(\eta-1))} Q^*[g^*, g^*].$$

As will be seen later, such a nondimensionalization is closely related to the choice of parameters in our numerical scheme.

Remark. Here we have introduced two dimensionless collision terms $Q^*[f^*, f^*]$ and $\hat{Q}[f^*, f^*]$, which differ only by a constant. In [10, 35], the dimensionless collision term $Q^*[f^*, f^*]$ is used, while in this work, we are going to use $\hat{Q}[f^*, f^*]$ in our further discussion. Since our numerical method is built based on the work [10], we point out the difference here to avoid confusion.

2.2. Burnett spectral method for the spatially homogeneous Boltzmann equation. For the spatially homogeneous Boltzmann equation, the Burnett spectral method has been introduced in [10], where the Burnett method is introduced in the dimensionless setting. In what follows, we will provide a brief review of this method. Since the flow is assumed to be spatially homogeneous, the variable x will be omitted temporarily in this subsection.

We will present the method based on the dimensionless collision term $\hat{Q}[f^*, f^*]$, where f^* is the dimensionless distribution function defined similar to (8):

$$(10) \qquad f^*(t^*, \boldsymbol{v}^*) = \frac{\mathfrak{m}\overline{\theta}^{3/2}}{\overline{\rho}} f\left(\frac{\overline{\mu}t^*}{\overline{\rho}\overline{\theta}}, \overline{\boldsymbol{u}} + \sqrt{\overline{\theta}}\boldsymbol{v}^*\right), \qquad \overline{\rho} = \mathfrak{m}\int_{\mathbb{R}^3} f(t, \boldsymbol{v})\,\mathrm{d}\boldsymbol{v}.$$

Note that $\overline{\rho}$ is independent of t since collision does not change the number density. To write down the spectral expansion, we first define the Burnett polynomials [9]:

$$(11) \qquad p_{lmn}(\boldsymbol{v}^*) = \sqrt{\frac{2^{1-l}\pi^{3/2}n!}{\Gamma(n+l+3/2)}} L_n^{(l+1/2)}\left(\frac{|\boldsymbol{v}^*|^2}{2}\right) |\boldsymbol{v}^*|^l Y_l^m\left(\frac{\boldsymbol{v}^*}{|\boldsymbol{v}^*|}\right),$$
$$l, n = 0, 1, \ldots, \quad m = -l, \ldots, l.$$

where $L_n^{(\alpha)}(\cdot)$ and $Y_l^m(\cdot)$ are, respectively, the Laguerre polynomials and spherical harmonics, whose definitions are given in detail in the appendix (see (37)). Let $\omega(\boldsymbol{v}^*)$ be the three-dimensional standard normal distribution

$$\omega(\boldsymbol{v}^*) = \frac{1}{(2\pi)^{3/2}} \exp\left(-\frac{|\boldsymbol{v}^*|^2}{2}\right).$$

Then the following orthogonality holds:

$$\int_{\mathbb{R}^3} [p_{l_1 m_1 n_1}(\boldsymbol{v}^*)]^\dagger p_{l_2 m_2 n_2}(\boldsymbol{v}^*)\omega(\boldsymbol{v}^*)\,\mathrm{d}\boldsymbol{v}^* = \delta_{l_1 l_2}\delta_{m_1 m_2}\delta_{n_1 n_2},$$

where $\dagger$ denotes the complex conjugate. The Petrov–Galerkin spectral method can be derived by approximating the dimensionless distribution function $f^*(t^*, \boldsymbol{v}^*)$ by

$$f_M^*(t^*, \boldsymbol{v}^*) = \sum_{l=0}^{M} \sum_{m=-l}^{l} \sum_{n=0}^{\lfloor (M-l)/2 \rfloor} \tilde{f}_{lmn}^*(t^*)p_{lmn}(\boldsymbol{v}^*)\omega(\boldsymbol{v}^*),$$

and then obtaining equations for the coefficients by calculating

$$(12) \quad \int_{\mathbb{R}^3} [p_{lmn}(\boldsymbol{v}^*)]^\dagger \frac{\partial f_M^*(t^*, \boldsymbol{v}^*)}{\partial t^*} \, \mathrm{d}\boldsymbol{v}^* = \int_{\mathbb{R}^3} [p_{lmn}(\boldsymbol{v}^*)]^\dagger \hat{Q}[f_M^*, f_M^*](t^*, \boldsymbol{v}^*) \, \mathrm{d}\boldsymbol{v}^*,$$
$$l = 0, \ldots, M, \qquad m = -l, \ldots, l, \qquad n = 0, \ldots, \lfloor (M-l)/2 \rfloor.$$

The general result is
(13)
$$\frac{\mathrm{d}\tilde{f}_{lmn}^*}{\mathrm{d}t^*} = \sum_{l_1=0}^{M} \sum_{m_1=-l_1}^{l_1} \sum_{n_1=0}^{\lfloor (M-l_1)/2 \rfloor} \sum_{l_2=0}^{M} \sum_{m_2=-l_2}^{l_2} \sum_{n_2=0}^{\lfloor (M-l_2)/2 \rfloor} A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2} \tilde{f}_{l_1 m_1 n_1}^* \tilde{f}_{l_2 m_2 n_2}^*,$$

where the constant coefficients $A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2}$ depend on the collision model. In [35, 10], the authors introduced an algorithm computing these coefficients for all inverse power law models, which is briefly described in the appendix. The computational cost of (13) with all l, m, n is $O(M^8)$ since $A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2} = 0$ when $m \neq m_1 + m_2$; and the discretization (13) automatically conserves mass, momentum, and energy.

Our numerical method is based on the discretization (13). However, such discretization is not readily applicable for the spatially inhomogeneous case. The main reason is that (13) does not preserve the equilibrium state. In detail, when $f^*(0, \boldsymbol{v}^*) = \exp(\alpha + \boldsymbol{\beta} \cdot \boldsymbol{v}^* + \gamma |\boldsymbol{v}^*|^2)$, after computing $f_M^*(0, \boldsymbol{v}^*)$ by projection, the right-hand side of (13) is nonzero. As mentioned in the introduction, one of the main contributions of this paper is to fix such a problem.

3. Numerical method. In this section, our numerical method will be provided in detail. To begin with, we will resume the discussion at the end of section 2.2 and develop an algorithm which preserves the steady state.

3.1. Restoring the dimension and preserving the steady state. In this section, we will still keep the spatial variable x omitted and focus on the homogeneous Boltzmann equation. From (12), it can be seen that the spectral method in section 2.2 works only when $f^*(t^*, \cdot) \in L^2(\mathbb{R}^3; \omega^{-1} \, \mathrm{d}\boldsymbol{v}^*)$ for all t^*, i.e.,

$$\int_{\mathbb{R}^3} |f^*(t^*, \boldsymbol{v}^*)|^2 [\omega(\boldsymbol{v}^*)]^{-1} \, \mathrm{d}\boldsymbol{v}^* < +\infty \qquad \forall t^* \geqslant 0.$$

By (10), it can be seen that the original distribution function $f(t, \boldsymbol{v})$ must satisfy

$$(14) \qquad \int_{\mathbb{R}^3} |f(t, \boldsymbol{v})|^2 \left[\omega \left(\frac{\boldsymbol{v} - \overline{\boldsymbol{u}}}{\sqrt{\overline{\theta}}} \right) \right]^{-1} \, \mathrm{d}\boldsymbol{v} < +\infty \qquad \forall t \geqslant 0.$$

This equation shows that the parameter $\overline{\theta}$ is not only a parameter in the nondimensionalization, but also plays the role of the scaling factor in the spectral method for problems on unbounded domains [32]. It is easy to see that when $\overline{\theta}$ is larger, (14) allows more distribution functions, i.e., $L^2(\mathbb{R}^3; \omega_2^{-1} \, \mathrm{d}\boldsymbol{v}) \subset L^2(\mathbb{R}^3; \omega_1^{-1} \, \mathrm{d}\boldsymbol{v})$ if $\overline{\theta}_1 > \overline{\theta}_2$, where

$$\omega_{1,2}(\boldsymbol{v}) = \omega \left((\boldsymbol{v} - \overline{\boldsymbol{u}}) / \sqrt{\overline{\theta}_{1,2}} \right).$$

In particular, the Maxwellian $\omega_1(\boldsymbol{v}) \in L^2(\mathbb{R}^3; \omega_1^{-1} \, \mathrm{d}\boldsymbol{v})$ will be in $L^2(\mathbb{R}^3; \omega_2^{-1} \, \mathrm{d}\boldsymbol{v})$ only when $\overline{\theta}_1 < 2\overline{\theta}_2$. Therefore, we need to choose a sufficiently large $\overline{\theta}$ to include all possible distribution functions. The existence of such $\overline{\theta}$ has been theoretically guaranteed in [4].

Remark. In the present work, $\overline{\theta}$ is a constant independent of t and $\boldsymbol{x}$. Its choice is based on an a priori understanding of the problem and shall be given in section 4 for each numerical test. With constant $\overline{\theta}$, it turns out that the spatial discretization of an advection term would be simplified a lot, as can be seen in section 3.3. Nevertheless, it is worth mentioning that an adaptive choice of $\overline{\theta}$, which is spatially and temporally dependent, may be more interesting, as it has some additional advantages, especially when $\overline{\theta}$ is selected as the local temperature [30]. However, the adaptive choice of $\overline{\theta}$ may also have a convergence issue [13] and needs to be further studied in a future work.

To emphasize the role of $\overline{\theta}$ in our algorithm, we will present our algorithm using the original distribution function $f(t, \boldsymbol{v})$. Thus the approximate distribution function is

(15)
$$f_M(t, \boldsymbol{v}) = \sum_{l=0}^{M} \sum_{m=-l}^{l} \sum_{n=0}^{\lfloor (M-l)/2 \rfloor} \tilde{f}_{lmn}(t) p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \omega^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}),$$

where
(16)
$$p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) = \overline{\theta}^{-(l+2n)/2} p_{lmn}\left(\frac{\boldsymbol{v} - \overline{\boldsymbol{u}}}{\sqrt{\overline{\theta}}}\right), \qquad \omega^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) = \frac{1}{\mathfrak{m}(2\pi\overline{\theta})^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \overline{\boldsymbol{u}}|^2}{2\overline{\theta}}\right).$$

The discretization (13) becomes

$$\frac{\mathrm{d}\tilde{f}_{lmn}}{\mathrm{d}t} = \frac{\overline{\theta}}{\overline{\mu}} \sum_{l_1=0}^{M} \sum_{m_1=-l_1}^{l_1} \sum_{n_1=0}^{\lfloor (M-l_1)/2 \rfloor} \sum_{l_2=0}^{M} \sum_{m=-l_2}^{l_2} \sum_{n_2=0}^{\lfloor (M-l_2)/2 \rfloor}$$
$$A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2} \overline{\theta}^{n-n_1-n_2+(l-l_1-l_2)/2} \tilde{f}_{l_1 m_1 n_1} \tilde{f}_{l_2 m_2 n_2}.$$

For simplicity, the right-hand side of the above equation will be named $\tilde{Q}_{lmn}[f_M]$ hereafter.

Now we are going to change the right-hand side of the above scheme such that the method preserves the steady state. When the distribution function $f(t, \boldsymbol{v})$ is a Maxwellian $\mathcal{M}(t, \boldsymbol{v})$, the Maxwellian can be determined by the first few coefficients in the series expansion:

$$\mathcal{M}(t, \boldsymbol{v}) = \frac{\rho(t)}{\mathfrak{m}[2\pi\theta(t)]^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}(t)|^2}{2\theta(t)}\right),$$

where

$$\rho(t) = \tilde{f}_{000}(t), \quad \boldsymbol{u}(t) = \overline{\boldsymbol{u}}(t) + \left(\frac{\tilde{f}_{100}(t)}{\tilde{f}_{000}(t)}, \quad -\sqrt{2}\,\mathrm{Re}\,\frac{\tilde{f}_{110}(t)}{\tilde{f}_{000}(t)}, \quad \sqrt{2}\,\mathrm{Im}\,\frac{\tilde{f}_{110}(t)}{\tilde{f}_{000}(t)}\right)^T,$$
$$\theta(t) = \overline{\theta} - \sqrt{\frac{2}{3}} \frac{\tilde{f}_{001}(t)}{\tilde{f}_{000}(t)} - \frac{|\tilde{f}_{100}(t)|^2 + 2|\tilde{f}_{110}(t)|^2}{3|\tilde{f}_{000}(t)|^2}.$$

Since $Q[\mathcal{M}, \mathcal{M}]$ always equals zero, one can rewrite the collision term as $Q[f, f] - Q[\mathcal{M}, \mathcal{M}]$. Thereby, the corresponding discretization turns out to be

(17)
$$\frac{\mathrm{d}\tilde{f}_{lmn}}{\mathrm{d}t} = \tilde{Q}_{lmn}[f_M] - \tilde{Q}_{lmn}[\mathcal{M}_M],$$

where $\mathcal{M}_M$ is the projection of the Maxwellian:

$$\mathcal{M}_M(t, \boldsymbol{v}) = \sum_{l=0}^{M} \sum_{m=-l}^{l} \sum_{n=0}^{\lfloor (M-l)/2 \rfloor} \tilde{\mathcal{M}}_{lmn}(t) p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \omega^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}),$$

and the coefficients $\tilde{\mathcal{M}}_{lmn}(t)$ can be obtained by the following theorem.

THEOREM 1. *For $\rho, \theta > 0$ and $\boldsymbol{u} = (u_x, u_y, u_z) \in \mathbb{R}^3$, let*

(18)
$$\mathcal{M}(\boldsymbol{v}) = \frac{\rho}{\mathfrak{m}(2\pi\theta)^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}|^2}{2\theta}\right).$$

If $\overline{\theta} > \theta/2$, then for any $\overline{\boldsymbol{u}} = (\bar{u}_x, \bar{u}_y, \bar{u}_z) \in \mathbb{R}^3$, there exist coefficients $\tilde{\mathcal{M}}_{lmn}$ such that

(19)
$$\mathcal{M}(\boldsymbol{v}) = \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} \sum_{n=0}^{+\infty} \tilde{\mathcal{M}}_{lmn} p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \omega^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v})$$

holds in $L^2(\mathbb{R}^3, [\omega^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v})]^{-1} \, d\boldsymbol{v})$, and when $n \geqslant 1$, the coefficients satisfy the following recursive formula:

(20)
$$\tilde{\mathcal{M}}_{lmn} = \frac{1}{\sqrt{n}\gamma_{l+1,m}} \Big(\sqrt{n+l+3/2}\gamma_{l+2,m}\tilde{\mathcal{M}}_{l+2,m,n-1} + (\bar{u}_x - u_x)\tilde{\mathcal{M}}_{l+1,m,n-1}$$
$$+ \sqrt{n+l+1/2}\gamma_{l+1,m}(\overline{\theta} - \theta)\tilde{\mathcal{M}}_{l,m,n-1} - \sqrt{n-1}\gamma_{l+2,m}(\overline{\theta} - \theta)\tilde{\mathcal{M}}_{l+2,m,n-2} \Big),$$

where the last term $\tilde{\mathcal{M}}_{l+2,m,n-2}$ is regarded as zero when $n = 1$, and the γ symbol is defined by

(21)
$$\gamma_{lm} = \sqrt{\frac{2(l-m)(l+m)}{(2l+1)(2l-1)}}.$$

When $n = 0$ and $|m| < l$, the recurrence formula is
(22)
$$\tilde{\mathcal{M}}_{lm0} = \frac{1}{\sqrt{l+1/2}\gamma_{lm}} \left((u_x - \bar{u}_x)\tilde{\mathcal{M}}_{l-1,m,0} - \sqrt{\frac{1}{l-1/2}}\gamma_{l-1,m}\frac{|\boldsymbol{u} - \overline{\boldsymbol{u}}|^2}{2}\tilde{\mathcal{M}}_{l-2,m,0} \right),$$

where $\tilde{\mathcal{M}}_{l-2,m,0}$ is regarded as zero if $|m| = l - 1$. When $n = 0$ and $m = \pm l$, we have

(23)
$$\tilde{\mathcal{M}}_{ll0} = \sqrt{\frac{1}{2^l l!}}\rho[(\bar{u}_y - u_y) - \mathrm{i}(\bar{u}_z - u_z)]^l, \qquad \tilde{\mathcal{M}}_{l,-l,0} = (-1)^l \tilde{\mathcal{M}}_{ll0}^{\dagger}.$$

The proof of this theorem is to be found in the appendix. By this theorem, we see that the computational cost for every coefficient is $O(1)$. Therefore the computational complexity for evaluating all the coefficients $\tilde{\mathcal{M}}_{lmn}$ with $l + 2n \leqslant M$ is $O(M^3)$. The detailed algorithm is as follows.

