Commun. Comput. Phys.
doi: 10.4208/cicp.220313.281013a

Vol. **15**, No. 5, pp. 1368-1406
May 2014

Dimension-Reduced Hyperbolic Moment Method for the Boltzmann Equation with BGK-Type Collision

Zhenning Cai[1], Yuwei Fan[1], Ruo Li[2] and Zhonghua Qiao[3,*]

[1] *School of Mathematical Sciences, Peking University, Beijing 100871, China.*
[2] *CAPT, LMAM & School of Mathematical Sciences, Peking University, Beijing 100871, China.*
[3] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong.*

Received 22 March 2013; Accepted (in revised version) 28 October 2013

Available online 5 March 2014

Abstract. We develop the dimension-reduced hyperbolic moment method for the Boltzmann equation, to improve solution efficiency using a numerical regularized moment method for problems with low-dimensional macroscopic variables and high-dimensional microscopic variables. In the present work, we deduce the globally hyperbolic moment equations for the dimension-reduced Boltzmann equation based on the Hermite expansion and a globally hyperbolic regularization. The numbers of Maxwell boundary condition required for well-posedness are studied. The numerical scheme is then developed and an improved projection algorithm between two different Hermite expansion spaces is developed. By solving several benchmark problems, we validate the method developed and demonstrate the significant efficiency improvement by dimension-reduction.

AMS subject classifications: 35L25, 65M08, 76P05, 78M05

Key words: Dimension-reduction, moment system, NRxx, global hyperbolicity, boundary condition, microflow.

1 Introduction

In kinetic theory, gas flows are often categorized into different regimes by the Knudsen number Kn. The classic Navier-Stokes-Fourier (NSF) equations are adequate to model the behavior of fluid in the continuum regime ($Kn < 0.001$), and can capture certain flow features in the slip regime ($0.001 < Kn < 0.1$) if the various velocity-slip technique and

*Corresponding author. *Email addresses:* caizn@pku.edu.cn (Z. Cai), ywfan@pku.edu.cn (Y. Fan), rli@math.pku.edu.cn (R. Li), zqiao@polyu.edu.hk (Z. Qiao)

temperature-jump boundary conditions are used. However, in rarefied fluids or microflows, the gas flows are in the transition regime ($Kn > 0.1$) and the NSF equations fail [20].

It is generally accepted that the Boltzmann equation can capture the physics of gases in transition regimes. Several methods have been devised to solve the Boltzmann equation directly, e.g., direct simulation Monte Carlo (DSMC) due to Bird [2], discrete velocity methods (DVM) [10, 12, 13, 25, 33], and the unified scheme for BGK and Shakhov models [17,31,32]. On the other hand, Grad [14] developed a class of models called the method of moments, to approximate the Boltzmann equation. Based on Grad's idea, a number of moment equations were put forward [5, 15, 16, 18, 26, 29]. In [5], the authors proposed a general numerical method, referred as NRxx method, for numerically solving the moment equations to arbitrary order. In the transition regime, it is well-known that the traditional no-slip wall boundary conditions do not apply, and the Boltzmann equation with the Maxwell boundary conditions [24] can depict the gas state-cf. [20, 21] for details. In [6], the Maxwell boundary condition for the NRxx method is studied, and [8, 9] further demonstrate the effectiveness of NRxx method in rarefied fluids simulation.

However, one flaw of Grad type moment equations is that they are not hyperbolic, which makes the system not well-posed and limits the application of the NRxx method. Recently, a new regularization was developed to make Grad's moment system globally hyperbolic [3, 4], and results in a class of hyperbolic moment equations (HME). A corresponding numerical scheme was proposed in [7], and the Maxwell boundary condition for the NRxx method in [6] can be extended to the HME without any modification. Given this progress both theoretically and numerically, the HME method is a valid and effective method to solve the Boltzmann equation for rarefied fluids and microflows for the one-dimensional case [7].

For a two-dimensional problem, the dimension of the molecular velocity is still 3. In [10], the author has proposed a dimension-reduction technique, by introducing an auxiliary distribution function to reduce the dimension of the molecular velocity the same as that of the space. As is well-known, this technique can reduce the computational cost in numerical simulation.

In the present work, we consider the two-dimensional case, and use the dimension-reduction technique on the Boltzmann equation to get two coupled dimension-reduced Boltzmann equations, and then invoke the NRxx method to achieve a dimension-reduced NRxx method. The globally hyperbolic regularization in [3] is first derived for the dimension-reduced NRxx method to obtain a class of dimension-reduced hyperbolic moment equations (DRHME). Both the BGK model and Shakhov model are studied for the collision term. The Maxwell boundary condition is also studied. In the numerical scheme, classic time splitting is adopted to handle the convection term and the collision term in the dimension-reduced Boltzmann equations, and the finite volume method is applied on the convection part.

During the derivation of the DRHME, the dimension-reduction technique is applied on the Boltzmann equation to get two reduced distribution functions and two coupled

reduced Boltzmann equations. We apply the Hermite expansion to the reduced distribution functions, and substitute the expansion into the reduced Boltzmann equations to get a moment system with infinite moments. Then following the result of the orders of magnitude of the moments in [9], we introduce a reasonable truncation to obtain a moment system with a finite number of equations. However, the moment closure remains to be an important issue. As in [3], we analyse the coefficient matrix of the equations and propose a globally hyperbolic regularization for the moment system. The Maxwell boundary condition for the NRxx method [6], which is also valid for the HME, is then extended for the DRHME. With the hyperbolic regularization, the structure of characteristics is clarified, so it is possible to prove that the number of the boundary conditions we provide and the number of the characteristics waves traveling inward to the domain are the same.

Similar as in [9], to develop the numerical scheme for the DRHME, we use a classic time splitting to split the DRHME into two parts: convection part and collision part. Since the hyperbolic regularization modifies a few terms in some equations of the moment system, these equations may not be written in conservation form. Thus the DLM theory [11] is introduced to deal with the nonconservative part and the numerical scheme in [27], which can be treated as a nonconservative version of the HLL scheme, is applied. As pointed out in [5], it is not trivial to add two distribution functions in spaces with different expansion parameters, so a projection between two different spaces has to be provided. Here we propose a more concise and efficient algorithm to implement the projection, where we only need to solve an ordinary differential system with a constant nilpotent coefficient matrix and the projection is invertible. And rigorous mathematical proof of the well-posedness of the projection is given. For the collision part, both the BGK model [1] and the Shakhov model [28] can be integrated analytically.

Two numerical examples are presented to validate our method. The first one is a one-dimensional shock tube test. It is used to demonstrate the validity of DRHME. The second example is cavity flow simulation. Good agreement with the DSMC solutions [15] shows the capability of the DRHME and some physical phenomenons can be observed.

The rest of the paper is organized as follows. Section 2 introduces the dimension-reduction technique on the Boltzmann equation. In Section 3, we derive the hyperbolic regularized moment system. The Maxwell boundary condition for DRHME is discussed in Section 4. The numerical scheme for the dimension-reduced moment system is developed in Section 5, and numerical examples are presented in Section 6. Finally, we conclude the paper in Section 7.

2 The governing equations

We study the Boltzmann equation with a BGK-type collision term

$$\frac{\partial f}{\partial t}+\sum_{j=1}^{3}\xi_j\frac{\partial f}{\partial x_j}=\frac{1}{\tau}(f^N-f), \tag{2.1}$$

where $f(t,\boldsymbol{x},\boldsymbol{\xi})$ is the velocity distribution function depending on time t, the spatial coordinate $\boldsymbol{x}=(x_1,x_2,x_3)^T\in\mathbb{R}^3$, and the molecular velocity $\boldsymbol{\xi}=(\xi_1,\xi_2,\xi_3)^T\in\mathbb{R}^3$, τ is the relaxation time and f^N is an appropriate distribution function depending on the collision model selected. The macroscopic density ρ, flow velocity $\boldsymbol{u}$, temperature θ and heat flux $\boldsymbol{q}$ of the gas are the first four moments of the distribution function, i.e., such that

$$\rho=\int_{\mathbb{R}^3} f\mathrm{d}\boldsymbol{\xi}, \qquad \rho u_i=\int_{\mathbb{R}^3}\xi_i f\mathrm{d}\boldsymbol{\xi}, \qquad i=1,2,3, \tag{2.2a}$$

$$\frac{3}{2}\rho\theta=\int_{\mathbb{R}^3}\frac{1}{2}|\boldsymbol{\xi}-\boldsymbol{u}|^2 f\mathrm{d}\boldsymbol{\xi}, \qquad q_i=\frac{1}{2}\int_{\mathbb{R}^3}|\boldsymbol{\xi}-\boldsymbol{u}|^2(\xi_i-u_i)f\mathrm{d}\boldsymbol{\xi}, \qquad i=1,2,3. \tag{2.2b}$$

In this paper, two well-known kinetic collision models for f^N: BGK model [1] and Shakhov model [28], are studied, and they both have the form:

$$f^N=\left[1+\frac{(1-\mathrm{Pr})(\boldsymbol{\xi}-\boldsymbol{u})\cdot\boldsymbol{q}}{5\rho\theta^2}\left(\frac{|\boldsymbol{\xi}-\boldsymbol{u}|^2}{\theta}-5\right)\right]f_M. \tag{2.3}$$

It is BGK model if Prandtl number $\mathrm{Pr}\equiv 1$. For Shakhov model, the Prandtl number depends on the gas and is equal to $2/3$ for a monatomic gas; f_M is the local Maxwellian

$$f_M=\frac{\rho}{(2\pi\theta)^{3/2}}\exp\left(-\frac{|\boldsymbol{\xi}-\boldsymbol{u}|^2}{2\theta}\right).$$

In kinetic theory, the issue of boundary conditions is complicated, especially in microflow simulation, but here we adopt the one proposed by Maxwell [24]. For a point $\boldsymbol{x}$ on the wall, suppose the velocity and temperature of the wall are $\boldsymbol{u}^W(t,\boldsymbol{x})$ and $\theta^W(t,\boldsymbol{x})$ at time t, respectively. The theory for hyperbolic problems indicates that boundary conditions have to be imposed for $(\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}<0$, where $\boldsymbol{n}$ is the unit outer normal on the boundary. The boundary condition proposed by Maxwell is then:

$$f^b(t,\boldsymbol{x},\boldsymbol{\xi})=\begin{cases}\chi f_M^W(t,\boldsymbol{x},\boldsymbol{\xi})+(1-\chi)f(t,\boldsymbol{x},\boldsymbol{\xi}^*), & (\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}<0,\\ f(t,\boldsymbol{x},\boldsymbol{\xi}), & (\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}\geq 0,\end{cases} \tag{2.4}$$

where $\chi\in[0,1]$ is a parameter relevant to the gas and material of the walls, and

$$\boldsymbol{\xi}^*=\boldsymbol{\xi}-2((\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n})\boldsymbol{n}, \quad f_M^W(t,\boldsymbol{x},\boldsymbol{\xi})=\frac{\rho^W(t,\boldsymbol{x})}{(2\pi\theta^W(t,\boldsymbol{x}))^{3/2}}\exp\left(-\frac{|\boldsymbol{\xi}-\boldsymbol{u}^W(t,\boldsymbol{x})|^2}{2\theta^W(t,\boldsymbol{x})}\right),$$

where ρ^W is determined by the condition that the normal mass flux on the boundary is zero, i.e.,

$$\int_{(\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}<0}[(\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}][f_M^W(t,\boldsymbol{x},\boldsymbol{\xi})-f(t,\boldsymbol{x},\boldsymbol{\xi}^*)]\mathrm{d}\boldsymbol{\xi}=0.$$

2.1 Dimension-reduced distribution functions

In this paper, we assume that the distribution function $f(t,\boldsymbol{x},\boldsymbol{\xi})$ is homogeneous in x_3 and has plane symmetry in the microscopic velocity ξ_3, i.e.,

$$f(t,\boldsymbol{x},\boldsymbol{\xi})=f(t,x_1,x_2,\boldsymbol{\xi})=f(t,x_1,x_2,\xi_1,\xi_2,\xi_r),\quad \text{where } \xi_r=|\xi_3|.$$

This is known as the "2+3"-Boltzmann equation [22]. Note also that the input data has to fulfill this plane symmetry, so the macroscopic flow velocity u_3 vanishes. With this assumption, we have

$$\frac{\partial f}{\partial x_3}=0,\quad u_3=0. \tag{2.5}$$

To study the numerical method, we introduce the reduced distribution functions [10,33]

$$g(t,x_1,x_2,\xi_1,\xi_2)=\int_{-\infty}^{\infty} f(t,x_1,x_2,\boldsymbol{\xi})\mathrm{d}\xi_3, \tag{2.6a}$$

$$h(t,x_1,x_2,\xi_1,\xi_2)=\int_{-\infty}^{\infty} \frac{\xi_3^2}{2} f(t,x_1,x_2,\boldsymbol{\xi})\mathrm{d}\xi_3. \tag{2.6b}$$

The macroscopic variables ρ, $\boldsymbol{u}$, θ and $\boldsymbol{q}$ are related to the reduced distribution functions by

$$\rho=\int_{\mathbb{R}^2} g\mathrm{d}\xi_1\mathrm{d}\xi_2, \qquad \rho u_i=\int_{\mathbb{R}^2}\xi_i g\mathrm{d}\xi_1\mathrm{d}\xi_2, \tag{2.7a}$$

$$\frac{3}{2}\rho\theta=\int_{\mathbb{R}^2}\Big(\frac{1}{2}|\boldsymbol{c}|^2 g+h\Big)\mathrm{d}\xi_1\mathrm{d}\xi_2, \qquad q_i=\int_{\mathbb{R}^2}(\xi_i-u_i)\Big(\frac{1}{2}|\boldsymbol{c}|^2 g+h\Big)\mathrm{d}\xi_1\mathrm{d}\xi_2, \tag{2.7b}$$

where $\boldsymbol{c}$ satisfies $c_i=\xi_i-u_i$ with $i=1,2$. The gas pressure p and the pressure tenser p_{ij} are defined by

$$p=\rho\theta,\quad p_{ij}=\int_{\mathbb{R}^2}\xi_i\xi_j g\mathrm{d}\xi_1\mathrm{d}\xi_2,\quad i,j=1,2.$$

Multiplying (2.1) by $(1,\xi_3^2/2)^T$ on both sides, and integrating over $\mathbb{R}$ with respect to ξ_3, we obtain the dimension reduced Boltzmann equations

$$\frac{\partial g}{\partial t}+\sum_{j=1}^{2}\xi_j\frac{\partial g}{\partial x_j}=\frac{1}{\tau}(g^N-g), \tag{2.8a}$$

$$\frac{\partial h}{\partial t}+\sum_{j=1}^{2}\xi_j\frac{\partial h}{\partial x_j}=\frac{1}{\tau}(h^N-h), \tag{2.8b}$$

where g^N and h^N stand for $g^N=\int_{\infty}^{\infty}f^N\mathrm{d}\xi_3$ and $h^N=\int_{\infty}^{\infty}2^{-1}\xi_3^2 f^N\mathrm{d}\xi_3$, respectively, and they read:

$$g^N=\left[1+\frac{(1-\mathrm{Pr})\boldsymbol{c}\cdot\boldsymbol{q}}{5\rho\theta^2}\Big(\frac{|\boldsymbol{c}|^2}{\theta}-4\Big)\right]g_M, \tag{2.9a}$$

$$h^N=\left[1+\frac{(1-\mathrm{Pr})\boldsymbol{c}\cdot\boldsymbol{q}}{5\rho\theta^2}\Big(\frac{|\boldsymbol{c}|^2}{\theta}-2\Big)\right]h_M, \tag{2.9b}$$

where $\boldsymbol{q}=(q_1,q_2)^T$. Hereafter without ambiguity notations are equipped with $\boldsymbol{x}=(x_1,x_2)^T$, $\boldsymbol{\xi}=(\xi_1,\xi_2)^T$ and $\boldsymbol{u}=(u_1,u_2)^T$.