Algorithm   Calculation of $\tilde{\mathcal{M}}_{lmn}$ for $l + 2n \leqslant M$.

1: **for** m from 0 to M **do**
2:       Compute $\tilde{\mathcal{M}}_{mm0}$ from (23)
3:       **for** l from $m + 1$ to M **do**
4:             Compute $\tilde{\mathcal{M}}_{lm0}$ from (22)
5:       **end for**
6:       **for** deg from m to M **do**
7:             **for** n from 1 to $\lfloor (deg - m)/2 \rfloor$ **do**
8:                   $l \leftarrow deg - 2n$
9:                   Compute $\tilde{\mathcal{M}}_{lmn}$ from (20)
10:             **end for**
11:       **end for**
12: **end for**
13: **for** m from $-M$ to -1 **do**
14:       **for** l from $|m|$ to M **do**
15:             **for** n from 0 to $\lfloor (M - l)/2 \rfloor$ **do**
16:                   $\tilde{\mathcal{M}}_{lmn} \leftarrow (-1)^m \tilde{\mathcal{M}}_{l,-m,n}^{\dagger}$
17:             **end for**
18:       **end for**
19: **end for**

The above algorithm gives a working order of computation to ensure that when the formulas in Theorem 1 are applied, no recursion is needed.

Obviously, the right-hand side of (17) vanishes when $f(t, \boldsymbol{v}) = \mathcal{M}(t, \boldsymbol{v})$, which indicates that the discretization of the collision term based on the equivalent expression $Q[f, f] - Q[\mathcal{M}, \mathcal{M}]$ preserves the steady state exactly. In fact, it is noted that a similar approach has been used to preserve the steady state in Fourier spectral method [19].

3.2. Modeling of the collision term. A complete algorithm has been described in the above subsection for the discretization of the collision term. However, the computational complexity for this algorithm is as high as $O(M^8)$, due to the nine indices appearing in the coefficients $A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2}$ and the constraint $m = m_1 + m_2$. This makes the simulation difficult when M is large. To reduce the computational cost, it has been proposed in [11, 35, 10] to introduce the "BGK modeling technique" to the collision, meaning that we only apply the quadratic collision to lower moments, which are considered to be important in the numerical computation, while for higher moments, we model their evolution by letting them decay to their equilibrium values at a constant rate. For monatomic gases, the famous BGK model [5] can be derived from the linearized collision model using such an idea. As discussed in [11, section 5.2], if the linearized collision model is applied only up to the second moments (stress tensor), and other higher moments are set as simple convergence to their equilibrium values, we can obtain the BGK model. Another well-known model is the Shakhov model [31], which adds heat fluxes to the part modeled by the linearized collision operator.

In this work, we are going to adopt the same technique when M is large. Assume that we want to apply quadratic modeling for the first M_0th moments, where M_0 is chosen as a constant positive integer less than or equal to M. Thus the spatially homogeneous Boltzmann equation is discretized by

$$(24) \qquad \frac{\mathrm{d}\tilde{f}_{lmn}}{\mathrm{d}t} = \overline{Q}_{lmn}^{M_0} := \begin{cases} \tilde{Q}_{lmn}^{M_0}[f_{M_0}] - \tilde{Q}_{lmn}^{M_0}[\mathcal{M}_{M_0}] & \text{if } l + 2n \leqslant M_0, \\ \nu(\tilde{\mathcal{M}}_{lmn} - \tilde{f}_{lmn}) & \text{if } l + 2n > M_0. \end{cases}$$

In the equations for $\tilde{f}_{lmn}$ with $l + 2n > M_0$, the coefficient ν indicates the rate of convergence to the equilibrium value $\tilde{\mathcal{M}}_{lmn}$, which is chosen following [11, 25] as

$$\nu = \frac{\rho\theta}{\overline{\mu}} \left(\frac{\overline{\theta}}{\theta}\right)^{\frac{1}{2}+\frac{2}{\eta-1}} \varrho_{M_0}.$$

To define ϱ_{M_0} in the above equation, we first define a sequence of matrices $A^l = (a_{nn'}^l) \in \mathbb{R}^{(N_l+1)\times(N_l+1)}$ for $l = 0, 1, \ldots, M_0$, where

$$N_l = \lfloor (M_0 - l)/2 \rfloor, \qquad a_{nn'}^l = A_{l0n}^{000,l0n'} + A_{l0n}^{l0n',000}, \quad n, n' = 0, 1, \ldots, N_l.$$

Thus the definition of ϱ_{M_0} is

$$\varrho_{M_0} = \max\{\varrho(A^l) \mid l = 0, 1, \ldots, M_0\},$$

where $\varrho(\cdot)$ is the spectral radius of the matrix. As is detailed in [11], such a ν is in fact the spectral radius of truncated linearized collision operator.

The total computational cost for (24) is $O(M_0^8 + M^3)$, which is obviously an improvement of the authors' previous work [25] using Hermite polynomials and a different technique to preserve the steady state, where the computational cost was $O(M_0^9 + M^4)$. Moreover, by avoiding two projections in (3.32) of [25], one of which may diverge if θ is far away from $\overline{\theta}$, it turns out that the above approach to preserve the steady state is more robust than the approach used in [25].

3.3. Adding back the spatial variable. From this section, we are going to recover the spatial variable x. Thus in (15), the function f_M on the left-hand side and the coefficients $\tilde{f}_{lmn}$ on the right-hand side should contain the parameter x. To discretize the advection term, we just need to compute $v_x \partial_x f_M$:

$$(25) \qquad v_x \frac{\partial f_M(t,x,\boldsymbol{v})}{\partial x} = \sum_{l=0}^{M} \sum_{m=-l}^{l} \sum_{n=0}^{\lfloor (M-l)/2 \rfloor} \frac{\partial \tilde{f}_{lmn}(t,x)}{\partial x} \left[v_x p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right] \omega^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}),$$

where the term in the square bracket can be expanded by

(26)
$$v_x p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})$$
$$= \overline{u}_x p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) + \overline{\theta} \left(\sqrt{n+l+3/2}\gamma_{l+1,m} p_{l+1,m,n}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) - \sqrt{n+1}\gamma_{-l,m} p_{l-1,m,n+1}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right)$$
$$+ \sqrt{n+l+1/2}\gamma_{-l,m} p_{l-1,m,n}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) - \sqrt{n}\gamma_{l+1,m} p_{l+1,m,n-1}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}).$$

By now, we can combine (15), (25), and (26) to get the complete semidiscrete equations:

$$\frac{\partial \tilde{f}_{lmn}}{\partial t} + \overline{u}_x \frac{\partial \tilde{f}_{lmn}}{\partial x} + \overline{\theta} \left(\sqrt{n+l+1/2}\gamma_{lm} \frac{\partial \tilde{f}_{l-1,m,n}}{\partial x} - \sqrt{n}\gamma_{-l-1,m} \frac{\partial \tilde{f}_{l+1,m,n-1}}{\partial x} \right)$$
$$+ \sqrt{n+l+3/2}\gamma_{-l-1,m} \frac{\partial \tilde{f}_{l+1,m,n}}{\partial x} - \sqrt{n+1}\gamma_{lm} \frac{\partial \tilde{f}_{l-1,m,n+1}}{\partial x} = \overline{Q}_{lmn}^{M_0},$$
$$l = 0, 1, \ldots, M, \quad m = -l, \ldots, l, \quad n = 0, 1, \ldots, \lfloor (M-l)/2 \rfloor.$$

Here $\tilde{f}_{l'm'n'}$ is regarded as zero when its indices are not in the range given by the last line of the above equations. Let $\boldsymbol{f}$ denote the vector whose components are all the coefficients $\tilde{f}_{lmn}$ appearing in the expansion of f_M. Then the above equations can be written as

$$(27) \qquad \frac{\partial \boldsymbol{f}}{\partial t} + \mathbf{A}\frac{\partial \boldsymbol{f}}{\partial x} = \boldsymbol{Q}(\boldsymbol{f}),$$

where $\mathbf{A}$ is a sparse constant matrix whose diagonal entries are $\overline{u}_x$. And each row of $\mathbf{A}$ has at most five nonzero entries.

To complete the problem, we need to add initial and boundary conditions for (27). The initial condition can be obtained by expanding (2) into a series. Alternatively, we can use the orthogonality of basis functions to write down the initial condition as

$$\tilde{f}_{lmn}(0,x) = \mathfrak{m}\overline{\theta}^{l+2n}\int_{\mathbb{R}^3}\left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\right]^{\dagger}f_0(x,\boldsymbol{v})\,\mathrm{d}\boldsymbol{v},$$

$$l = 0, 1, \ldots, M, \quad m = -l, \ldots, l, \quad n = 0, 1, \ldots, \left\lfloor\frac{M-l}{2}\right\rfloor.$$

When the solid wall exists in the problem, we need to formulate wall boundary conditions for (27), which will be detailed in the next subsection.

3.4. Wall boundary conditions. We only consider the wall boundary condition of (27) at $x = a$. The basic idea is the same as the construction of the initial condition. We multiply (3) by

$$\mathfrak{m}\overline{\theta}^{l+2n}\left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\right]^{\dagger}$$

and integrate with respect to $\boldsymbol{v}$. Note that (3) holds only for $v_x > 0$, and therefore the integral domain is the half space:

(28)
$$\mathfrak{m}\overline{\theta}^{l+2n}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{0}^{+\infty}\left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\right]^{\dagger}f_M(t,a,\boldsymbol{v})\,\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z$$
$$= \mathfrak{m}\overline{\theta}^{l+2n}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{0}^{+\infty}\left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\right]^{\dagger}\left(\chi_a f_a^W(t,a,\boldsymbol{v}) + (1-\chi_a)f_M(t,a,\overline{\boldsymbol{v}})\right)\,\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z.$$

Here, the distribution function $f(t,a,\boldsymbol{v})$ in (3) has been replaced by the numerical solution $f_M(t,a,\boldsymbol{v})$, and the "wall Maxwellian" $f_a^W(t,\boldsymbol{v})$ is still defined by (4), while in the definition of $n_a^W(t)$ (5), the distribution function $f(t,a,\boldsymbol{v})$ should be again replaced by $f_M(t,a,\boldsymbol{v})$. However, (28) does not complete the statement of the boundary conditions, since if (28) with all l,m,n satisfying $l+2n \leqslant M$ are imposed at $x = a$, the number of boundary conditions will exceed the number required by the hyperbolicity, which should be the number of characteristics pointing into the domain. To fix the issue, we first choose $\overline{\boldsymbol{u}}$ such that $\overline{u}_x = 0$. Thus, as in [23, 12, 25], we can get the correct number of boundary conditions if we only take into account (28) for l,m,n satisfying $p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) = -p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\overline{\boldsymbol{v}})$ and $l+2n \leqslant M$. The symmetry condition requires that $l+m$ must be odd. Below we are going to omit the spatial variable, which is fixed at $x = a$.

To make the boundary conditions (28) more explicit, we adopt the idea in [34] to split the distribution function into an odd part and an even part:

(29)

$$f_M^{(\text{odd})}(t, \boldsymbol{v}) = \frac{f_M(\boldsymbol{v}) - f_M(\overline{\boldsymbol{v}})}{2} = \sum_{l=0}^{M} \sum_{\substack{m=-l \\ l+m \text{ odd}}}^{l} \sum_{n=0}^{\lfloor (M-l)/2 \rfloor} \tilde{f}_{lmn}(t) p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \omega^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}),$$

(30)

$$f_M^{(\text{even})}(t, \boldsymbol{v}) = \frac{f_M(\boldsymbol{v}) + f_M(\overline{\boldsymbol{v}})}{2} = \sum_{l=0}^{M} \sum_{\substack{m=-l \\ l+m \text{ even}}}^{l} \sum_{n=0}^{\lfloor (M-l)/2 \rfloor} \tilde{f}_{lmn}(t) p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \omega^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}).$$

Thus the boundary condition (28) can be rewritten as

(31)

$$\mathfrak{m} \overline{\theta}^{l+2n} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left[p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} f_M^{(\text{odd})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z$$

$$= \frac{\chi_a \mathfrak{m} \overline{\theta}^{l+2n}}{2 - \chi_a} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left[p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \left(f_a^W(t, \boldsymbol{v}) - f_M^{(\text{even})}(t, \boldsymbol{v}) \right) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z,$$

$l + m$ is odd.

Further simplification requires the following result.