By (2.6), the boundary condition (2.4) for $\phi=g,h$ become

$$\phi^b(t,\boldsymbol{x},\boldsymbol{\xi})=\begin{cases}\chi\phi_M^W(t,\boldsymbol{x},\boldsymbol{\xi})+(1-\chi)\phi(t,\boldsymbol{x},\boldsymbol{\xi}^*), & (\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}<0,\\ \phi(t,\boldsymbol{x},\boldsymbol{\xi}), & (\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}\geq 0,\end{cases}\tag{2.10}$$

where $\chi\in[0,1]$ is a parameter for different gases and walls, $\boldsymbol{n}$ is the unit outer normal of the boundary, and

$$g_M^W(t,\boldsymbol{x},\boldsymbol{\xi})=\frac{\rho^W(t,\boldsymbol{x})}{2\pi\theta^W(t,\boldsymbol{x})}\exp\Big(-\frac{|\boldsymbol{\xi}-\boldsymbol{u}^W(t,\boldsymbol{x})|^2}{2\theta^W(t,\boldsymbol{x})}\Big),$$
$$\boldsymbol{\xi}^*=\boldsymbol{\xi}-2((\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n})\boldsymbol{n},\quad h_M^W(t,\boldsymbol{x},\boldsymbol{\xi})=\theta^W(t,\boldsymbol{x})g_M^W(t,\boldsymbol{x},\boldsymbol{\xi})/2,$$

where ρ^W is determined by the condition that the normal mass flux on the boundary is zero, i.e.,

$$\int_{(\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}<0}[(\boldsymbol{\xi}-\boldsymbol{u}^W)\cdot\boldsymbol{n}][g_M^W(t,\boldsymbol{x},\boldsymbol{\xi})-g(t,\boldsymbol{x},\boldsymbol{\xi}^*)]\mathrm{d}\boldsymbol{\xi}=0.\tag{2.11}$$

Now let us consider the reduced distribution functions g and h. It is clear that g and h are coupled together due to the collision term on the right-hand side of the dimension-reduced Boltzmann equations (2.8). In the collision term, the local Maxwellian plays an important role, but the forms of both the local Maxwellian of g and h are not straightforward. Briefly, noting that $h_M=\theta g_M/2$ we introduce

$$k_{\theta'}=h-\theta' g/2,\tag{2.12}$$

where θ' is an arbitrary function depending on the time t and the spatial coordinate $\boldsymbol{x}$. In particular, when $\theta'=\theta$ we have

$$k_{\theta,M}=h_M-\theta g_M/2=0.$$

Since the functions $\{g,h\}$ and $\{g,k_\theta\}$ can be mutually transformed, the function k_θ is to be used in the analysis in the next section. For convenience, the subscript of $k_{\theta'}$ will be omitted hereafter, and if not specified explicitly k stands for k_θ.

3 Dimension-reduced hyperbolic moment system

In 1949, Grad [14] introduced a sequence of approximation to the Boltzmann equation by expanding the distribution function in terms of Hermite polynomials in velocity space. Recently, we developed the method more generally, and referred it as the NRxx method [5, 9]. In this section, we aim to extend the NRxx method to the dimension-reduced Boltzmann equations (2.8) and develop a hyperbolic regularization for it. First,

we derive a moment system with an infinite number of equations. Then with help of the result of the orders of magnitude of the moments in [9], a reasonable truncation is proposed to obtain a moment system with a finite number of equations. Finally, to close the moment system and address the fatal flaw of the moment system (lack of hyperbolicity), a globally hyperbolic regularization is proposed. Moreover, the kinetic boundary condition (2.10) is also studied.

3.1 Hermite expansion of the reduced distribution functions

We expand the reduced distribution functions g, h and k into Hermite polynomial series as

$$\phi(t,\boldsymbol{x},\boldsymbol{\xi}) = \sum_{\alpha\in\mathbb{N}^2} \phi_\alpha(t,\boldsymbol{x})\mathcal{H}_{\theta(t,\boldsymbol{x}),\alpha}(\boldsymbol{v}), \quad \phi = g,h,k, \tag{3.1}$$

where $\alpha=(\alpha_1,\alpha_2)$ is a 2-dimensional multi-index, and $\boldsymbol{v}$ is dimensionless velocity

$$\boldsymbol{v} = \frac{\boldsymbol{\xi}-\boldsymbol{u}(t,\boldsymbol{x})}{\sqrt{\theta(t,\boldsymbol{x})}}. \tag{3.2}$$

The basis functions $\mathcal{H}_{\theta,\alpha}$ are defined as

$$\mathcal{H}_{\theta,\alpha}(\boldsymbol{v}) = \prod_{d=1}^{2} \frac{1}{\sqrt{2\pi}}\theta^{-\frac{\alpha_d+1}{2}} He_{\alpha_d}(v_d)\exp\Big(-\frac{v_d^2}{2}\Big),$$

where $He_n(x)$ is the n-th degree Hermite polynomial defined by

$$He_n(x) = (-1)^n \exp\Big(\frac{x^2}{2}\Big)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\exp\Big(-\frac{x^2}{2}\Big), \quad n\in\mathbb{N}.$$

It is clear that

$$k_\alpha = h_\alpha - \frac{\theta}{2}g_\alpha. \tag{3.3}$$

If we use other arbitrary variables $\boldsymbol{u}'(t,\boldsymbol{x})$ and $\theta'(t,\boldsymbol{x})$ to expand the reduced distribution functions as

$$\phi(t,\boldsymbol{x},\boldsymbol{\xi}) = \sum_{\alpha\in\mathbb{N}^2} \phi'_\alpha(t,\boldsymbol{x})\mathcal{H}_{\theta'(t,\boldsymbol{x}),\alpha}(\boldsymbol{v}'), \quad \phi = g,h,k_{\theta'}, \tag{3.4}$$

where $\boldsymbol{v}'$ is

$$\boldsymbol{v}' = \frac{\boldsymbol{\xi}-\boldsymbol{u}'(t,\boldsymbol{x})}{\sqrt{\theta'(t,\boldsymbol{x})}},$$

then the relations between the macroscopic quantities ρ, $\boldsymbol{u}$, θ, ϕ_α and $\boldsymbol{u}'$, θ', ϕ'_α are:

$$\rho = g_0 = g'_0, \tag{3.5a}$$

$$\rho\boldsymbol{u} = \rho\boldsymbol{u}' + (g'_{e_d})_{d=1,2}^T, \tag{3.5b}$$

$$\rho|\boldsymbol{u}-\boldsymbol{u}'|^2 + 3\rho\theta = 2h_0 + 2g_{2e_1} + 2g_{2e_2} + 2\rho\theta', \tag{3.5c}$$

where e_d, $d=1,2$, is the unit multi-index with its d-th entry 1. Similar as (3.3), we have

$$k_{\theta',\alpha}=h_\alpha-\frac{\theta'}{2}g_\alpha. \tag{3.6}$$

If $\boldsymbol{u}'=\boldsymbol{u}$ and $\theta'=\theta$, then Eq. (3.5) gives

$$g_0=\rho,\quad g_{e_d}=0,\quad d=1,2,\quad h_0+g_{2e_1}+g_{2e_2}=\frac{\rho\theta}{2}. \tag{3.7}$$

Using the same trick on p_{ij}, $i,j=1,2$ and $\boldsymbol{q}$, we have

$$p_{ij}=(1+\delta_{ij})g_{e_i+e_j}+\delta_{ij}\rho\theta,\quad q_i=2g_{3e_i}+\sum_{j=1}^{2}g_{e_i+2e_j}+h_{e_i}, \tag{3.8}$$

and substituting (3.3) into the last equation of (3.7) gives

$$k_0+g_{2e_1}+g_{2e_2}=0. \tag{3.9}$$

It is notable that three temperatures θ defined in (2.7), including the subscript in k_θ and the subscript in $\mathcal{H}_{\theta,\alpha}$, have been used. The latter two always are kept the same, equal to the temperature defined in (2.7), if not otherwise stated.

For convenience, we list some useful relations of Hermite polynomials as follows.

1. Orthogonality: $\int_{\mathbb{R}}He_m(x)He_n(x)\exp(-x^2/2)\mathrm{d}x=m!\sqrt{2\pi}\delta_{m,n}$.
2. Recursion relation: $He_{n+1}(x)=xHe_n(x)-nHe_{n-1}(x)$.
3. Differential relation: $He_n'(x)=nHe_{n-1}(x)$.

Due to the orthogonality of Hermite polynomials, the moments ϕ_α can be denoted by

$$\phi_\alpha=\frac{2\pi\theta^{|\alpha|+2}}{\alpha!}\int_{\mathbb{R}^2}\phi\mathcal{H}_{\theta,\alpha}(\boldsymbol{v})\exp\Big(\frac{\boldsymbol{v}^2}{2}\Big)\mathrm{d}\boldsymbol{v},\quad \phi=g,h \text{ or } k, \tag{3.10}$$

where $\boldsymbol{v}$ is defined in (3.2), and $|\alpha|=\alpha_1+\alpha_2$, $\alpha!=\alpha_1!\alpha_2!$.

3.2 Moment equations

In this subsection, we derive equations for the moments $\{g_\alpha,h_\alpha\}_{\alpha\in\mathbb{N}^2}$ and $\{k_\alpha\}_{\alpha\in\mathbb{N}^2}$. The strategy is to substitute the Hermite expansion (3.1) into the dimension reduced Boltzmann equations (2.8), and then match the coefficients of the same basis functions to get the equations of $\{g_\alpha,h_\alpha\}$. The relation (3.3) is then invoked to obtain the equations of $\{k_\alpha\}$. By (3.7), we have

$$g_M=g_0\mathcal{H}_{\theta,0},\quad h_M=\frac{\theta}{2}g_0\mathcal{H}_{\theta,0}.$$

So the collision term yields

$$\frac{1}{\tau}(g^N-g)=\frac{1-\mathrm{Pr}}{5\tau}\sum_{i,j=1}^{2}q_i\mathcal{H}_{\theta,e_i+2e_j}-\frac{1}{\tau}\sum_{|\alpha|\geq 2}g_\alpha\mathcal{H}_{\theta,\alpha}, \tag{3.11a}$$

$$\begin{aligned}\frac{1}{\tau}(h^N-h)=&\frac{1-\mathrm{Pr}}{5\tau}\sum_{i=1}^{2}q_i\left(2\mathcal{H}_{\theta,e_i}+\sum_{j=1}^{2}\theta\mathcal{H}_{\theta,e_i+2e_j}\right)\\&+\frac{1}{\tau}\left(\frac{1}{2}\theta g_0-h_0\right)\mathcal{H}_{\theta,0}-\frac{1}{\tau}\sum_{|\alpha|\geq 1}h_\alpha\mathcal{H}_{\theta,\alpha}.\end{aligned} \tag{3.11b}$$

Substituting the Hermite expansion (3.1) into the left-hand side of the dimension-reduced Boltzmann equations (2.8), we have

$$\begin{aligned}\sum_{\alpha\in\mathbb{N}^2}\Bigg\{&\left(\frac{\partial\phi_\alpha}{\partial t}+\sum_{d=1}^{2}\frac{\partial u_d}{\partial t}\phi_{\alpha-e_d}+\frac{1}{2}\frac{\partial\theta}{\partial t}\sum_{d=1}^{2}\phi_{\alpha-2e_d}\right)\\&+\sum_{j=1}^{2}\Bigg[\left(\theta\frac{\partial\phi_{\alpha-e_j}}{\partial x_j}+u_j\frac{\partial\phi_\alpha}{\partial x_j}+(\alpha_j+1)\frac{\partial\phi_{\alpha+e_j}}{\partial x_j}\right)\\&+\sum_{d=1}^{2}\frac{\partial u_d}{\partial x_j}\left(\theta\phi_{\alpha-e_d-e_j}+u_j\phi_{\alpha-e_d}+(\alpha_j+1)\phi_{\alpha-e_d+e_j}\right)\\&+\frac{1}{2}\frac{\partial\theta}{\partial x_j}\sum_{d=1}^{2}\left(\theta\phi_{\alpha-2e_d-e_j}+u_j\phi_{\alpha-2e_d}+(\alpha_j+1)\phi_{\alpha-2e_d+e_j}\right)\Bigg]\Bigg\}\mathcal{H}_{\theta,\alpha},\end{aligned} \tag{3.12}$$

where $\phi=g$ or h.

Collecting (3.11) and (3.12), and comparing the coefficients of the basis functions $\mathcal{H}_{\theta,\alpha}$, we obtain the following moment equations:

$$\begin{aligned}&\frac{\partial\phi_\alpha}{\partial t}+\sum_{d=1}^{2}\frac{\partial u_d}{\partial t}\phi_{\alpha-e_d}+\frac{1}{2}\frac{\partial\theta}{\partial t}\sum_{d=1}^{2}\phi_{\alpha-2e_d}\\&\quad+\sum_{j=1}^{2}\Bigg[\left(\theta\frac{\partial\phi_{\alpha-e_j}}{\partial x_j}+u_j\frac{\partial\phi_\alpha}{\partial x_j}+(\alpha_j+1)\frac{\partial\phi_{\alpha+e_j}}{\partial x_j}\right)\\&\quad+\sum_{d=1}^{2}\frac{\partial u_d}{\partial x_j}\left(\theta\phi_{\alpha-e_d-e_j}+u_j\phi_{\alpha-e_d}+(\alpha_j+1)\phi_{\alpha-e_d+e_j}\right)\\&\quad+\frac{1}{2}\frac{\partial\theta}{\partial x_j}\sum_{d=1}^{2}\left(\theta\phi_{\alpha-2e_d-e_j}+u_j\phi_{\alpha-2e_d}+(\alpha_j+1)\phi_{\alpha-2e_d+e_j}\right)\Bigg]\\=&\frac{1}{\tau}\left(\Delta_\phi(\alpha)-\phi_\alpha\right),\quad \phi=g,h,\end{aligned} \tag{3.13}$$

where $\Delta_\phi(\alpha)$ denotes, for $i,j=1,2$,

$$\Delta_g(\alpha)=\begin{cases} g_0, & \alpha=0,\\ (1-\mathrm{Pr})q_i/5, & \alpha=e_i+2e_j,\\ 0, & \text{otherwise,}\end{cases} \qquad \Delta_h(\alpha)=\begin{cases} \theta g_0/2, & \alpha=0,\\ 2(1-\mathrm{Pr})q_i/5, & \alpha=e_i,\\ (1-\mathrm{Pr})\theta q_i/5, & \alpha=e_i+2e_j,\\ 0, & \text{otherwise.}\end{cases}$$

With (3.3), we have

$$\frac{\partial k_\alpha}{\partial s}=\frac{\partial h_\alpha}{\partial s}-\frac{\theta}{2}\frac{\partial g_\alpha}{\partial s}-\frac{g_\alpha}{2}\frac{\partial \theta}{\partial s}, \quad s=t,x_1,x_2.$$