THEOREM 2. *Suppose $l + m$ is odd. Then*

$$\mathfrak{m} \overline{\theta}^{l+2n} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left[p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} f_M^{(\text{odd})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z = \frac{1}{2} \tilde{f}_{lmn}(t).$$

Proof. Since $l+m$ is odd, we have $p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) = -p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\overline{\boldsymbol{v}})$. Noting that $f_M^{(\text{odd})}(t, \boldsymbol{v}) = -f_M^{(\text{odd})}(t, \overline{\boldsymbol{v}})$, we obtain by a change of variables that

$$\mathfrak{m} \overline{\theta}^{l+2n} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left[p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} f_M^{(\text{odd})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z$$

(32)

$$= \mathfrak{m} \overline{\theta}^{l+2n} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} \left[p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\overline{\boldsymbol{v}}) \right]^{\dagger} f_M^{(\text{odd})}(t, \overline{\boldsymbol{v}}) \, \mathrm{d}\overline{v}_x \, \mathrm{d}\overline{v}_y \, \mathrm{d}\overline{v}_z$$

$$= \mathfrak{m} \overline{\theta}^{l+2n} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} \left[p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} f_M^{(\text{odd})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z.$$

Let J_{lmn} be the above quantity. Then by adding up the first and third lines in (32), we obtain

$$2 J_{lmn} = \mathfrak{m} \overline{\theta}^{l+2n} \int_{\mathbb{R}^3} \left[p_{lmn}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} f_M^{(\text{odd})}(t, \boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} = \tilde{f}_{lmn}(t),$$

which implies the conclusion of the theorem.   □

The above theorem gives the left-hand side of (31). To proceed, we first consider a special case $(l, m, n) = (1, 0, 0)$. In this case, $p_{100}^{[\overline{\boldsymbol{u}}, \overline{\theta}]}(\boldsymbol{v}) = v_x/\overline{\theta}$, and the right-hand

side of (31) can be computed by

$$\frac{\chi_a \mathfrak{m}}{2 - \chi_a} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} v_x \left(f_a^W(t, \boldsymbol{v}) - f_M^{(\mathrm{even})}(t, \boldsymbol{v}) \right) \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z$$

$$= \frac{\chi_a \mathfrak{m}}{2 - \chi_a} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x \left(f_M^{(\mathrm{even})}(t, \boldsymbol{v}) - f_a^W(t, \boldsymbol{v}) \right) \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z$$

$$= \frac{\chi_a \mathfrak{m}}{2 - \chi_a} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x \left(f_M(\boldsymbol{v}) - f_M^{(\mathrm{odd})}(t, \boldsymbol{v}) - f_a^W(t, \boldsymbol{v}) \right) \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z$$

$$= \frac{\chi_a \mathfrak{m}}{2 - \chi_a} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x \left(f_M(t, \boldsymbol{v}) - f_a^W(t, \boldsymbol{v}) \right) \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z - \frac{\chi_a}{2 - \chi_a} \frac{1}{2} \tilde{f}_{100}(t).$$

Here the first equality uses the symmetry of f_a^W and $f_M^{(\mathrm{even})}$; the second equality uses the decomposition of f_M; and the third equality uses Theorem 2 and the symmetry of $f_M^{(\mathrm{odd})}$. Now, by using (4) and (5), straightforward calculation yields

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x \left(f_M(t, \boldsymbol{v}) - f_a^W(t, \boldsymbol{v}) \right) \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z = 0.$$

Thus the boundary condition (31) for $(l, m, n) = (1, 0, 0)$ turns out to be

$$\frac{1}{2} \tilde{f}_{100}(t) = -\frac{\chi_a}{2 - \chi_a} \frac{1}{2} \tilde{f}_{100}(t),$$

which is equivalent to

$$\tilde{f}_{100}(t) = 0.$$

Such a result agrees with the requirement that the perpendicular momentum or velocity must equal zero. By this result, we also know that

$$n_a^W(t) = \sqrt{\frac{2\pi}{\theta_a^W(t)}} \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x f_M^{(\mathrm{odd})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z \right.$$

$$\left. + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x f_M^{(\mathrm{even})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z \right)$$

$$= \sqrt{\frac{2\pi}{\theta_a^W(t)}} \left(\frac{1}{2\mathfrak{m}} \tilde{f}_{100}(t) + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x f_M^{(\mathrm{even})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z \right)$$

$$= \sqrt{\frac{2\pi}{\theta_a^W(t)}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x f_M^{(\mathrm{even})}(t, \boldsymbol{v}) \, \mathrm{d}v_x \, \mathrm{d}v_y \, \mathrm{d}v_z,$$

which means that the right-hand side of (31) is completely independent of $f_M^{(\mathrm{odd})}$.

By the above results, in general, (31) can be written as

$$\tilde{f}_{lmn}(t) = \frac{2\chi_a}{2 - \chi_a} \sum_{l'=0}^{M} \sum_{\substack{m'=-l' \\ l'+m' \text{ even}}}^{l'} \sum_{n'=0}^{\lfloor (M-l')/2 \rfloor} B_{lmn}^{l'm'n'} \tilde{f}_{l'm'n'}(t), \qquad l + m \text{ is odd.}$$

The constants $B_{lmn}^{l'm'n'}$ are given by

$$B_{lmn}^{l'm'n'} = \mathfrak{m}\tilde{\mathcal{M}}_{lmn}^W \sqrt{\frac{2\pi}{\theta_a^W(t)}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{0} v_x p_{l'm'n'}^{[\overline{\boldsymbol{u}},\overline{\theta}]} \omega^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\, \mathrm{d}v_x\, \mathrm{d}v_y\, \mathrm{d}v_z$$

$$- \mathfrak{m}\overline{\theta}^{l+2n} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^\dagger p_{l'm'n'}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \omega^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\, \mathrm{d}v_x\, \mathrm{d}v_y\, \mathrm{d}v_z,$$

where

$$\tilde{\mathcal{M}}_{lmn}^W = \frac{\overline{\theta}^{l+2n}}{[2\pi\theta_a^W(t)]^{3/2}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^\dagger \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}_a^W(t)|^2}{2\theta_a^W(t)} \right)\, \mathrm{d}v_x\, \mathrm{d}v_y\, \mathrm{d}v_z.$$

4. Numerical examples. By now, we are ready to carry out numerical tests to see the performance of the method. In all our numerical tests, we choose $\eta = 10$, $\mathfrak{m} = 6.63 \times 10^{-26}\,\mathrm{kg}$, and $\kappa = 3.469\,46 \times 10^{-113}\,\mathrm{kg\,m^{11}\,s^{-2}}$. To define the Knudsen number, we employ the variable hard sphere (VHS) model [7]. At the reference temperature 273.15 K, if the reference diameter of the gas molecule is $d = 4.17 \times 10^{-10}\,\mathrm{m}$, then the viscosity of the VHS gas matches the viscosity of the inverse power law gas. Thus the mean free path and the Knudsen number can be defined by

$$(33) \qquad \lambda = \frac{\mathfrak{m}}{\sqrt{2}\pi\rho d^2}, \qquad Kn = \lambda/L,$$

where ρ is the reference gas density and L is the characteristic length. For spatial discretization, we use the linearly reconstructed finite volume method, the nodal discontinuous Galerkin method, or the finite volume WENO scheme, depending on the specific simulations below. In more detail, for convenient processing of the wall boundary condition, the linearly reconstructed finite volume method or the nodal discontinuous Galerkin method will be used if such a boundary condition is applied, while the WENO scheme will be used for other cases. Both steady-state and unsteady flows are to be carried out in our numerical tests. Note that although only $(1+3)$-dimensional flows are simulated, all the examples below are quite challenging due to the existence of a high Mach number or large temperature ratio, which makes it difficult to capture the profile of the distribution function over the whole computational domain.

4.1. Simulation of steady-state flows. To study the steady-state flows, we start from a given initial state and use time-stepping to evolve the system for a sufficiently long time. To describe the stopping criterion, we define

$$\|f(t_1,\cdot,\cdot) - f(t_2,\cdot,\cdot)\| := \left(\int_a^b \int_{\mathbb{R}^3} |f(t_1,x,\boldsymbol{v}) - f(t_2,x,\boldsymbol{v})|^2 \left[\omega^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^{-1} \mathrm{d}\boldsymbol{v}\, \mathrm{d}x \right)^{1/2}.$$

We consider that the steady state is achieved at the nth time step if the numerical solution satisfies

$$(34) \qquad \frac{\|f((n+1)\Delta t,\cdot,\cdot) - f(n\Delta t,\cdot,\cdot)\|}{\|f(\Delta t,\cdot,\cdot) - f(0,\cdot,\cdot)\|} < \epsilon,$$

where Δt is the time step.

4.1.1. Steady shock structure. The plane wave shock structure is a classical example frequently used to test the kinetic models or the Boltzmann solver [37, 33]. The domain is unbounded ($a = -\infty$ and $b = +\infty$) and the boundary conditions are given by

$$\lim_{x \to -\infty} f(x, \boldsymbol{v}) = f_a(\boldsymbol{v}) := \frac{\rho_a}{\mathfrak{m}(2\pi\theta_a)^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}_a|^2}{2\theta_a}\right),$$

$$\lim_{x \to +\infty} f(x, \boldsymbol{v}) = f_b(\boldsymbol{v}) := \frac{\rho_b}{\mathfrak{m}(2\pi\theta_b)^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}_b|^2}{2\theta_b}\right),$$

where

$$\frac{\rho_b}{\rho_a} = \frac{4Ma^2}{Ma^2 + 3}, \quad \boldsymbol{u}_a = \left(\sqrt{\frac{5\theta_a}{3}}Ma, 0, 0\right)^T,$$

$$\boldsymbol{u}_b = \left(\sqrt{\frac{5\theta_a}{3}}\frac{Ma^2 + 3}{4Ma}, 0, 0\right)^T, \quad \frac{\theta_b}{\theta_a} = \frac{(5Ma^2 - 1)(Ma^2 + 3)}{16Ma^2},$$

and Ma is the Mach number of the shock wave. In our numerical tests, we set

$$\rho_a = 9.282 \times 10^{-6}\,\mathrm{kg\,m^{-3}}, \qquad \theta_a = k_B/\mathfrak{m} \times 273.15\,\mathrm{K} = 5.688 \times 10^4\,\mathrm{m^2\,s^{-2}}.$$

To obtain the steady state, we set the initial condition to be

$$f(0, x, \boldsymbol{v}) = \begin{cases} f_a(\boldsymbol{v}) & \text{if } x < 0, \\ f_b(\boldsymbol{v}) & \text{if } x > 0, \end{cases}$$

and we evolve the distribution functions until (34) is achieved with $\epsilon = 10^{-6}$. Numerically, the computational domain is set to be $[-30\lambda, 30\lambda]$, where λ is the mean free path defined in (33) with $\rho = \rho_a$. The computational domain is divided into 60 grid cells, and the fifth-order WENO finite volume method is applied for the spatial discretization. For the velocity discretization, we use

$$\overline{\boldsymbol{u}} = (\boldsymbol{u}_a + \boldsymbol{u}_b)/2, \qquad \overline{\theta} = 0.7\theta_b.$$

Two Mach numbers $Ma = 6.5$ and $Ma = 9.0$ are considered in our numerical tests. The corresponding solution of density ρ, temperature T, normal stress σ_{xx}, and heat flux q_x, obtained by our method with $M_0 = 10$ and $M = 20$, are presented in Figure 1 for $Ma = 6.5$ and in Figure 2 for $Ma = 9.0$, respectively. The results are compared with the solution obtained by the DSMC method [7]. Even for such high Mach number cases, the results show that the shock structure profiles, including the shock thickness and the peak values of heat flux and normal stress, coincide perfectly well between the solution given by our method and the DSMC method.

Both simulations are performed on a cluster with the CPU model Intel Xeon E5-2697A V4 @ 2.6GHz. Ten threads are used for each simulation. Details of the simulations, including the number of time steps and the total elapsed time, are listed in Table 1, which shows the efficiency of the proposed method. Moreover, it is noted that for the proposed method, most of the computational cost is devoted to the computation of the approximate collision operator. And in [10], a comparison with the Fourier spectral method has been carried out for the spatially homogeneous problems, which shows the computation of the collision operator used in the present work is competitive with the Fourier spectral method. Thus, the present method would still be very efficient in comparison to the Fourier spectral method for the spatially inhomogeneous problems.
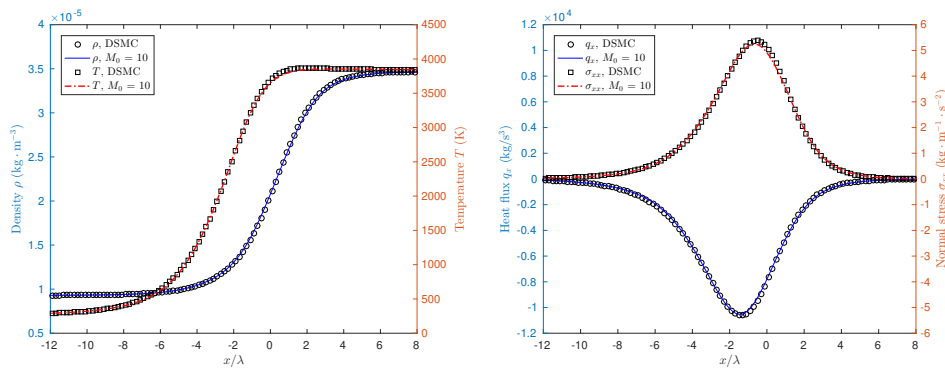
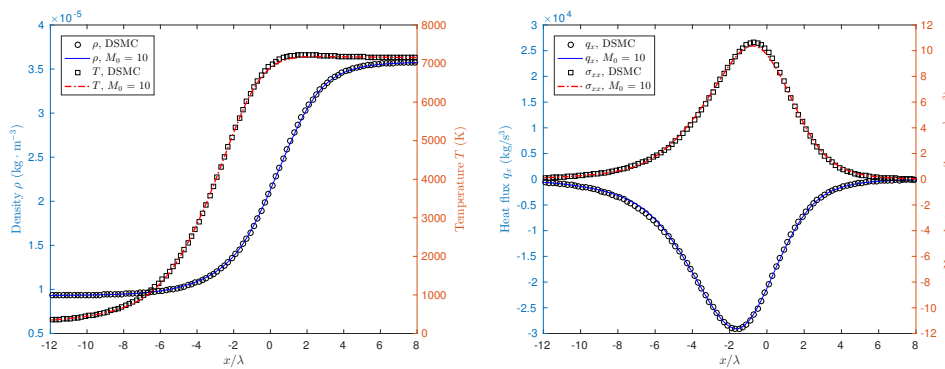FIG. 1. *Solution of the shock structure for* $Ma = 6.5$ *and* $M = 20$.

FIG. 2. *Solution of the shock structure for* $Ma = 9.0$ *and* $M = 20$.

TABLE 1
Run-time data for shock structure simulations with 60 *grid cells,* $M_0 = 10$, *and* $M = 20$.

Test case	$Ma = 6.5$	$Ma = 9.0$
Number of coefficients	1771	1771
Time step (Δt)	6.47×10^{-7} s	4.74×10^{-7} s
Number of time steps	2605	2690
Total elapsed time	96.31 s	99.99 s
Elapsed time per time step	3.70×10^{-2} s	3.72×10^{-2} s

4.1.2. Fourier flow. This is another benchmark test for problems with boundary conditions [25]. The fluid locates between two stationary and infinitely large parallel plates with different temperature. At the steady state, a significant temperature jump can be observed for rarefied gases. The parameters of this problem include

- L: distance between two plates;
- T_a^W, T_b^W: the temperature of the left and right plates;
- ρ_0: the average density of the fluid.

In our tests, we always choose $T_a^W = 273.15$ K and $\rho_0 = 9.282 \times 10^{-6}$ kg m^{-3}. The computational domain is defined by $a = -L/2$ and $b = L/2$ with the accommodation coefficients in the boundary condition being $\chi_a = \chi_b = 1$. The domain is decomposed into 10 uniform grid cells, and the fourth-order nodal discontinuous Galerkin method [24] is used for spatial discretization. For velocity space, we discretize it using $\overline{\boldsymbol{u}} = 0$ and $\overline{\theta} = (\theta_a^W + \theta_b^W)/2$, where $\theta_a^W = k_B T_a^W/\mathfrak{m}$ and $\theta_b^W = k_B T_b^W/\mathfrak{m}$. We compute the

steady state by starting from the initial condition

$$f(0, x, \boldsymbol{v}) = \frac{\rho_0}{\mathfrak{m}(2\pi\overline{\theta})^{3/2}} \exp\left(-\frac{|\boldsymbol{v}|^2}{2\overline{\theta}}\right),$$

and the stopping criterion is again (34) with $\epsilon = 10^{-6}$. In order to compare our results, the DSMC method [7] is also employed to produce the reference solution. Below we are going to consider two different choices of T_b^W.

(1) $T_b^W = 4T_a^W$. We first set the temperature ratio of two plates to be 4. Three distances $L = 0.092\,456$ m, $0.018\,491$ m, and $0.003\,698$ m are considered. They correspond to Knudsen number 0.1, 0.5, and 2.5, respectively. For numerical results presented in this paper, we adopt $M = 20$, 30, and 40, as the Knudsen number increases from 0.1 to 2.5.