Adding (3.13) with $\phi=h$ to $-\theta/2$ times (3.13) with $\phi=g$, we obtain

$$\begin{aligned}&\frac{\partial k_\alpha}{\partial t}+\sum_{d=1}^{2}\frac{\partial u_d}{\partial t}k_{\alpha-e_d}+\frac{1}{2}\frac{\partial\theta}{\partial t}\Big(\sum_{d=1}^{2}k_{\alpha-2e_d}+g_\alpha\Big)\\ &\quad+\sum_{j=1}^{2}\Big[\Big(\theta\frac{\partial k_{\alpha-e_j}}{\partial x_j}+u_j\frac{\partial k_\alpha}{\partial x_j}+(\alpha_j+1)\frac{\partial k_{\alpha+e_j}}{\partial x_j}\Big)\\ &\quad+\sum_{d=1}^{2}\frac{\partial u_d}{\partial x_j}\Big(\theta k_{\alpha-e_d-e_j}+u_j k_{\alpha-e_d}+(\alpha_j+1)k_{\alpha-e_d+e_j}\Big)\\ &\quad+\frac{1}{2}\frac{\partial\theta}{\partial x_j}\sum_{d=1}^{2}\Big(\theta k_{\alpha-2e_d-e_j}+u_j k_{\alpha-2e_d}+(\alpha_j+1)k_{\alpha-2e_d+e_j}\Big)\\ &\quad+\frac{1}{2}\frac{\partial\theta}{\partial x_j}\Big(\frac{\theta}{2}g_{\alpha-e_j}+\frac{u_j}{2}g_\alpha+\frac{\alpha_j+1}{2}g_{\alpha+e_j}\Big)\Big]=\frac{1}{\tau}(\Delta_k(\alpha)-k_\alpha),\end{aligned} \tag{3.14}$$

where

$$\Delta_k(\alpha)=\begin{cases} 2(1-\mathrm{Pr})q_i/5, & \alpha=e_i, \quad i=1,2,\\ 0, & \text{otherwise.}\end{cases}$$

In particular, on setting $\alpha=0$ and $\phi=g$ in (3.13) we deduce the mass conservation

$$\frac{\partial\rho}{\partial t}+\sum_{j=1}^{2}\Big(u_j\frac{\partial\rho}{\partial x_j}+\rho\frac{\partial u_j}{\partial x_j}\Big)=0. \tag{3.15}$$

By setting $\alpha=e_d$, with $d=1,2$ and $\phi=g$ in (3.13), noting that $g_{e_d}=0$, and using Eq. (3.8), we obtain

$$\rho\Big(\frac{\partial u_d}{\partial t}+\sum_{j=1}^{2}u_j\frac{\partial u_d}{\partial x_j}\Big)+\sum_{j=1}^{2}\frac{\partial p_{jd}}{\partial x_j}=0. \tag{3.16}$$

By setting $\alpha=2e_d$, with $d=1,2$ and $\phi=g$ in (3.13) and noting that $g_{e_d}=0$, we obtain

$$\frac{\partial g_{2e_d}}{\partial t}+\frac{\rho}{2}\left(\frac{\partial\theta}{\partial t}+\sum_{j=1}^{2}u_j\frac{\partial\theta}{\partial x_j}\right)+\rho\theta\frac{\partial u_d}{\partial x_d}+\sum_{j,l=1}^{2}(1+\alpha_j)g_{2e_d-e_l+e_j}\frac{\partial u_l}{\partial x_j}$$
$$+\sum_{j=1}^{2}u_j\frac{\partial g_{2e_d}}{\partial x_j}+(1+2\delta_{jd})\frac{\partial g_{2e_d+e_j}}{\partial x_j}=-\frac{1}{\tau}g_{2e_d}, \tag{3.17}$$

and by setting $\alpha=0$ in (3.14), we obtain

$$\frac{\partial k_0}{\partial t}+\frac{\rho}{2}\frac{\partial\theta}{\partial t}+\sum_{j=1}^{2}\left(u_j\frac{\partial k_0}{\partial x_j}+\frac{\partial k_{e_j}}{\partial x_j}+k_0\frac{\partial u_j}{\partial x_j}+\frac{\rho u_j}{4}\frac{\partial\theta}{\partial x_j}\right)=-\frac{1}{\tau}k_0. \tag{3.18}$$

From (3.9) and (3.8), on adding (3.17) with $d=1,2$ to (3.18) we obtain

$$\rho\left(\frac{\partial\theta}{\partial t}+\sum_{j=1}^{2}u_j\frac{\partial\theta}{\partial x_j}\right)+\frac{2}{3}\sum_{j=1}^{2}\left(\frac{\partial q_j}{\partial x_j}+\sum_{d=1}^{2}p_{jd}\frac{\partial u_d}{\partial x_j}\right)=0. \tag{3.19}$$

Eqs. (3.15), (3.16) and (3.19) are the classical hydrodynamic equations, where right-hand sides are all zero, so the collision term does not affect the variables ρ, $\boldsymbol{u}$ and θ.

3.3 Truncation

Since the moment system (3.13) and (3.14) contains an infinite number of equations, a truncation has to be applied. However, (3.13) and (3.14) indicate that the governing equations of ϕ_α with $\phi=g,k$ are related to $\phi_{\alpha+e_j}$, $j=1,2$, so the resulting system is not closed. One simple way is to choose two positive integer $M_g\geq 3$, $M_k\geq 1$ and let $\{g_\alpha\}_{|\alpha|\leqslant M_g}\bigcup\{k_\alpha\}_{|\alpha|\leq M_k}$ be a finite subset of the moments, and $g_\alpha=0$ with $|\alpha|>M_g$ and $k_\alpha=0$ with $|\alpha|>M_k$. This is precisely what Grad [14] has done, and it preserves the Galilean invariance on the space $\boldsymbol{x}$. However, the relation between M_g and M_k is not yet clarified. In this subsection, we use the result of the order of magnitude of moments in [9], and choose the truncation, such that the order of magnitude of $\{k_\alpha\}|_{|\alpha|=M_k}$ equals that of $\{g_\alpha\}|_{|\alpha|=M_g}$.

Same as that in [9], we expand the distribution function f into Hermite polynomial series as

$$f(t,\boldsymbol{x},\boldsymbol{\xi})=\sum_{\beta\in\mathbb{N}^3}f_\beta(t,\boldsymbol{x})\mathcal{H}_{\theta,\beta}(\tilde{v}),$$

here β is a 3-dimensional multi-index, and $\tilde{v}$ is dimensionless velocity in 3-dimension microscopic space

$$\tilde{v}_i=\frac{\xi_i-u_i}{\sqrt{\theta}},\quad i=1,2,3.$$

The basis functions $\mathcal{H}_{\theta,\beta}$ are defined as

$$\mathcal{H}_{\theta,\beta}=\prod_{d=1}^{3}\frac{1}{2\pi}\theta^{-\frac{\beta_d+1}{2}}He_{\beta_d}(\tilde{v}_d)\exp\Big(-\frac{\tilde{v}_d^2}{2}\Big).$$

Then (3.10) indicates

$$\begin{aligned}g_\alpha&=\frac{2\pi\theta^{|\alpha|+2}}{\alpha!}\int_{\mathbb{R}^2}g\mathcal{H}_{\theta,\alpha}(\boldsymbol{v})\exp\Big(\frac{\boldsymbol{v}^2}{2}\Big)\mathrm{d}\boldsymbol{v}\\&=\frac{(2\pi)^{3/2}\theta^{|\alpha|+3}}{\alpha!}\int_{\mathbb{R}^3}f\mathcal{H}_{\theta,(\alpha,0)}(\tilde{\boldsymbol{v}})\exp\Big(\frac{\tilde{\boldsymbol{v}}^2}{2}\Big)\mathrm{d}\tilde{\boldsymbol{v}}=f_{(\alpha,0)}.\end{aligned}\tag{3.20}$$

Similarly, we can obtain

$$k_\alpha=f_{(\alpha,2)}.\tag{3.21}$$

In [9], the authors point out that the order of magnitude of f_β is $f_\beta=\mathcal{O}(\tau^{\lceil|\beta|/3\rceil})$. Thus, we have

$$g_\alpha=\mathcal{O}(\tau^{\lceil|\alpha|/3\rceil}),\quad k_\alpha=\mathcal{O}(\tau^{\lceil(|\alpha|+2)/3\rceil}).\tag{3.22}$$

Hence, k_α, $|\alpha|=m$ has the same order of magnitude as g_α, $|\alpha|=m+2$. It is then reasonable to choose the truncation with $M_g=M$, $M_k=M-2$, and let $\{g_\alpha\}_{|\alpha|\le M}\bigcup\{k_\alpha\}_{|\alpha|\le M-2}$ as the finite subset, and discard all equations containing the term $\partial g_\alpha/\partial t$ with $|\alpha|>M$ and $\partial k_\alpha/\partial t$ with $|\alpha|>M-2$. Since the moments set $\{g_\alpha,h_\alpha\}$ and $\{g_\alpha,k_\alpha\}$ are equivalent, we let the moments set be $\{g_\alpha\}_{|\alpha|\le M}\bigcup\{h_\alpha\}_{|\alpha|\le M-2}$ for a truncated system.

Remark 3.1. In the truncation, we take $g_\alpha=0$ with $|\alpha|>M$ and $k_\alpha=0$ with $|\alpha|>M-2$. Since $h_\alpha=k_\alpha+\theta g_\alpha/2$, we have

$$h_\alpha=\begin{cases}\theta g_\alpha/2+k_\alpha, & |\alpha|\le M-2,\\ \theta g_\alpha/2, & M-2<|\alpha|\le M,\\ 0, & |\alpha|>M.\end{cases}$$

Thus we let $M_h=M$. In a numerical scheme, we need only to update g_α with $|\alpha|\le M$ and h_α with $|\alpha|\le M-2$, for h_α with $M-2<|\alpha|\le M$ can be obtained from g_α. Hereafter, we keep in mind that h_α with $M-2<|\alpha|\le M$ is not zero but $\theta g_\alpha/2$.

3.4 Closure with globally hyperbolic regularization

It is well-known that lack of hyperbolicity is one of the flaws of the Grad-type moment method. Many works has been carried out on this topic, see e.g., [3, 9, 29, 30]. Here we study the hyperbolicity of the moment system obtained above, in the context of [3]. And we need only investigate the equations without the collision term.

Substituting (3.16) and (3.19) into (3.13) to eliminate the time derivatives of $\boldsymbol{u}$ and θ, we obtain

$$\begin{aligned}&\frac{\partial g_\alpha}{\partial t}+\sum_{j=1}^{2}\Big(\theta\frac{\partial g_{\alpha-e_j}}{\partial x_j}+u_j\frac{\partial g_\alpha}{\partial x_j}+(\alpha_j+1)\frac{\partial g_{\alpha+e_j}}{\partial x_j}\Big)+\sum_{j=1}^{2}\Big(-\frac{\theta}{2\rho}C_{\theta,\alpha}^{(j)}\Big)\frac{\partial \rho}{\partial x_j}\\&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\frac{\partial u_d}{\partial x_j}\Big(\theta g_{\alpha-e_d-e_j}+(\alpha_j+1)g_{\alpha-e_d+e_j}-\frac{C_\alpha}{2\rho}p_{jd}\Big)\\&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\Big(\Big(-\frac{g_{\alpha-e_d}}{\rho}\Big)\frac{\partial p_{jd}}{\partial x_j}+\frac{C_{\theta,\alpha}^{(j)}}{6\rho}\frac{\partial p_{dd}}{\partial x_j}\Big)+\Big(-\frac{C_\alpha}{3\rho}\Big)\sum_{j=1}^{2}\frac{\partial q_j}{\partial x_j}=0,\end{aligned}\tag{3.23}$$

where C_α and $C_{\theta,\alpha}^{(j)}$ are defined by

$$C_\alpha=\sum_{k=1}^{2}g_{\alpha-2e_k},\tag{3.24a}$$

$$C_{\theta,\alpha}^{(j)}=\sum_{k=1}^{2}\big(\theta g_{\alpha-2e_k-e_j}+(\alpha_j+1)g_{\alpha-2e_k+e_j}\big).\tag{3.24b}$$

Collecting (3.8), (3.15), (3.19) and (3.23), we get for $i=1,2$,

$$\begin{aligned}&\frac{\partial p_{ii}/2}{\partial t}+\sum_{j=1}^{2}u_j\frac{\partial p_{ii}/2}{\partial x_j}+\sum_{j=1}^{2}\Big(\frac{1}{2}+\delta_{ij}\Big)\rho\theta\frac{\partial u_j}{\partial x_j}\\&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}(2\delta_{ij}+1)g_{2e_i-e_d+e_j}\frac{\partial u_d}{\partial x_j}+\sum_{j=1}^{2}(2\delta_{ij}+1)\frac{\partial g_{2e_i+e_j}}{\partial x_j}=0.\end{aligned}\tag{3.25}$$

In particular, as in [3], we regularize Eq. (3.23) to obtain

$$\begin{aligned}&\frac{\partial g_\alpha}{\partial t}+\sum_{j=1}^{2}\Big(\theta\frac{\partial g_{\alpha-e_j}}{\partial x_j}+u_j\frac{\partial g_\alpha}{\partial x_j}+(1-\delta_{\alpha,M})(\alpha_j+1)\frac{\partial g_{\alpha+e_j}}{\partial x_j}\Big)+\sum_{j=1}^{2}\Big(-\frac{\theta}{2\rho}C_{\theta,\alpha}^{(j)}\Big)\frac{\partial \rho}{\partial x_j}\\&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\frac{\partial u_d}{\partial x_j}\Big(\theta g_{\alpha-e_d-e_j}+(1-\delta_{\alpha,M})(\alpha_j+1)g_{\alpha-e_d+e_j}-\frac{C_\alpha}{2\rho}p_{jd}\Big)\\&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\Big(\Big(-\frac{g_{\alpha-e_d}}{\rho}\Big)\frac{\partial p_{jd}}{\partial x_j}+\frac{C_{\theta,\alpha}^{(j)}}{6\rho}\frac{\partial p_{dd}}{\partial x_j}\Big)+\Big(-\frac{C_\alpha}{3\rho}\Big)\sum_{j=1}^{2}\frac{\partial q_j}{\partial x_j}=0,\end{aligned}\tag{3.26}$$

where C_α is defined in (3.24a), while $C_{\theta,\alpha}^{(j)}$ is redefined as

$$C_{\theta,\alpha}^{(j)}=\sum_{k=1}^{2}\big(\theta g_{\alpha-2e_k-e_j}+(1-\delta_{\alpha,M})(\alpha_j+1)g_{\alpha-2e_k+e_j}\big),$$

and

$$\delta_{\alpha,M}=\begin{cases} 0, & |\alpha|<M, \\ 1, & |\alpha|=M. \end{cases}$$

Following [3], we define $\mathcal{N}(\alpha)=(\alpha_1+\alpha_2+1)(\alpha_1+\alpha_2)/2+\alpha_2+1$, and then the cardinality of $\{g_\alpha||\alpha|\leq M\}$ is $N_g=\mathcal{N}(Me_2)$. Let $\boldsymbol{w}^{(g)}\in\mathbb{R}^{N_g}$ and for each $i=1,2$,

$$w_1^{(g)}=\rho, \qquad w_{\mathcal{N}(e_i)}^{(g)}=u_i, \tag{3.27a}$$

$$w_{\mathcal{N}(2e_i)}^{(g)}=\frac{p_{ii}}{2}, \qquad w_{\mathcal{N}(e_1+e_2)}^{(g)}=p_{12}, \tag{3.27b}$$

$$w_{\mathcal{N}(\alpha)}^{(g)}=g_\alpha, \quad 3\leq|\alpha|\leq M. \tag{3.27c}$$

Combining (3.15) with (3.16), (3.25) and (3.26), we have a quasi-linear system

$$\frac{\partial \boldsymbol{w}^{(g)}}{\partial t}+\sum_{j=1}^{2}\mathbf{A}_j^{(g)}\frac{\partial \boldsymbol{w}^{(g)}}{\partial x_1}=0. \tag{3.28}$$

For convenience, we first recall an essential result proposed in [3] in the following lemma:

Lemma 3.1. *For any unit vector $\boldsymbol{n}=(n_1,n_2)^T\in\mathbb{R}^2$, the matrix $\sum_{j=1}^2 n_j\mathbf{A}_j^{(g)}$ is diagonalisable. Precisely, its characteristic polynomial is*

$$\left|\lambda\mathbf{I}-\sum_{j=1}^{2}n_j\mathbf{A}_j^{(g)}\right|=\prod_{m=1}^{M+1}\theta^{m/2}He_m\Big(\frac{\lambda-\boldsymbol{u}\cdot\boldsymbol{n}}{\sqrt{\theta}}\Big), \tag{3.29}$$

and its eigenvalues are as

$$\boldsymbol{u}\cdot\boldsymbol{n}+\mathrm{C}_{i,m}\sqrt{\theta}, \quad 1\leq i\leq m\leq M+1, \tag{3.30}$$

where $\mathrm{C}_{i,m}$ is the i-th root of $He_m(z)$, noticing that $He_m(z)$, $m\in\mathbb{N}$ has m different zeros, which read $\mathrm{C}_{1,m},\cdots,\mathrm{C}_{m,m}$, and satisfy $\mathrm{C}_{1,m}<\cdots<\mathrm{C}_{m,m}$.