Numerical solutions of density ρ, temperature T, normal stress σ_{xx}, and heat flux q_x, obtained by our method with three choices of M_0, i.e., $M_0 = 5$, 10, and 15, are presented in Figures 3–5 for $Kn = 0.1$, 0.5, and 2.5, respectively. It can be seen that all our results agree well with the DSMC solution. In particular, most of our solutions become much closer and closer to the reference DSMC solution, as M_0 increases. And the relative deviations of our solution away from the DSMC solution are actually quite small, i.e., for the heat flux q_x, these deviations are all less than 1.4% for $M_0 = 10$ and 0.6% for $M_0 = 15$.

(2) $T_b^W = 10T_a^W$. Now we set the temperature ratio to be 10, which is obviously tougher to simulate due to the wide spread of the distribution functions. Two distances $L = 0.018\,491$ m and $0.003\,698$ m, with the corresponding Knudsen number 0.5 and 2.5, respectively, are considered. As the previous case, we set $M = 30$ for $Kn = 0.5$ and $M = 40$ for $Kn = 2.5$ in our tests.

Numerical solutions obtained by our method with $M_0 = 5$, 10, and 15 and the DSMC method are given in Figures 6 and 7 for $Kn = 0.5$ and $Kn = 2.5$, respectively. The results still show a good agreement between our solutions and the DSMC solutions, although more obvious deviation can be observed, especially for the case with $Kn = 2.5$. However, as more moments are modeled accurately by the quadratic collision model in our method, which indicates M_0 is increased, remarkable improvement of our results can be obtained as shown in these figures.

At last, the run-time data of partial simulations, which are obtained on the same cluster with 10 threads for each simulation as shock structure simulations, are also provided in Table 2 to show the efficiency of our method.

4.2. Simulation of unsteady flows. Now we use another two numerical examples to study the evolution of the flow. In both cases, we need to employ the steady shock structure simulated in section 4.1.1 in the initial condition, while in this section, we expect that the shock wave moves at the given Mach number. This can be achieved by the following steps:

- Perform the same simulation as in section 4.1.1 using $\overline{\boldsymbol{u}} = (u_a, 0, 0)^T$.
- In the numerical results, we perform a translation of the distribution function such that the fluid state in front of the shock has velocity zero. Such a translation can be simply implemented by setting $\overline{\boldsymbol{u}} = 0$.

The second step is in fact a transform of the frame of reference, which turns a steady shock structure to a moving shock structure. Below we are going to study the collision of two shock structures and the interaction between the shock and the solid wall.

4.2.1. Collision of two shocks. We first consider the collision of two shock waves which move in opposite directions with the Mach numbers $Ma = 3.8$ and 6.5,

(a) Density, ρ (kg $\cdot$ m^{-3})

(b) Temperature, T (K)

(c) Normal stress, σ_{xx} (kg $\cdot$ m^{-1} $\cdot$ s^{-2})
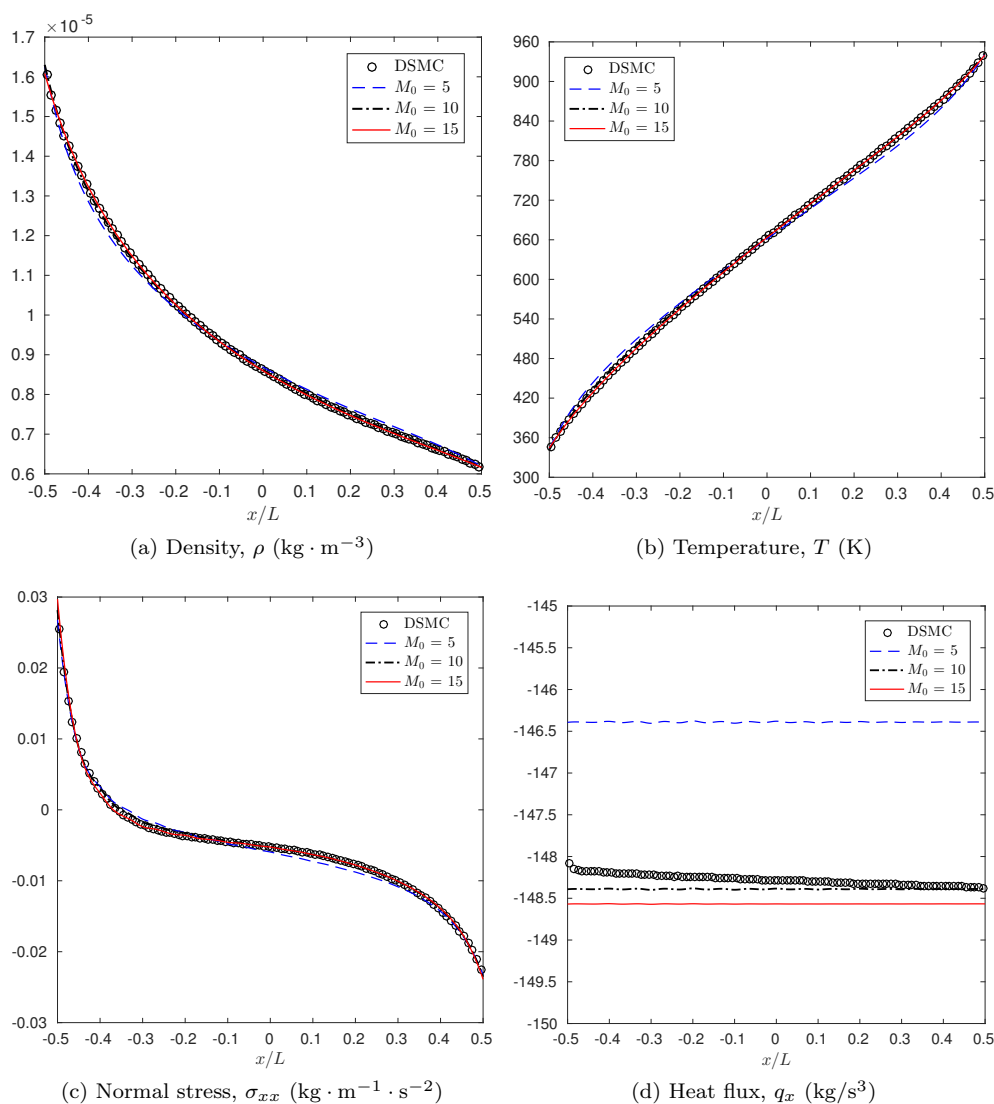
(d) Heat flux, q_x (kg/s^3)

FIG. 3. *Solution of the Fourier flow for $T_b^W = 4T_a^W$ with $Kn = 0.1$ and $M = 20$.*

respectively. Precisely speaking, the shock with $Ma = 3.8$ is on the left of the domain and moves to the right, while the shock with $Ma = 6.5$ is on the right of the domain and moves to the left. In our method, the parameters $\overline{\boldsymbol{u}}$ and $\overline{\theta}$ must be constants for all grid cells. Thus, the initial shock profiles are obtained by the simulations in section 4.1.1 using $\overline{\boldsymbol{u}} = \boldsymbol{u}_a$ with the corresponding Mach numbers and $\overline{\theta} = 0.8\theta_b$, where θ_b is computed with $Ma = 6.5$. After resetting $\overline{\boldsymbol{u}} = 0$ and reversing the velocity of the shock with $Ma = 3.8$, we then obtain the initial state of this test. Due to the existence of two shocks in opposite directions, the distribution function can spread widely after the merging of two shocks, which makes the problem highly challenging.

Our numerical solutions of density ρ, temperature T, normal stress σ_{xx}, and heat flux q_x with $M_0 = 10$ and $M = 20$ are presented in Figure 8 for various time instants from $t = 0$ to 1.2. As shown in the figure, the black solid line gives the corresponding
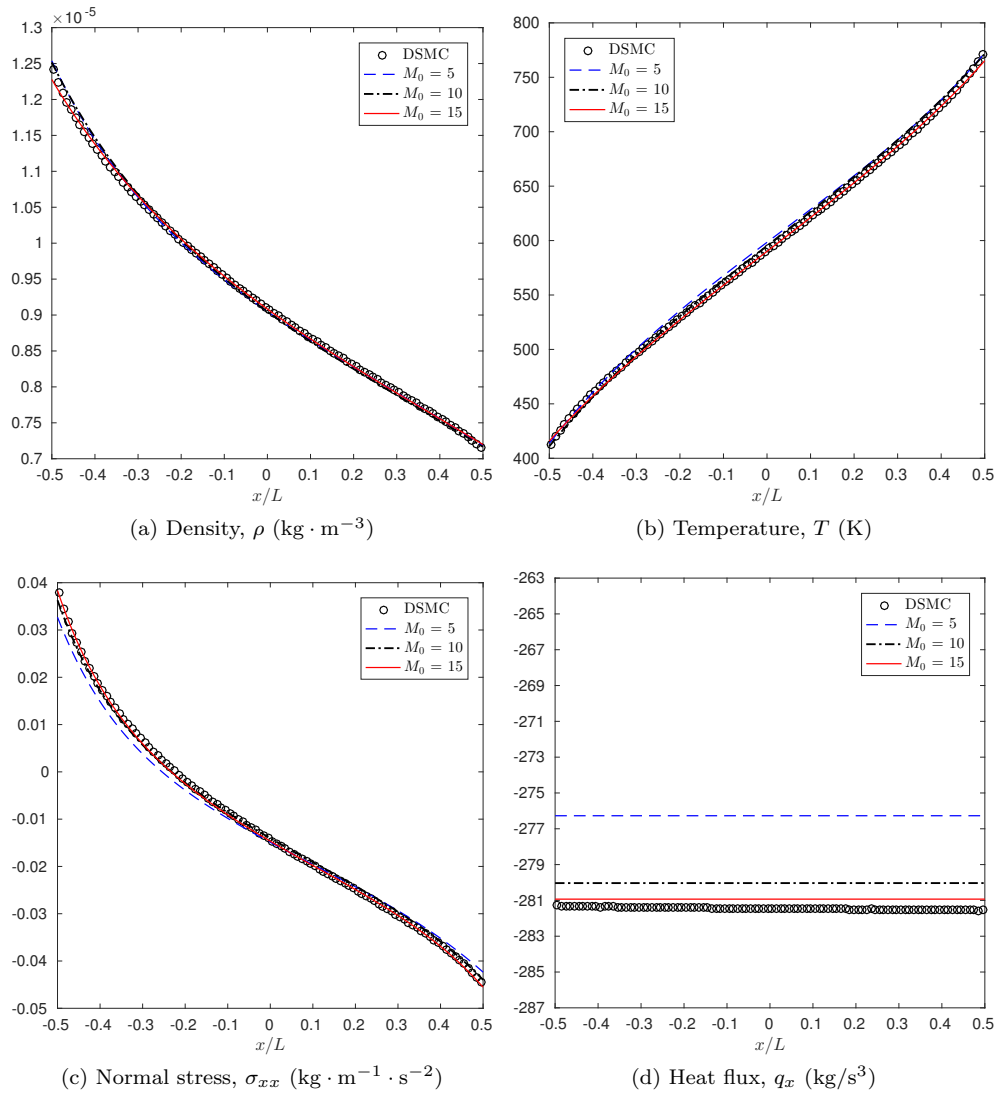
(a) Density, ρ (kg $\cdot$ m^{-3})

(b) Temperature, T (K)

(c) Normal stress, σ_{xx} (kg $\cdot$ m^{-1} $\cdot$ s^{-2})

(d) Heat flux, q_x (kg/s^3)

FIG. 4. *Solution of the Fourier flow for $T_b^W = 4T_a^W$ with $Kn = 0.5$ and $M = 30$.*

initial state of these quantities. After the collision of two shocks around the time $t = 0.6$, two new shock waves will be generated with a rarefaction wave standing between them. The left shock wave moves to the left from the right with the speed smaller than $Ma = 6.5$, while the right shock wave moves from left to right with the speed smaller than $Ma = 3.8$. Besides, the left shock wave moves much faster than the right shock wave. These solution structures are consistent with the corresponding solutions obtained by the classical Euler equations, whose numerical results at $t = 1.2$ are also given in Figure 8 by gray solid line.

4.2.2. Interaction of a shock and a solid wall. This test investigates the interaction of a shock wave and a solid wall. The shock wave moves from right to left, and the solid wall is fixed on the left of the computational domain $[-30\lambda, 30\lambda]$ with
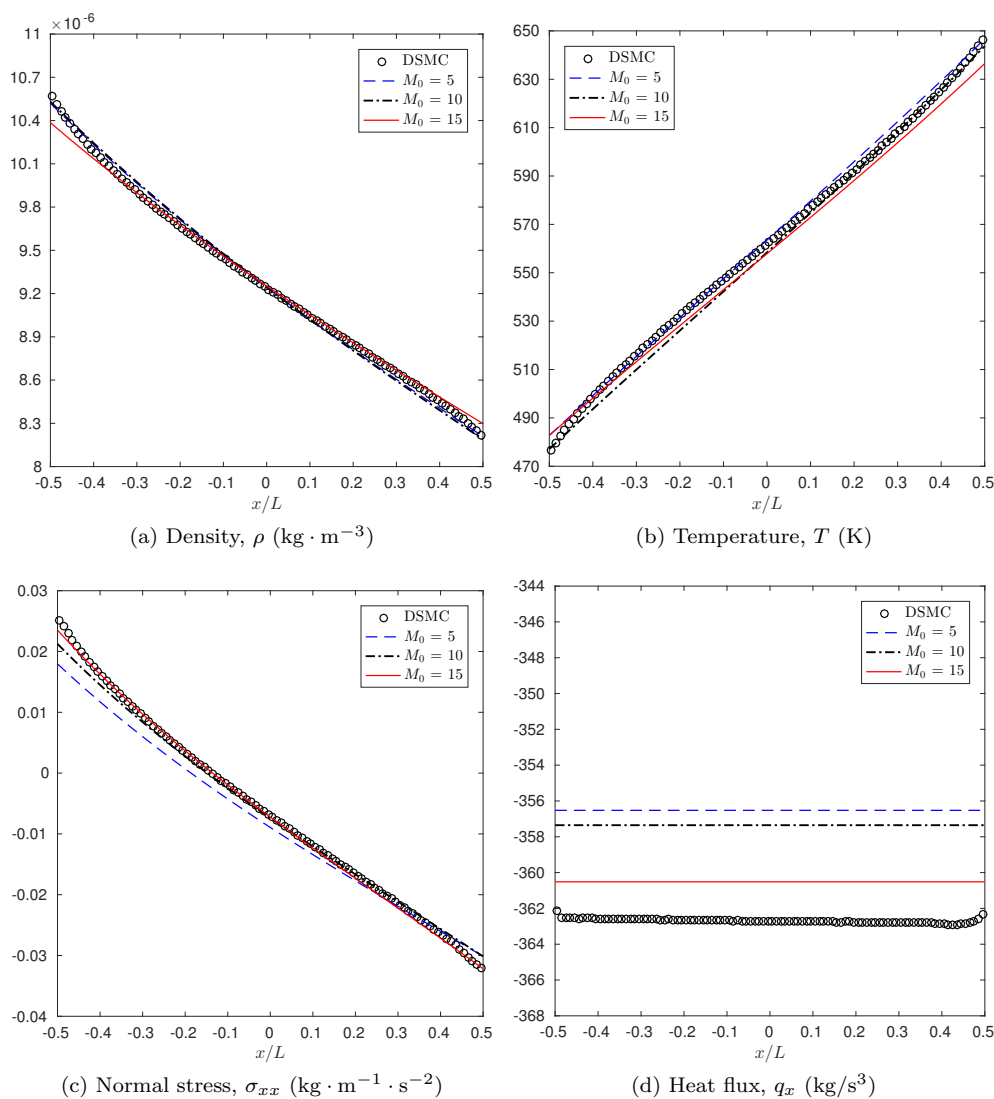
(a) Density, ρ (kg $\cdot$ m^{-3})

(b) Temperature, T (K)

(c) Normal stress, σ_{xx} (kg $\cdot$ m^{-1} $\cdot$ s^{-2})

(d) Heat flux, q_x (kg/s^3)

Fig. 5. *Solution of the Fourier flow for $T_b^W = 4T_a^W$ with $Kn = 2.5$ and $M = 40$.*

the wall temperature $T_a^W = 273.15$ K and the accommodation coefficient $\chi_a = 1$. It is observed that the temperature would be doubled after interacting with the wall. So we use $\overline{\boldsymbol{u}} = \boldsymbol{u}_a$ and $\overline{\theta} = \theta_b$ together with $M_0 = 10$ and $M = 40$ to prepare the initial shock profile. Then the shock wave will move to the left with the expected speed by setting $\overline{\boldsymbol{u}} = 0$. For easier processing of the boundary condition, a linearly reconstructed finite volume method with 600 uniform grid cells is performed, instead of the fifth-order WENO finite volume method used in section 4.1.1.

This example is even more challenging due to the existence of the solid wall, which introduces discontinuity into the distribution function. Numerical solutions of density ρ, temperature T, normal stress σ_{xx}, and heat flux q_x are presented in Figure 9 at various time instants from $t = 0$ to 1.2 with $Ma = 3.8$. The black solid line in the figure represents the corresponding initial state of these quantities. It is shown that
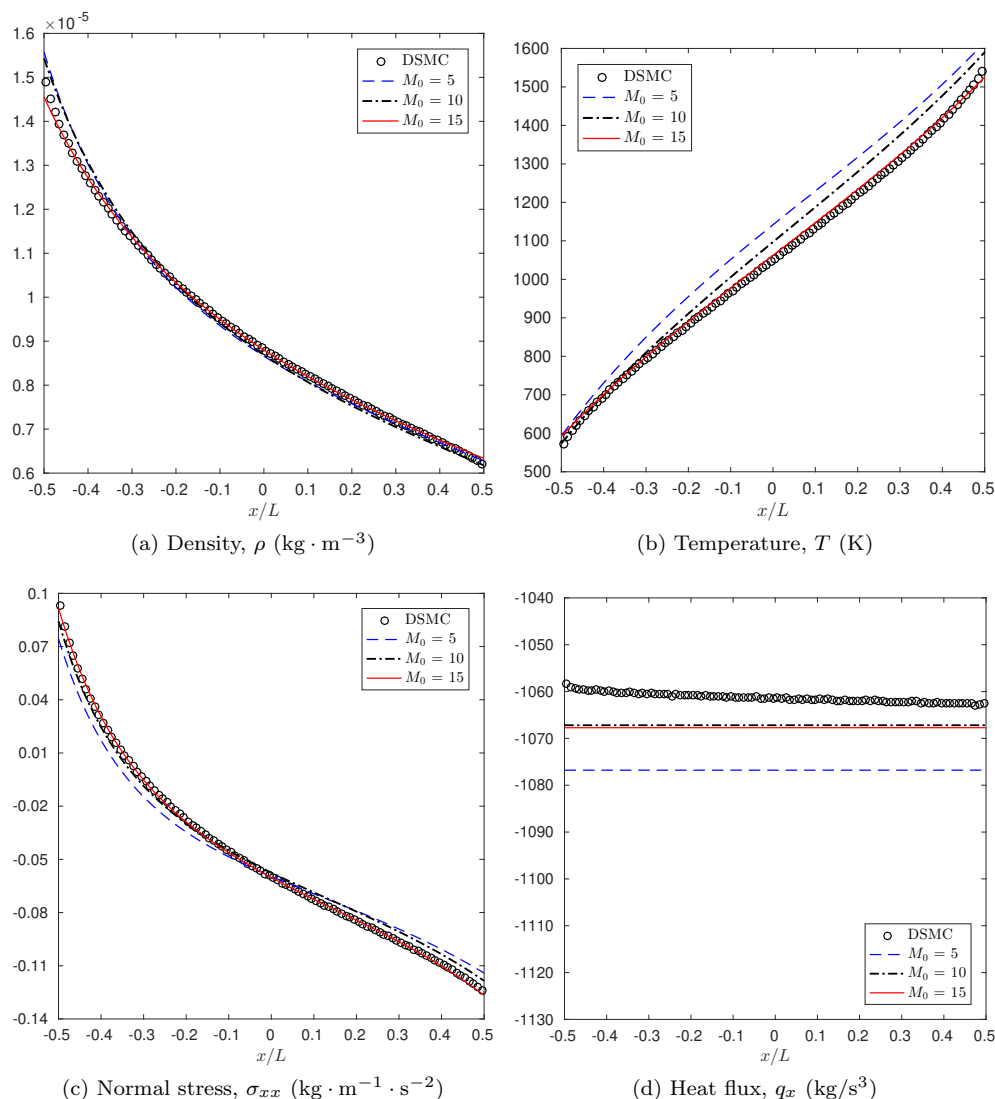
(a) Density, ρ (kg·m^{-3})   (b) Temperature, T (K)

(c) Normal stress, σ_{xx} (kg·m^{-1}·s^{-2})   (d) Heat flux, q_x (kg/s^3)

FIG. 6. *Solution of the Fourier flow for $T_b^W = 10 T_a^W$ with $Kn = 0.5$ and $M = 30$.*

the shock collides with the wall around the time $t = 0.6$. By the interaction with the wall, the particles accumulate in a small region around the wall, resulting in a significant increase in density. A shock wave is then bounced back, with the speed much slower than the original shock. Similar results can be obtained for the case $Ma = 6.5$, as given in Figure 10 at various time instants from $t = 0$ to 0.8.

5. Conclusion. We have developed an efficient numerical scheme for the spatially inhomogeneous Boltzmann equation based on the Burnett spectral method. Two major numerical strategies are used: (1) we coupled the quadratic collision operator and the BGK collision operator to balance the computational cost and the modeling accuracy; (2) we introduced a zero term in the collision term to preserve the steady state. Both steady and unsteady flows are solved as numerical examples.
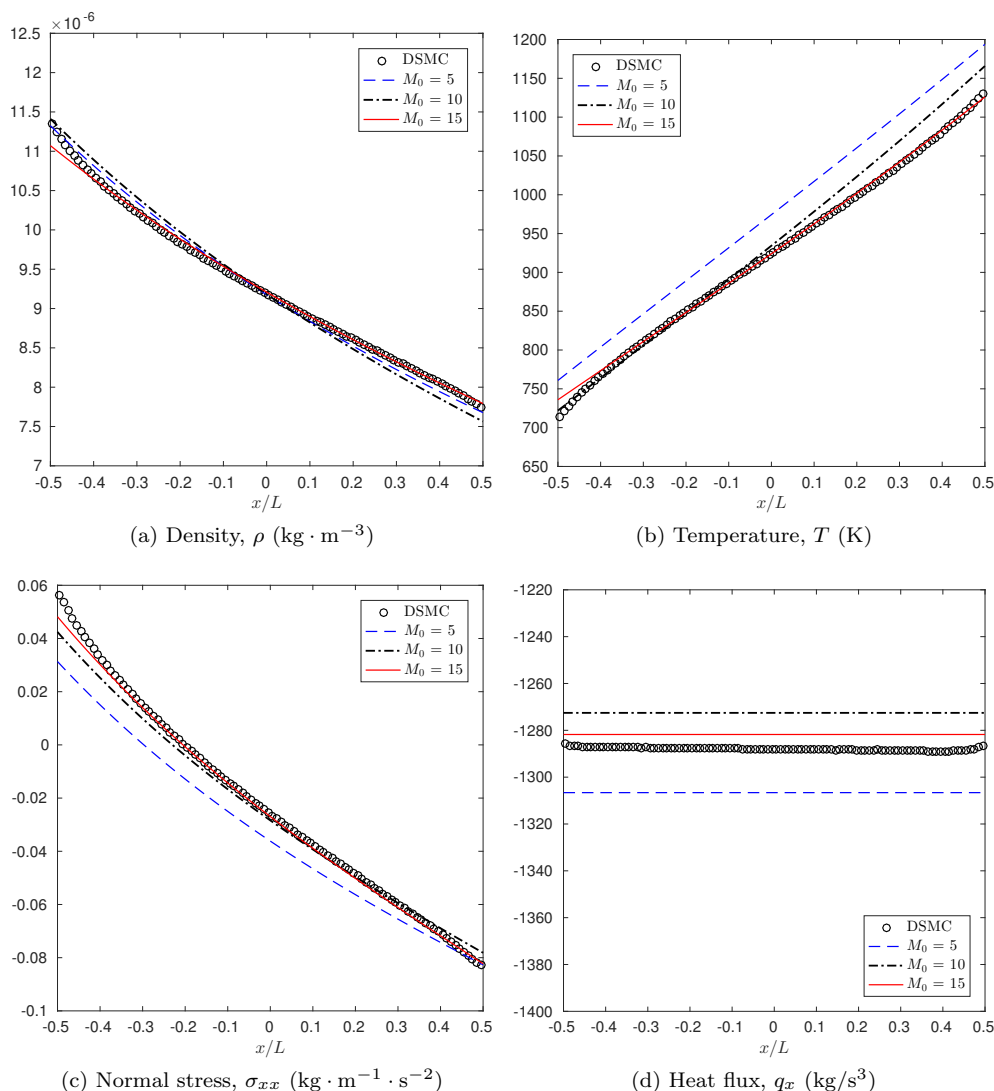
(a) Density, ρ (kg $\cdot$ m^{-3})

(b) Temperature, T (K)

(c) Normal stress, σ_{xx} (kg $\cdot$ m^{-1} $\cdot$ s^{-2})

(d) Heat flux, q_x (kg/s^3)

FIG. 7. *Solution of the Fourier flow for $T_b^W = 10T_a^W$ with $Kn = 2.5$ and $M = 40$.*

TABLE 2
Run-time data for Fourier simulations with 10 *grid cells.*

Test case	$T_b^W = 4T_a^W$			$T_b^W = 10T_a^W$	
Knudsen number (Kn)	0.1	0.5	2.5	0.5	2.5
M_0 and M	10, 20	10, 30	10, 40	15, 30	15, 40
Number of coefficients	1771	5456	12341	5456	12341
Time step (Δt)	1.73×10^{-7} s	2.74×10^{-8} s	4.67×10^{-9} s	1.85×10^{-8} s	3.15×10^{-9} s
Number of time steps	8754	13230	31899	15403	35888
Total elapsed time	164.68 s	251.13 s	1360.74 s	4367.15 s	13 266 s
Elapsed time per time step	1.88×10^{-2} s	1.90×10^{-2} s	4.27×10^{-2} s	0.284 s	0.370 s

(a) Density, ρ (kg · m^{-3})   (b) Temperature, T (K)

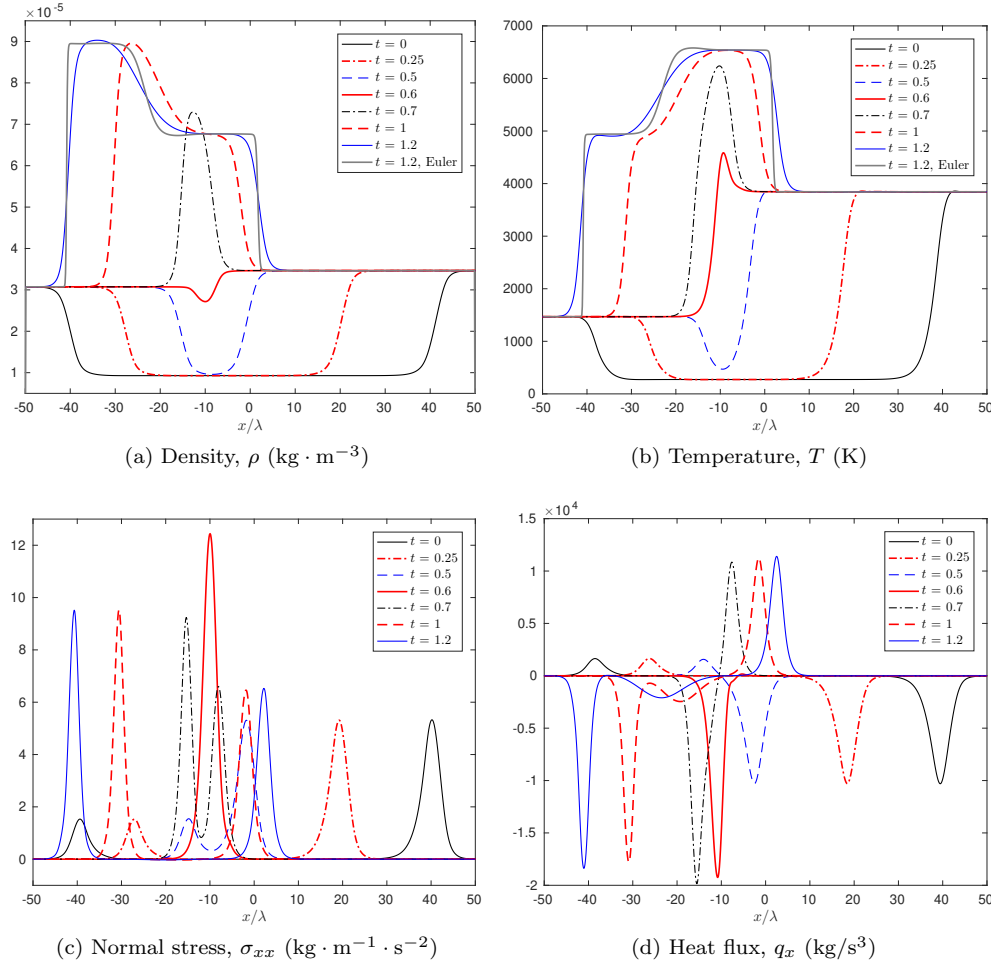(c) Normal stress, σ_{xx} (kg · m^{-1} · s^{-2})   (d) Heat flux, q_x (kg/s^3)

FIG. 8. *Solution for the collision of two shocks with Ma = 3.8 and 6.5, respectively, at various times with $M_0 = 10$ and $M = 20$. The gray solid line is the corresponding results from the Euler equations at $t = 1.2$.*

Although only one spatial dimension (with three velocity dimensions) is considered in the simulation, all the examples include large temperature contrast or fast flow speed, which are still quite challenging. The computational time shows the high numerical efficiency of our method.