Now let us study the moment equations of k_α (3.14). As before, we use (3.16) and (3.19) to eliminate the time derivatives of $\boldsymbol{u}$ and θ in (3.14) and modify the equation as

$$\begin{aligned}
&\frac{\partial k_\alpha}{\partial t}+\sum_{j=1}^{2}\Big(\theta\frac{\partial k_{\alpha-e_j}}{\partial x_j}+u_j\frac{\partial k_\alpha}{\partial x_j}+(1-\delta_{\alpha,M})(\alpha_j+1)\frac{\partial k_{\alpha+e_j}}{\partial x_j}\Big)+\sum_{j=1}^{2}\Big(-\frac{\theta}{2\rho}C_{\theta,\alpha}^{(j)}\Big)\frac{\partial\rho}{\partial x_j}\\
&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\frac{\partial u_d}{\partial x_j}\Big(\theta k_{\alpha-e_d-e_j}+(1-\delta_{\alpha,M})(\alpha_j+1)k_{\alpha-e_d+e_j}-\frac{C_\alpha}{3\rho}p_{jd}\Big)\\
&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\Big(\Big(-\frac{k_{\alpha-e_d}}{\rho}\Big)\frac{\partial p_{jd}}{\partial x_j}+\frac{C_{\theta,\alpha}^{(j)}}{6\rho}\frac{\partial p_{dd}}{\partial x_j}\Big)+\Big(-\frac{C_\alpha}{3\rho}\Big)\sum_{j=1}^{2}\frac{\partial q_j}{\partial x_j}=0,
\end{aligned} \tag{3.31}$$

where C_α and $C^{(j)}_{\theta,\alpha}$ are redefined as

$$C_\alpha=\sum_{k=1}^{2}k_{\alpha-2e_k}+g_\alpha,$$
$$C^{(j)}_{\theta,\alpha}=\theta\Big(\sum_{k=1}^{2}k_{\alpha-2e_k-e_j}+g_{\alpha-e_j}\Big)+(1-\delta_{\alpha,M})(\alpha_j+1)\Big(\sum_{k=1}^{2}k_{\alpha-2e_k+e_j}+g_{\alpha+e_j}\Big).$$

We denote by $N_k=\mathcal{N}((M-2)e_2)$ the total number of elements in $\{k_\alpha|_{|\alpha|\le M-2}\}$, and let $\boldsymbol{w}^{(k)}\in\mathbb{R}^{N_k}$ and

$$w^{(k)}_{\mathcal{N}(\alpha)}=k_\alpha,\quad |\alpha|\le M-2,$$

where $\boldsymbol{w}^{(g)}$ and $\boldsymbol{w}^{(k)}$ contain all the moments of the system. Let

$$\tilde{\boldsymbol{w}}=\begin{pmatrix}\boldsymbol{w}^{(g)}\\ \boldsymbol{w}^{(k)}\end{pmatrix}.$$

Collecting (3.28) and (3.31), we obtain the regularized moment system

$$\frac{\partial\tilde{\boldsymbol{w}}}{\partial t}+\sum_{j=1}^{2}\mathbf{A}_j\frac{\partial\tilde{\boldsymbol{w}}}{\partial x_1}=0. \tag{3.32}$$

It is clear that the matrix $\mathbf{A}_j$ has the form

$$\mathbf{A}_j=\begin{pmatrix}\mathbf{A}^{(g)}_j & 0\\ * & \mathbf{A}^{(k)}_j\end{pmatrix}, \tag{3.33}$$

where $\mathbf{A}^{(k)}_j\in\mathbb{R}^{N_k\times N_k}$ depends on (3.31), and satisfies

$$\mathbf{A}^{(k)}_j(\mathcal{N}(\alpha),\mathcal{N}(\alpha))=u_j, \tag{3.34a}$$
$$\mathbf{A}^{(k)}_j(\mathcal{N}(\alpha),\mathcal{N}(\alpha-e_j))=\theta,\qquad \text{if } \alpha_j>0, \tag{3.34b}$$
$$\mathbf{A}^{(k)}_j(\mathcal{N}(\alpha),\mathcal{N}(\alpha+e_j))=\alpha_j+1,\qquad \text{if } |\alpha|<M-2, \tag{3.34c}$$

and any entries of $\mathbf{A}^{(k)}_j$ not defined above are zero. We detach the matrix $\mathbf{A}^{(g)}_j$ and $\mathbf{A}^{(k)}_j$ from the moment system and study the case $M_k=M_g$. For the case $g=g_M$, we have $g_\alpha=0$, $|\alpha|>0$ and $\mathbf{A}^{(g)}_j=\mathbf{A}^{(k)}_j$, hence $\mathbf{A}^{(k)}_j$ is diagonalisable with real eigenvalues. We have the following theorem:

Theorem 3.1. *For any unit vector $\boldsymbol{n}=(n_1,n_2)^T\in\mathbb{R}^2$, the matrix $\sum_{j=1}^{2}n_j\mathbf{A}_j$ is diagonalisable. More precisely, the characteristic polynomial of $\sum_{j=1}^{2}n_j\mathbf{A}_j$ is*

$$\Big|\lambda\mathbf{I}-\sum_{j=1}^{2}n_j\mathbf{A}_j\Big|=\prod_{m=1}^{M_g+1}\theta^{m/2}He_m\Big(\frac{\lambda-\boldsymbol{u}\cdot\boldsymbol{n}}{\sqrt{\theta}}\Big)\cdot\prod_{m=1}^{M_k+1}\theta^{m/2}He_m\Big(\frac{\lambda-\boldsymbol{u}\cdot\boldsymbol{n}}{\sqrt{\theta}}\Big), \tag{3.35}$$

and its eigenvalues are

$$\boldsymbol{u}\cdot\boldsymbol{n}+\mathrm{C}_{i,m}\sqrt{\theta},\quad 1\le i\le m\le M+1, \tag{3.36}$$

where $\mathrm{C}_{i,m}$ *is the i-th root of* $He_m(z)$.

Proof. We first consider the case $\boldsymbol{n}=(1,0)$, and then apply rotational invariance to get the general result.

Lemma 3.1 shows $\mathbf{A}_1^{(g)}$ is diagonalisable, and the discussion above indicates that $\mathbf{A}_1^{(h)}$ is also diagonalisable. If $\boldsymbol{r}^{(g)}$ is a right eigenvector of $\mathbf{A}_1^{(g)}$, then $\boldsymbol{r}=(\boldsymbol{r}^{(g),T},0)^T$ is a right eigenvector of $\mathbf{A}_1$, and for $\mathbf{A}_1^{(k)}$ a similar result holds, hence the matrix $\mathbf{A}$ is diagonalisable.