Appendix A. Inverse power law model. In this section we provide a full description of the collision kernel of the inverse power law model. For given κ, $\mathfrak{m}$, and η, we have

$$B(\boldsymbol{v} - \boldsymbol{v}_1, \boldsymbol{n}) = \left(\frac{2\kappa}{\mathfrak{m}}\right)^{\frac{2}{\eta-1}} |\boldsymbol{v} - \boldsymbol{v}_1|^{\frac{\eta-5}{\eta-1}} \frac{W_0(\vartheta)W_0'(\vartheta)}{\sin \vartheta},$$

where

$$\vartheta = \arccos \frac{|(\boldsymbol{v} - \boldsymbol{v}_1) \cdot \boldsymbol{n}|}{|\boldsymbol{v} - \boldsymbol{v}_1|},$$

(a) Density, ρ (kg $\cdot$ m^{-3})

(b) Temperature, T (K)

(c) Normal stress, σ_{xx} (kg $\cdot$ m^{-1} $\cdot$ s^{-2})
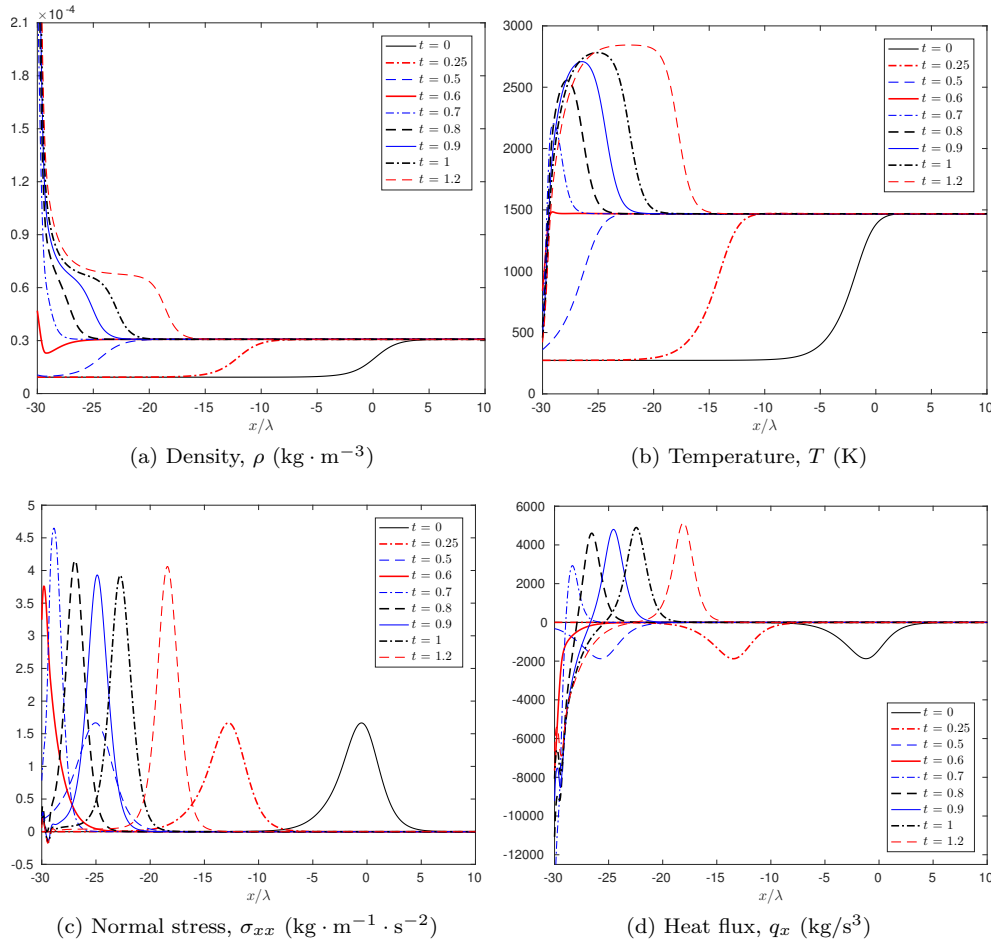
(d) Heat flux, q_x (kg/s^3)

FIG. 9. *Solution for the interaction of a shock with Ma = 3.8 and a solid wall at various times with $M_0 = 10$ and $M = 40$.*

and the function $W_0(\cdot)$ is defined by

$$W_0(\vartheta) = \sqrt{1 - y(\vartheta)} \left(\frac{\eta - 1}{2} y(\vartheta) \right)^{-\frac{1}{\eta-1}}$$

with $y(\cdot)$ being defined implicitly as

$$(35) \qquad \int_0^1 \left(1 - x^2[1 - y(\vartheta)] - x^{\eta-1} y(\vartheta) \right)^{-1/2} \sqrt{1 - y(\vartheta)} \, \mathrm{d}x = \vartheta.$$

The viscosity coefficient $\overline{\mu}$ can be derived from the first-order Chapman–Enskog expansion. In (9), the function $A_2(\eta)$ is defined by

$$(36) \qquad A_2(\eta) = \int_0^{\pi/2} W_0(\vartheta) W_0'(\vartheta) \sin^2(2\vartheta) \, \mathrm{d}\vartheta.$$

When $\eta = 5$, the corresponding inverse power law model is also called the Maxwell model. In this case, for any l, m, n, the function $p_{lmn}(\boldsymbol{v}^*)\omega(\boldsymbol{v}^*)$ is the eigenfunction

(a) Density, ρ (kg $\cdot$ m^{-3})

(b) Temperature, T (K)

(c) Normal stress, σ_{xx} (kg $\cdot$ m^{-1} $\cdot$ s^{-2})
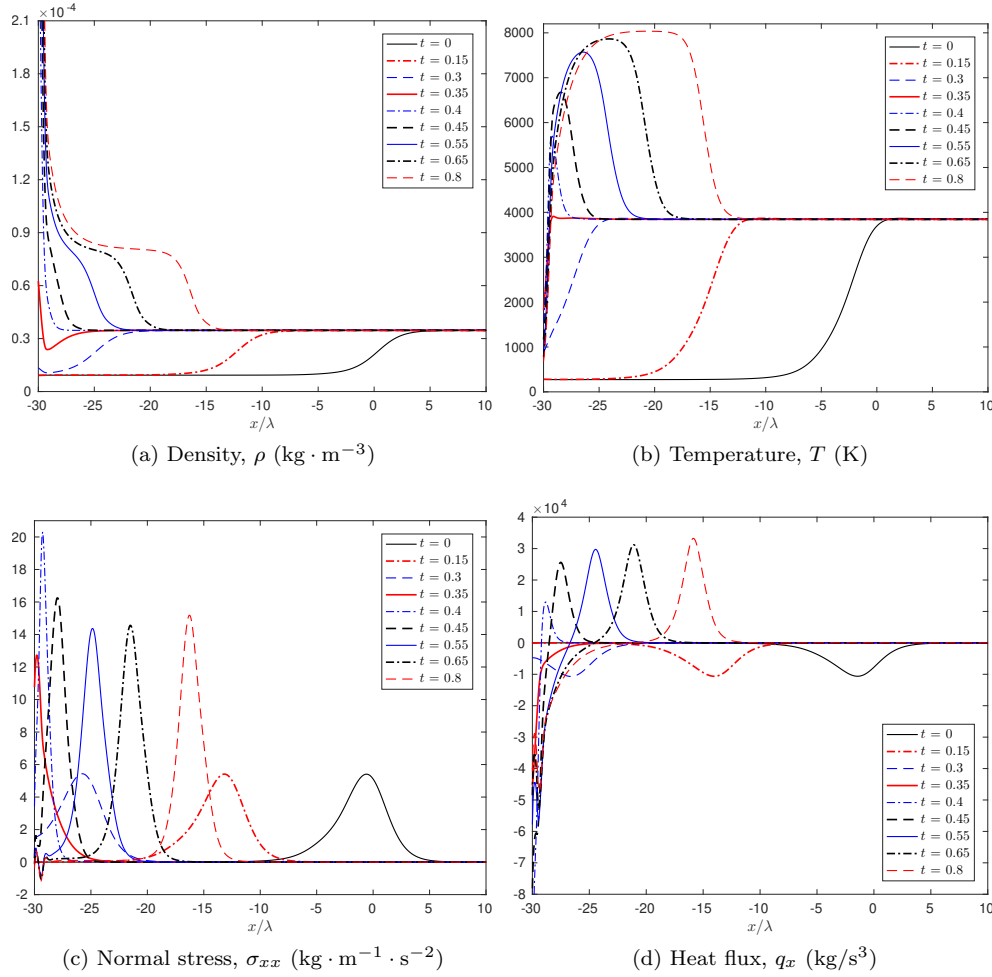
(d) Heat flux, q_x (kg/s^3)

FIG. 10. *Solution for the interaction of a shock with Ma = 6.5 and a solid wall at various time with $M_0 = 10$ and $M = 40$.*

of the linearized collision operator $L[f] = Q(f, \omega) + Q(\omega, f)$. Here we complete the definition of p_{lmn} by providing the definition of the Laguerre polynomials and spherical harmonics, which appear in (11):

$$(37) \qquad L_n^{(\alpha)}(x) = \frac{x^{-\alpha} \exp(x)}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}[x^{n+\alpha} \exp(-x)],$$

$$Y_l^m(\boldsymbol{n}) = Y_l^m(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\vartheta) \exp(\mathrm{i}m\varphi),$$

where (ϑ, φ) is the spherical coordinates of $\boldsymbol{n}$, i.e., $\boldsymbol{n} = (\cos\vartheta, \sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi)^T$, and $P_l^m(\cdot)$ is the associated Legendre function defined by

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}[(x^2-1)^l].$$

When $m = 0$, the polynomial $P_l^0(x)$ is the Legendre polynomial of degree l, which is often denoted by $P_l(x)$.

Appendix B. Coefficients in the collision operator for inverse power law models. To complete the description of the ODE system (13), we summarize in this section the results in [35, 10] and provide the complete process to compute these coefficients. The computational cost for computing all these coefficients is $O(M^{14})$ (see [10] for details), but this needs to be done only once before the numerical simulation.

Supposing the index η in the power potential is given, below we are going to provide a sequence of formulas, by which the coefficients $A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2}$ can be computed step by step:

- Compute the following integral $\mathcal{I}_k$ for $k \leqslant 2M$:

$$
\mathcal{I}_k = \int_0^1 \left[P_k\left(-\cos 2\vartheta(y) \right) - 1 \right] \left[2(1-y) + (\eta-1)y \right] \left[(\eta-1)y \right]^{-\frac{\eta+1}{\eta-1}} \, \mathrm{d}y,
$$

where $\vartheta(y)$ is the inverse function of $y(\vartheta)$ defined in (35), and P_k is the Legendre polynomial of degree k.

- Compute K_{mn}^{kl} for $k \leqslant 2M$, $l \leqslant M$, $m \leqslant \lfloor k/2 \rfloor$, $n \leqslant \lfloor l/2 \rfloor$, $k - 2m = l - 2n$ by

$$
K_{mn}^{kl} = (-1)^{m+n} 2^{\frac{\eta-3}{\eta-1}+k-2m} \Gamma\left(k - 2m + 2 - \frac{2}{\eta-1} \right) C(\eta) \mathcal{I}_{k-2m}
$$

$$
\times \sum_{i=0}^{\min(m,n)} \binom{\frac{1}{2} - \frac{2}{\eta-1}}{m-i} \binom{\frac{1}{2} - \frac{2}{\eta-1}}{n-i} \binom{k - 2m + i + 1 - \frac{2}{\eta-1}}{i},
$$

where

$$
C(\eta) = \frac{5}{4^{\frac{3}{2} - \frac{2}{\eta-1}} \sqrt{\pi} A_2(\eta) \Gamma(4 - 2(\eta-1))}.
$$

- For all $k \leqslant M$, define the homogeneous polynomial $S_k(\boldsymbol{v}, \boldsymbol{w})$ for $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ by the following recursive formulas:

$$
S_0(\boldsymbol{v}, \boldsymbol{w}) = 1, \qquad S_1(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w},
$$

$$
S_{k+1}(\boldsymbol{v}, \boldsymbol{w}) = \frac{2k+1}{k+1} (\boldsymbol{v} \cdot \boldsymbol{w}) S_k(\boldsymbol{v}, \boldsymbol{w}) - \frac{k}{k+1} (|\boldsymbol{v}||\boldsymbol{w}|)^2 S_{k-1}(\boldsymbol{v}, \boldsymbol{w}), \quad k \geqslant 2.
$$

By the above definition, write $S_k(\cdot, \cdot)$ as

$$
S_k(\boldsymbol{v}, \boldsymbol{w}) = \sum_{k_1+k_2+k_3=k} \sum_{l_1+l_2+l_3=k} S_{k_1 k_2 k_3}^{l_1 l_2 l_3} v_1^{k_1} v_2^{k_2} v_3^{k_3} w_1^{l_1} w_2^{l_2} w_3^{l_3},
$$

and find all the coefficients $S_{k_1 k_2 k_3}^{l_1 l_2 l_3}$.

- Compute $C_{m_1 m_2 m_3}^{k_1 k_2 k_3}$ for $k_1 + k_2 + k_3 \leqslant 2M$ and $m_1 + m_2 + m_3 \leqslant M$ by

$$C_{m_1 m_2 m_3}^{k_1 k_2 k_3} = \frac{(-1)^{m_1+m_2+m_3} 4\pi (m_1 + m_2 + m_3)!}{(2(k_1 + k_2 + k_3 - m_1 - m_2 - m_3) + 1)!} \frac{k_1! k_2! k_3!}{m_1! m_2! m_3!}.$$

- Compute $\gamma_{k_1 k_2 k_3}^{l_1 l_2 l_3}$ for $k_1 + k_2 + k_3 \leqslant 2M$ and $l_1 + l_2 + l_3 \leqslant M$ by

$$\gamma_{k_1 k_2 k_3}^{l_1 l_2 l_3} = \sum_{m_1=0}^{\lfloor k_1/2 \rfloor} \sum_{m_2=0}^{\lfloor k_2/2 \rfloor} \sum_{m_3=0}^{\lfloor k_3/2 \rfloor} \sum_{n_1=0}^{\lfloor l_1/2 \rfloor} \sum_{n_2=0}^{\lfloor l_2/2 \rfloor} \sum_{n_3=0}^{\lfloor l_3/2 \rfloor} (2k - 4m + 1)$$
$$\times C_{m_1 m_2 m_3}^{k_1 k_2 k_3} C_{n_1 n_2 n_3}^{l_1 l_2 l_3} S_{k_1-2m_1, k_2-2m_2, k_3-2m_3}^{l_1-2n_1, l_2-2n_2, l_3-2n_3} K_{mn}^{kl},$$

where

$$k = k_1 + k_2 + k_3, \quad l = l_1 + l_2 + l_3, \quad m = m_1 + m_2 + m_3, \quad n = n_1 + n_2 + n_3.$$

- Compute $a_{i'j'}^{ij}$ for $i \leqslant M$, $j \leqslant M$, $i' \leqslant 2M$, $j' \leqslant 2M$, $i + j = i' + j'$ by

$$a_{i'j'}^{ij} = 2^{-(i'+j')/2} i! j! \sum_{s=\max(0,i'-j)}^{\min(i',i)} \frac{(-1)^{j'-i+s}}{s!(i-s)!(i'-s)!(j'-i+s)!}.$$

- Compute $B_{k_1 k_2 k_3}^{i_1 i_2 i_3, j_1 j_2 j_3}$ for $i_1 + i_2 + i_3 \leqslant M$, $j_1 + j_2 + j_3 \leqslant M$, $k_1 + k_2 + k_3 \leqslant M$ by

$$B_{k_1 k_2 k_3}^{i_1 i_2 i_3, j_1 j_2 j_3}$$
$$= \sum_{i_1'=0}^{\min(i_1+j_1, k_1)} \sum_{i_2'=0}^{\min(i_2+j_2, k_2)} \sum_{i_3'=0}^{\min(i_3+j_3, k_3)} \frac{2^{-k/2}}{2^3 \pi^{3/2}} \frac{1}{l_1'! l_2'! l_3'!} a_{i_1'j_1'}^{i_1 j_1} a_{i_2'j_2'}^{i_2 j_2} a_{i_3'j_3'}^{i_3 j_3} \gamma_{j_1'j_2'j_3'}^{l_1'l_2'l_3'},$$

where

$$j_s' = i_s + j_s - i_s', \quad l_s' = k_s - i_s', \qquad s = 1, 2, 3.$$

- Given that $C_{000}^{000} = 1$ and

$$C_{lmn}^{k_1 k_2 k_3} = 0, \qquad |m| > l \text{ or } l < 0 \text{ or } n < 0,$$

compute the coefficients $C_{lmn}^{k_1 k_2 k_3}$ for $k_1 + k_2 + k_3 \leqslant M$, $l + 2n = k_1 + k_2 + k_3$, $-l \leqslant m \leqslant l$ by solving the following equations:
(38)
$$a_{l,m+1,n}^{(-1)} C_{l+1,m,n}^{k_1 k_2 k_3} + b_{l,m+1,n}^{(-1)} C_{l-1,m,n+1}^{k_1 k_2 k_3} = \frac{1}{2} k_1 C_{l,m+1,n}^{k_1-1,k_2,k_3} - \frac{\mathrm{i}}{2} k_2 C_{l,m+1,n}^{k_1,k_2-1,k_3},$$
$$a_{l,m,n}^{(0)} C_{l+1,m,n}^{k_1 k_2 k_3} + b_{l,m,n}^{(0)} C_{l-1,m,n+1}^{k_1 k_2 k_3} = k_3 C_{l,m,n}^{k_1,k_2,k_3-1},$$
$$a_{l,m-1,n}^{(1)} C_{l+1,m,n}^{k_1 k_2 k_3} + b_{l,m-1,n}^{(1)} C_{l-1,m,n+1}^{k_1 k_2 k_3} = -\frac{1}{2} k_1 C_{l,m-1,n}^{k_1-1,k_2,k_3} - \frac{\mathrm{i}}{2} k_2 C_{l,m-1,n}^{k_1,k_2-1,k_3},$$

where

$$a_{lmn}^{(\mu)} = \frac{1}{2^{|\mu|}} \sqrt{\frac{(2(n+l)+3)[l+(2\delta_{1,\mu}-1)m+\delta_{1,\mu}+1][l+(2\delta_{1,\mu}-1)m+\delta_{1,\mu}+1]}{(2l+1)(2l+3)}},$$

$$b_{lmn}^{(\mu)} = \frac{(-1)^{\mu+1}}{2^{|\mu|}} \sqrt{\frac{2(n+1)[l-(2\delta_{1,\mu}-1)m-\delta_{1,\mu}][l-(2\delta_{1,\mu}-1)m-\delta_{1,\mu}]}{(2l-1)(2l+1)}}, \ \mu = -1, 0, 1.$$

The equations of (38) can be applied recursively. Note that (38) includes three equations. They are always consistent so that we can use (38) to solve two coefficients $C_{l+1,m,n}^{k_1 k_2 k_3}$ and $C_{l-1,m,n+1}^{k_1 k_2 k_3}$ based on the knowledge of $C_{l,m-1,n}^{k_1-1,k_2,k_3}$, $C_{l,m-1,n}^{k_1,k_2-1,k_3}$, and $C_{l,m-1,n}^{k_1,k_2,k_3-1}$.

- Compute the coefficients $A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2}$ for $l+2n \leqslant M$, $l_1+2n_1 \leqslant M$, $l_2+2n_2 \leqslant M$, $-l_1 \leqslant m_1 \leqslant l_1$, $-l_2 \leqslant m_2 \leqslant l_2$, $m = m_1 + m_2$ by

$$A_{lmn}^{l_1 m_1 n_1, l_2 m_2 n_2}$$
$$= \sum_{k_1+k_2+k_3=l+2n} \sum_{i_1+i_2+i_3=l} \sum_{j \in I_{l_2+2n_2}} \overline{C_{lmn}^{k_1 k_2 k_3}} C_{l_1 m_1 n_1}^{i_1 i_2 i_3} C_{l_2 m_2 n_2}^{j_1 j_2 j_3} B_{k_1 k_2 k_3}^{i_1 i_2 i_3, j_1 j_2 j_3}.$$

Appendix C. Some lemmas. In this section we present some lemmas as the preparation for the proof of Theorem 1. In the equations appearing in the lemmas below, we always assume that the indices l, m, n satisfy $(l, m, n) \in \mathbb{N} \times \mathbb{Z} \times \mathbb{N}$ and $-l \leqslant m \leqslant l$. Any quantity with indices l', m', n' is considered as zero if $|m'| > l'$ or $l' < 0$ or $n' < 0$ (e.g., $p_{l-1,m,n}$ with $l = m$).

LEMMA 1. *Burnett polynomials $p_{lmn}(\cdot)$ satisfy the following properties:*

(39)
$$v_x^* p_{lmn}(\boldsymbol{v}^*) = \sqrt{n+l+3/2}\gamma_{l+1,m}p_{l+1,m,n}(\boldsymbol{v}^*) - \sqrt{n}\gamma_{l+1,m}p_{l+1,m,n-1}(\boldsymbol{v}^*)$$
$$+ \sqrt{n+l+1/2}\gamma_{lm}p_{l-1,m,n}(\boldsymbol{v}^*) - \sqrt{n+1}\gamma_{lm}p_{l-1,m,n+1}(\boldsymbol{v}^*),$$

(40)
$$v_x^* p_{lm0}(\boldsymbol{v}^*) = \sqrt{l+3/2}\gamma_{l+1,m}p_{l+1,m,0}(\boldsymbol{v}^*) + \sqrt{\frac{1}{l+1/2}}\gamma_{lm}\frac{|\boldsymbol{v}^*|^2}{2}p_{l-1,m,0}(\boldsymbol{v}^*),$$

(41)
$$\frac{\mathrm{d}p_{lmn}(\boldsymbol{v}^*)}{\mathrm{d}v_x^*} = \sqrt{n+l+1/2}\gamma_{lm}p_{l-1,m,n}(\boldsymbol{v}^*) - \sqrt{n}\gamma_{l+1,m}p_{l+1,m,n-1}(\boldsymbol{v}^*),$$

where the γ symbol is defined in (21).

Proof. The proof of this lemma requires the following identities [1, equations (8.5.3), (22.7.31), (22.7.30)]:

$$(l-m+1)P_{l+1}^m(x) = (2l+1)xP_l^m(x) - (l+m)P_{l-1}^m(x),$$
$$xL_n^{(\alpha+1)}(x) = (n+\alpha+1)L_n^{(\alpha)}(x) - (n+1)L_{n+1}^{(\alpha)}(x),$$
$$L_n^{(\alpha-1)}(x) = L_n^{(\alpha)}(x) - L_{n-1}^{(\alpha)}(x).$$

Representing $\boldsymbol{v}^*$ by spherical coordinates $(r\cos\vartheta, r\sin\vartheta\cos\varphi, r\sin\vartheta\sin\varphi)$, we get

$$
\begin{aligned}
v_x^* p_{lmn}(\boldsymbol{v}^*) &= r\cos\vartheta \cdot \sqrt{\frac{2^{1-l}\pi^{3/2}n!}{\Gamma(n+l+3/2)}} L_n^{(l+1/2)}\left(\frac{r^2}{2}\right) r^l \\
&\quad \cdot \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\vartheta)\exp(\mathrm{i}m\varphi) \\
&= \sqrt{\frac{2^{1-l}\pi^{3/2}n!}{\Gamma(n+l+3/2)}}\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} L_n^{(l+1/2)}\left(\frac{r^2}{2}\right) r^{l+1} \\
&\quad \times \left(\frac{l-m+1}{2l+1}P_{l+1}^m(\cos\vartheta) + \frac{l+m}{2l+1}P_{l-1}^m(\cos\vartheta)\right)\exp(\mathrm{i}m\varphi) \\
&= \sqrt{\frac{2^{1-l}\pi^{3/2}n!}{\Gamma(n+l+3/2)}}\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} \\
&\quad \times \left(\frac{l-m+1}{2l+1}\left[L_n^{(l+3/2)}\left(\frac{r^2}{2}\right) - L_{n-1}^{(l+3/2)}\left(\frac{r^2}{2}\right)\right] r^{l+1}P_{l+1}^m(\cos\vartheta)\right. \\
&\quad\quad + \frac{2(l+m)}{2l+1}\left[(n+l+1/2)L_n^{(l-1/2)}\left(\frac{r^2}{2}\right)\right. \\
&\quad\quad\quad \left.\left.- (n+1)L_{n+1}^{(l-1/2)}\left(\frac{r^2}{2}\right)\right] r^{l-1}P_{l-1}^m(\cos\vartheta)\right)\exp(\mathrm{i}m\varphi) \\
&= \sqrt{\frac{2(l-m+1)(l+m+1)(n+l+3/2)}{(2l+1)(2l+3)}}p_{l+1,m,n}(\boldsymbol{v}^*) \\
&\quad - \sqrt{\frac{2(l-m+1)(l+m+1)n}{(2l+1)(2l+3)}}p_{l+1,m,n}(\boldsymbol{v}^*) \\
&\quad + \sqrt{\frac{2(l-m)(l+m)(n+l+1/2)}{(2l+1)(2l-1)}}p_{l-1,m,n}(\boldsymbol{v}^*) \\
&\quad - \sqrt{\frac{2(l-m)(l+m)(n+1)}{(2l+1)(2l-1)}}p_{l-1,m,n+1}(\boldsymbol{v}^*).
\end{aligned}
$$

Equation (39) is a direct result of the above equality by inserting the definitions of γ_{lm} (21). When $n = 0$,

$$
\begin{aligned}
v_x^* p_{lm0}(\boldsymbol{v}^*) &= \sqrt{\frac{2^{1-l}\pi^{3/2}}{\Gamma(l+3/2)}}\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} r^{l+1} \\
&\quad \times \left(\frac{l-m+1}{2l+1}P_{l+1}^m(\cos\vartheta) + \frac{l+m}{2l+1}P_{l-1}^m(\cos\vartheta)\right)\exp(\mathrm{i}m\varphi) \\
&= \sqrt{\frac{(l-m+1)(l+m+1)(l+3/2)}{(2l+1)(2l+3)}}p_{l+1,m,0}(\boldsymbol{v}^*) \\
&\quad + \sqrt{\frac{(l-m)(l+m)}{2(l+1/2)(2l+1)(2l-1)}} r^2 p_{l-1,m,0}(\boldsymbol{v}^*).
\end{aligned}
$$

Again, the equality (40) can be obtained by inserting the definition of γ_{lm}.

Now we prove (41). It is clear that the equality holds for $l = n = 0$. If $l > 0$ or $n > 0$, we just need to compute the following integral for any l', m', n' satisfying $l' + 2n' < l + 2n$:

$$
\int_{\mathbb{R}^3} [p_{l'm'n'}(\boldsymbol{v}^*)]^\dagger \frac{\mathrm{d}p_{lmn}(\boldsymbol{v}^*)}{\mathrm{d}v_x^*} \omega(\boldsymbol{v}^*) \, \mathrm{d}\boldsymbol{v}^*
$$
$$
= - \int_{\mathbb{R}^3} \left(\frac{\mathrm{d}p_{l'm'n'}(\boldsymbol{v}^*)}{\mathrm{d}v_x^*} \right)^\dagger p_{lmn}(\boldsymbol{v}^*) \omega(\boldsymbol{v}^*) \, \mathrm{d}\boldsymbol{v}^* - \int_{\mathbb{R}^3} [p_{l'm'n'}(\boldsymbol{v}^*)]^\dagger p_{lmn}(\boldsymbol{v}^*) \frac{\mathrm{d}\omega(\boldsymbol{v}^*)}{\mathrm{d}v_x^*} \, \mathrm{d}\boldsymbol{v}^*,
$$

where the first term on the right-hand side is zero since $p_{l'm'n'}(\cdot)$ is an orthogonal polynomial. Thus,

$$
\int_{\mathbb{R}^3} [p_{l'm'n'}(\boldsymbol{v}^*)]^\dagger \frac{\mathrm{d}p_{lmn}(\boldsymbol{v}^*)}{\mathrm{d}v_x^*} \omega(\boldsymbol{v}^*) \, \mathrm{d}\boldsymbol{v}^* = - \int_{\mathbb{R}^3} [p_{l'm'n'}(\boldsymbol{v}^*)]^\dagger p_{lmn}(\boldsymbol{v}^*) \frac{\mathrm{d}\omega(\boldsymbol{v}^*)}{\mathrm{d}v_x^*} \, \mathrm{d}\boldsymbol{v}^*
$$
$$
= \int_{\mathbb{R}^3} [p_{l'm'n'}(\boldsymbol{v}^*)]^\dagger v_x^* p_{lmn}(\boldsymbol{v}^*) \omega(\boldsymbol{v}^*) \, \mathrm{d}\boldsymbol{v}^*.
$$

Now we insert (39) into the above equation. By the orthogonality of p_{lmn}, it is not difficult to see that

$$
\int_{\mathbb{R}^3} [p_{l'm'n'}(\boldsymbol{v}^*)]^\dagger \frac{\mathrm{d}p_{lmn}(\boldsymbol{v}^*)}{\mathrm{d}v_x^*} \omega(\boldsymbol{v}^*) \, \mathrm{d}\boldsymbol{v}^*
$$
$$
= \sqrt{n+l+1/2}\,\gamma_{lm}\delta_{l-1,l'}\delta_{mm'}\delta_{nn'} - \sqrt{n}\,\gamma_{l+1,m}\delta_{l+1,l'}\delta_{mm'}\delta_{n+1,n'},
$$

which implies the equality (41). $\square$

An immediate corollary of the above lemma is the parallel properties for $p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\cdot)$.

LEMMA 2. *The polynomials $p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}$ defined in (16) satisfy the following properties:*

(42)

$$
v_x p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})
$$
$$
= \overline{u}_x p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) + \overline{\theta}\left(\sqrt{n+l+3/2}\,\gamma_{l+1,m} p_{l+1,m,n}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) - \sqrt{n+1}\,\gamma_{-l,m} p_{l-1,m,n+1}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right)
$$
$$
+ \sqrt{n+l+1/2}\,\gamma_{-l,m} p_{l-1,m,n}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) - \sqrt{n}\,\gamma_{l+1,m} p_{l+1,m,n-1}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}),
$$

(43)

$$
v_x p_{lm0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) = \sqrt{l+3/2}\,\gamma_{l+1,m} p_{l+1,m,0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) + \sqrt{\frac{1}{l+1/2}}\,\gamma_{lm} \frac{|\boldsymbol{v}|^2}{2} p_{l-1,m,0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}),
$$

(44)

$$
\frac{\mathrm{d}p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})}{\mathrm{d}v_x} = \overline{\theta}^{-1}\left[\sqrt{n+l+1/2}\,\gamma_{lm} p_{l-1,m,n}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) - \sqrt{n}\,\gamma_{l+1,m} p_{l+1,m,n-1}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right].
$$

These equations can be directly obtained from Lemma 1 by replacing $\boldsymbol{v}^*$ with $(\boldsymbol{v} - \overline{\boldsymbol{u}})/\sqrt{\overline{\theta}}$. The detail of the proof is omitted. Note that (42) is the same as (26).

LEMMA 3. *For*

$$
\mathcal{M}(\boldsymbol{v}) = \frac{\rho}{\mathfrak{m}(2\pi\theta)^{3/2}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}|^2}{2\theta} \right),
$$

it holds that

(45)
$$
\mathfrak{m} \int_{\mathbb{R}^3} \left[p_{lm0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^\dagger \mathcal{M}(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} = \rho \left[p_{lm0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{u}) \right]^\dagger.
$$

Proof. The proof uses the following formula [14]:

$$p_{lm0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) = \sum_{\lambda=0}^{l} \sum_{\mu=-\lambda}^{\lambda} \binom{l+m}{\lambda+\mu}^{1/2} \binom{l-m}{\lambda-\mu}^{1/2}$$

$$\times \sqrt{\frac{\Gamma(l-\lambda+1/2)\Gamma(\lambda+1/2)}{\sqrt{\pi}\Gamma(l+1/2)}} p_{\lambda\mu0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{u}) p_{l-\lambda,m-\mu,0}^{[\boldsymbol{u},\overline{\theta}]}(\boldsymbol{v}).$$

Since

$$\mathfrak{m} \int_{\mathbb{R}^3} \left[p_{lm0}^{[\boldsymbol{u},\overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \mathcal{M}(\boldsymbol{v})\, \mathrm{d}\boldsymbol{v}$$

$$= \frac{\rho}{(2\pi\theta)^{3/2}} \sqrt{\frac{2^{1-l}\pi^{3/2}}{\Gamma(l+3/2)}} \int_{\mathbb{R}^3} \left| \frac{\boldsymbol{v}-\boldsymbol{u}}{\sqrt{\overline{\theta}}} \right|^l \left[Y_l^m \left(\frac{\boldsymbol{v}-\boldsymbol{u}}{|\boldsymbol{v}-\boldsymbol{u}|} \right) \right]^{\dagger} \exp\left(-\frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{2\theta} \right) \mathrm{d}\boldsymbol{v}$$

$$= \frac{\rho}{(2\pi\theta)^{3/2}} \sqrt{\frac{2^{1-l}\pi^{3/2}}{\Gamma(l+3/2)}} \int_0^{+\infty} \left(\frac{r}{\sqrt{\overline{\theta}}} \right)^l \exp\left(-\frac{r^2}{2\theta} \right) r^2\, \mathrm{d}r \int_{\mathbb{S}^2} Y_l^m(\boldsymbol{n})\, \mathrm{d}\boldsymbol{n}$$

$$= \frac{\rho}{\theta^{3/2}} \sqrt{\frac{2^{-l}}{\Gamma(l+3/2)\sqrt{\pi}}} \delta_{l0}\delta_{m0} \int_0^{+\infty} \left(\frac{r}{\sqrt{\overline{\theta}}} \right)^l \exp\left(-\frac{r^2}{2\theta} \right) r^2\, \mathrm{d}r = \rho\delta_{l0}\delta_{m0},$$

we obtain

$$\mathfrak{m} \int_{\mathbb{R}^3} \left[p_{lm0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \mathcal{M}(\boldsymbol{v})\, \mathrm{d}\boldsymbol{v} = \sum_{\lambda=0}^{l} \sum_{\mu=-\lambda}^{\lambda} \binom{l+m}{\lambda+\mu}^{1/2} \binom{l-m}{\lambda-\mu}^{1/2}$$

$$\times \sqrt{\frac{\Gamma(l-\lambda+1/2)\Gamma(\lambda+1/2)}{\sqrt{\pi}\Gamma(l+1/2)}} \left[p_{\lambda\mu0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{u}) \right]^{\dagger} \rho\delta_{l-\lambda,0}\delta_{m-\mu,0} = \rho \left[p_{lm0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{u}) \right]^{\dagger},$$

which completes the proof.   □

LEMMA 4. *For* $\mathcal{M}(\boldsymbol{v})$ *defined in* (18), *it holds that*

(46)
$$\begin{aligned}
&\mathfrak{m} \int_{\mathbb{R}^3} v_x \left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \mathcal{M}(\boldsymbol{v})\, \mathrm{d}\boldsymbol{v} \\
&= \mathfrak{m} \int_{\mathbb{R}^3} \left[u_x p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) + \sqrt{n+l+1/2}\gamma_{lm}(\theta/\overline{\theta}) p_{l-1,m,n}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right. \\
&\qquad\qquad \left. - \sqrt{n}\gamma_{l+1,m}(\theta/\overline{\theta}) p_{l+1,m,n-1}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \mathcal{M}(\boldsymbol{v})\, \mathrm{d}\boldsymbol{v}.
\end{aligned}$$

Proof. By integration by parts,

$$\mathfrak{m} \int_{\mathbb{R}^3} v_x \left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \mathcal{M}(\boldsymbol{v})\, \mathrm{d}\boldsymbol{v}$$

$$= \mathfrak{m} \int_{\mathbb{R}^3} u_x \left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \mathcal{M}(\boldsymbol{v})\, \mathrm{d}\boldsymbol{v} - \mathfrak{m}\theta \int_{\mathbb{R}^3} \left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) \right]^{\dagger} \frac{\mathrm{d}\mathcal{M}(\boldsymbol{v})}{\mathrm{d}v_x}\, \mathrm{d}\boldsymbol{v}$$

$$= \mathfrak{m} \int_{\mathbb{R}^3} \left[u_x p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) + \theta\frac{\mathrm{d}p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})}{\mathrm{d}v_x} \right]^{\dagger} \mathcal{M}(\boldsymbol{v})\, \mathrm{d}\boldsymbol{v}.$$

Then (46) can be obtained by inserting (44) into the above equation.   □

Appendix D. Proof of Theorem 1.

Proof. Since $\overline{\theta} > \theta/2$, by straightforward calculation, we get

$$\int_{\mathbb{R}^3} [\mathcal{M}(\boldsymbol{v})]^2 [\omega^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})]^{-1} \, \mathrm{d}\boldsymbol{v} = \frac{\rho^2 \overline{\theta}^3}{\mathfrak{m}[\theta(2\overline{\theta}-\theta)]^{3/2}} \exp\left(\frac{|\boldsymbol{u}-\overline{\boldsymbol{u}}|^2}{2\overline{\theta}-\theta}\right) < +\infty.$$

Therefore the expansion (19) holds for

(47) $$\tilde{\mathcal{M}}_{lmn} = \mathfrak{m}\overline{\theta}^{l+2n} \int_{\mathbb{R}^3} \left[p_{lmn}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\right]^{\dagger} \mathcal{M}(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v}.$$

When $n > 0$, we can apply (42) to get

$$\tilde{\mathcal{M}}_{lmn} = \frac{1}{\sqrt{n}\gamma_{l+1,m}} \Bigg(\sqrt{n+l+3/2}\gamma_{l+2,m}\tilde{\mathcal{M}}_{l+2,m,n-1} - \sqrt{n-1}\gamma_{l+2,m}\overline{\theta}\tilde{\mathcal{M}}_{l+2,m,n-2}$$

$$+ \sqrt{n+l+1/2}\gamma_{l+1,m}\overline{\theta}\tilde{\mathcal{M}}_{l,m,n-1} + \bar{u}_x \tilde{\mathcal{M}}_{l+1,m,n-1}$$

$$- \mathfrak{m}\int_{\mathbb{R}^3} v_x \left[p_{l+1,m,n-1}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v})\right]^{\dagger} \mathcal{M}(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} \Bigg).$$

The integral term in the above equation can be calculated by applying (46), and the result will be (20) after simplification. Similarly, combining (43), (45), and (47) yields the iterative formula (22).

The proof of the initial condition (23) requires the following formula,

$$|\boldsymbol{v}^*|^l Y_l^{\pm l}\left(\frac{\boldsymbol{v}^*}{|\boldsymbol{v}^*|}\right) = \frac{(\mp 1)^l}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}}(v_y^* \pm \mathrm{i}v_z^*)^l,$$

from which we know that
(48)
$$p_{l,\pm l,0}^{[\overline{\boldsymbol{u}},\overline{\theta}]}(\boldsymbol{v}) = \overline{\theta}^{-l}\frac{(\pm 1)^l}{2^l l!}\sqrt{\frac{2^{1-l}\pi^{3/2}}{\Gamma(l+3/2)}}\sqrt{\frac{(2l+1)!}{4\pi}}(v_y^* \pm \mathrm{i}v_z^*)^l = (\mp 1)^l\sqrt{\frac{1}{2^l l!}}\overline{\theta}^{-l}(v_y^* \pm \mathrm{i}v_z^*)^l,$$

where we have used

$$\Gamma(l+3/2) = \frac{(2l+2)!}{4^{l+1}(l+1)!}\sqrt{\pi}$$

in the second equality of (48). Then (23) is an immediate result of (45), (47), and (48). □

REFERENCES

[1] M. ABRAMOWITZ AND I. STEGUN, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, 1964.

[2] A. ALEKSEENKO AND E. JOSYULA, *Deterministic solution of the spatially homogeneous Boltzmann equation using discontinuous Galerkin discretizations in the velocity space*, J. Comput. Phys., 272 (2014), pp. 170–188.

[3] A. ALEKSEENKO AND J. LIMBACHER, *Evaluating high order discontinuous Galerkin discretization of the Boltzmann collision integral in $O(N^2)$ operations using the discrete Fourier transform*, Kinet. Relat. Models, 12 (2019), pp. 703–726.

[4] R. ALONSO, I. M. GAMBA, AND M. TASKOVIC̀, *Exponentially-Tailed Regularity and Time Asymptotic for the Homogeneous Boltzmann Equation*, preprint, https://arxiv.org/abs/1711.06596, 2017.

[5]   P. L. Bhatnagar, E. P. Gross, and M. Krook, *A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems*, Phys. Rev., 94 (1954), pp. 511–525.

[6]   G. A. Bird, *Approach to translational equilibrium in a rigid sphere gas*, Phys. Fluids, 6 (1963), pp. 1518–1519.

[7]   G. A. Bird, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford Eng. Sci. Ser. 42, Clarendon Press, Oxford, 1994.

[8]   A. Bobylev and S. Rjasanow, *Difference scheme for the Boltzmann equation based on the fast Fourier transform*, Eur. J. Mech. B Fluids, 16 (1997), pp. 293–306.

[9]   D. Burnett, *The distribution of molecular velocities and the mean motion in a non-uniform gas*, Proc. Lond. Math. Soc., 40 (1936), pp. 382–435.

[10] Z. Cai, Y. Fan, and Y. Wang, *Burnett spectral method for the spatially homogeneous Boltzmann equation*, Comput. & Fluids, 200 (2020), 104456.

[11] Z. Cai and M. Torrilhon, *Approximation of the linearized Boltzmann collision operator for hard-sphere and inverse-power-law models*, J. Comput. Phys., 295 (2015), pp. 617–643.

[12] Z. Cai and M. Torrilhon, *Numerical simulation of microflows using moment methods with linearized collision operator*, J. Sci. Comput., 74 (2018), pp. 336–374.

[13] Z. Cai and M. Torrilhon, *On the Holway-Weiss debate: Convergence of the Grad-moment-expansion in kinetic gas theory*, Phys. Fluids, 31 (2020), 126105.

[14] M. J. Caola, *Solid harmonics and their addition theorems*, J. Phys. A, 11 (1978), pp. L23–L25.

[15] G. Dechristé and L. Mieussens, *A Cartesian cut cell method for rarefied flow simulations around moving obstacles*, J. Comput. Phys., 314 (2016), pp. 465–488.

[16] G. Dimarco, C. Hauck, and R. Loubère, *A class of low dissipative schemes for solving kinetic equations*, J. Sci. Comput., 78 (2019), pp. 393–432.

[17] G. Dimarco, R. Loubère, J. Narski, and T. Rey, *An efficient numerical method for solving the Boltzmann equation in multidimensions*, J. Comput. Phys., 353 (2018), pp. 46–81.

[18] F. Filbet and S. Jin, *An asymptotic preserving scheme for the ES-BGK model of the Boltzmann equation*, J. Sci. Comput., 46 (2011), pp. 204–224.

[19] F. Filbet, L. Pareschi, and T. Rey, *On steady-state preserving spectral methods for homogeneous Boltzmann equations*, C. R. Math., 353 (2015), pp. 309–314.

[20] I. M. Gamba, J. R. Haack, C. D. Hauck, and J. Hu, *A fast spectral method for the Boltzmann collision operator with general collision kernels*, SIAM J. Sci. Comput., 39 (2017), pp. B658–B674.

[21] I. M. Gamba and S. Rjasanow, *Galerkin-Petrov approach for the Boltzmann equation*, J. Comput. Phys., 366 (2018), pp. 341–365.

[22] I. M. Gamba and S. H. Tharkabhushanam, *Spectral-Lagrangian methods for collisional models of non-equilibrium statistical states*, J. Comput. Phys., 228 (2009), pp. 2012–2036.

[23] H. Grad, *On the kinetic theory of rarefied gases*, Comm. Pure Appl. Math., 2 (1949), pp. 331–407.

[24] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods*, Springer, New York, 2008.

[25] Z. Hu, Z. Cai, and Y. Wang, *Numerical simulation of microflows using Hermite spectral methods*, SIAM J. Sci. Comput., 42 (2020), pp. B105–B134.

[26] J. Huang, K. Xu, and P. Yu, *A unified gas-kinetic scheme for continuum and rarefied flows II: Multi-dimensional cases*, Commun. Comput. Phys., 12 (2012), pp. 662–690.

[27] G. Kitzler and J. Schröberl, *A polynomial spectral method for the spatially homogeneous Boltzmann equation*, SIAM J. Sci. Comput., 41 (2019), pp. B27–B49.

[28] C. Mouhot and L. Pareschi, *Fast algorithms for computing the Boltzmann collision operator*, Math. Comp., 75 (2006), pp. 1833–1852.

[29] L. Pareschi and B. Perthame, *A Fourier spectral method for homogeneous Boltzmann equations*, Transp. Theor. Stat., 25 (1996), pp. 369–382.

[30] C. R. Scullard, A. Hickok, J. O. Sotiris, B. M. Tzolova, R. L. V. Heyningen, and F. R. Graziani, *Adaptive spectral solution method for the Landau and Lenard-Balescu equations*, J. Comput. Phys., 402 (2020), 109110.

[31] E. M. Shakhov, *Generalization of the Krook kinetic relaxation equation*, Fluid Dyn., 3 (1968), pp. 95–96.

[32] T. Tang, *The Hermite spectral method for Gaussian-type functions*, SIAM J. Sci. Comput., 14 (1993), pp. 594–606.

[33] M. Y. Timokhin, H. Struchtrup, A. A. Kokhanchik, and Y. A. Bondar, *Different variants of R13 moment equations applied to the shock-wave structure*, Phys. Fluids, 29 (2017), 037105.

[34] M. TORRILHON, *Convergence study of moment approximations for boundary value problems of the Boltzmann-BGK equation*, Commun. Comput. Phys., 18 (2015), pp. 529–557.

[35] Y. WANG AND Z. CAI, *Approximation of the Boltzmann collision operator based on Hermite spectral method*, J. Comput. Phys., 397 (2019), 108815.

[36] L. WU, C. WHITE, T. SCANLONA, J. REESE, AND Y. ZHANG, *Deterministic numerical solutions of the Boltzmann equation using the fast spectral method*, J. Comput. Phys., 250 (2013), pp. 27–52.

[37] K. XU AND J. HUANG, *An improved unified gas-kinetic scheme and the study of shock structures*, IMA J. Appl. Math., 76 (2011), pp. 698–711.