Since $\mathbf{A}_1$ is a block lower triangular matrix, the determinant of $\lambda\mathbf{I}-\mathbf{A}_1$ is

$$\left|\lambda\mathbf{I}-\mathbf{A}_1\right|=\left|\lambda\mathbf{I}-\mathbf{A}_1^{(g)}\right|\cdot\left|\lambda\mathbf{I}-\mathbf{A}_1^{(k)}\right|.$$

Lemma 3.1 gives

$$\left|\lambda\mathbf{I}-\mathbf{A}_1^{(g)}\right|=\prod_{m=1}^{M_g+1}\theta^{m/2}He_m\Big(\frac{\lambda-u_1}{\sqrt{\theta}}\Big).$$

From the simple form of $\mathbf{A}_1^{(k)}$, it is not difficult to calculate its characteristic polynomial with the recursion relation of Hermite polynomial $He_{n+1}(x)=xHe_n(x)-nHe_{n-1}(x)$, to obtain

$$\left|\lambda\mathbf{I}-\mathbf{A}_1^{(k)}\right|=\prod_{m=1}^{M_k+1}\theta^{m/2}He_m\Big(\frac{\lambda-u_1}{\sqrt{\theta}}\Big);$$

so for the case $\boldsymbol{n}=(1,0)$ the theorem holds.

Previously, we have proved that all terms in (3.23) and (3.31) are rotational invariant, so the moment system (3.32) conserves rotational invariance-cf. [3] for details. □

Since $h_\alpha=k_\alpha+\theta g_\alpha/2$, on combining (3.26) and (3.31) we get the regularized equations of h_α:

$$\begin{aligned}
&\frac{\partial h_\alpha}{\partial t}+\sum_{j=1}^{2}\Big(\theta\frac{\partial h_{\alpha-e_j}}{\partial x_j}+u_j\frac{\partial h_\alpha}{\partial x_j}+(\alpha_j+1)\frac{\partial h_{\alpha+e_j}}{\partial x_j}\Big)+\sum_{j=1}^{2}\Big(-\frac{\theta}{2\rho}C_{\theta,\alpha}^{(j)}\Big)\frac{\partial\rho}{\partial x_j}\\
&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\frac{\partial u_d}{\partial x_j}\Big(\theta h_{\alpha-e_d-e_j}+(1-\delta_{\alpha,M})(\alpha_j+1)h_{\alpha-e_d+e_j}-\frac{C_\alpha}{2\rho}p_{jd}\Big)\\
&\quad+\sum_{j=1}^{2}\sum_{d=1}^{2}\Big(\Big(-\frac{h_{\alpha-e_d}}{\rho}\Big)\frac{\partial p_{jd}}{\partial x_j}+\frac{C_{\theta,\alpha}^{(j)}}{6\rho}\frac{\partial p_{dd}}{\partial x_j}\Big)+\Big(-\frac{C_\alpha}{3\rho}\Big)\sum_{j=1}^{2}\frac{\partial q_j}{\partial x_j}\\
&\quad-\delta_{\alpha,M}\sum_{j=1}^{2}(\alpha_j+1)\Big(\frac{\partial k_{\alpha+e_j}}{\partial x_j}+\sum_{d=1}^{2}k_{\alpha-e_d+e_j}\frac{\partial u_d}{\partial x_j}\Big)=0,
\end{aligned}\tag{3.37}$$

where C_α and $C^{(j)}_{\theta,\alpha}$ are redefined as

$$C_\alpha = \sum_{d=1}^{2} h_{\alpha-2e_k},$$
$$C^{(j)}_{\theta,\alpha} = \theta \sum_{k=1}^{2} h_{\alpha-2e_k-e_j} + (1-\delta_{\alpha,M})(\alpha_j+1) \sum_{k=1}^{2} h_{\alpha-2e_k+e_j}.$$

Let $\boldsymbol{w}^{(h)} \in \mathbb{R}^{N_k}$,

$$w^{(h)}_{\mathcal{N}(\alpha)} = h_\alpha, \quad |\alpha| \le M-2,$$

and

$$\boldsymbol{w} = \begin{pmatrix} \boldsymbol{w}^{(g)} \\ \boldsymbol{w}^{(h)} \end{pmatrix}. \tag{3.38}$$

Then the moment system (3.32) can be written as

$$\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{2} \mathbf{B}_j \frac{\partial \boldsymbol{w}}{\partial x_1} = 0, \tag{3.39}$$

and one can observe that it is equivalent to the moment system (3.32), hence Theorem 3.1 implies the moment system (3.39) is also hyperbolic. Consequently, we have obtained a regularized hyperbolic moment system for the dimension reduced Boltzmann equation.

On including the collision part in the moment system (3.39), we obtain

$$\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{2} \mathbf{B}_j \frac{\partial \boldsymbol{w}}{\partial x_j} = \frac{1}{\tau} \boldsymbol{Q}\boldsymbol{w}. \tag{3.40}$$

Here $\boldsymbol{Q}\boldsymbol{w}/\tau$ stands for the collision part, and $\boldsymbol{Q}$ is determined by (3.13).

4 Boundary condition

We use the same method as in [6] to study the boundary condition (2.10). Although only the distribution functions with the three-dimensional velocity are considered in [6], it is trivial to expand it to the reduced distribution functions g and h. In this subsection, we first list the result of the construction of the boundary conditions, and then prove that the hyperbolic moment equations are subject to the correct number of boundary conditions.

Without loss of generality, consider the case with the outer normal vector $\boldsymbol{n} = (1,0)$. (Any other case can be included by a rotation.) The boundary distributions g^b and h^b is as

$$\phi^b(t,\boldsymbol{x},\boldsymbol{\xi}) = \sum_{|\alpha| \le M_\phi} \phi^b_\alpha \mathcal{H}_{\theta^b,\alpha}\Big(\frac{\boldsymbol{\xi}-\boldsymbol{u}^b}{\sqrt{\theta^b}}\Big), \quad \phi = g,h, \tag{4.1}$$

where $\boldsymbol{u}^b=(u_1^W,u_2)$, $\theta^b=\theta^W$, and ϕ_α^b is defined by

$$\phi_\alpha^b=\begin{cases}\phi_\alpha, & \alpha_1 \text{ is even},\\ \dfrac{2\chi}{2-\chi}\Big[P_{\phi,\alpha}+\displaystyle\sum_{l=0}^{K(\alpha)}S(\alpha_1,2l)\theta^{\alpha_1/2-l}\phi_{2le_1+\alpha_2e_2}\Big], & \alpha_1 \text{ is odd},\end{cases} \tag{4.2}$$

where $K(\alpha)=\lceil(M-\alpha_1)/2\rceil$. $P_{g,\alpha}$, $P_{h,\alpha}$ and $S(\cdot,\cdot)$ are all defined in a recursive manner. The symbol $S(\cdot,\cdot)$ is defined as

$$S(r,s)=\begin{cases}1/2, & r=s=0,\\ K(1,s-1), & r=0 \text{ and } s\neq 0,\\ K(r,0), & r\neq 0 \text{ and } s=0,\\ K(r,s)+S(r-1,s-1)\cdot s/r, & \text{otherwise},\end{cases} \tag{4.3}$$

where

$$K(r,s)=\frac{\sqrt{2\pi}}{r!}He_{r-1}(0)He_s(0). \tag{4.4}$$

Clearly, the values of $S(\cdot,\cdot)$ are not relevant to the distribution function; so they can be calculated and stored beforehand. Now we define

$$P_{g,\alpha}=J_{\alpha_1}(u_1^W-u_1)\hat{J}_{\alpha_2}\sqrt{\frac{2\pi}{\theta^W}}\sum_{k=0}^{\lfloor M/2\rfloor}S(1,2k)\theta^{1/2-k}g_{2ke_1}, \tag{4.5}$$

where $J_n(\cdot)$ is also defined in a recursive manner:

$$J_0(x)=1, \qquad J_1(x)=x, \tag{4.6a}$$

$$J_n(x)=\frac{1}{n}\Big[(\theta^W-\theta)J_{n-2}(x)+xJ_{n-1}(x)\Big], \quad n\geq 2. \tag{4.6b}$$

For $\hat{J}_s$, the expression is a little more complex:

$$\hat{H}_0=0, \qquad \hat{H}_1=\sqrt{\frac{\theta^W}{2\pi}}, \qquad \hat{H}_n=-\frac{n-2}{n(n-1)}\theta\hat{H}_{n-2}, \qquad n\geq 2, \tag{4.7a}$$

$$\hat{J}_0=1/2, \qquad \hat{J}_1=-\hat{H}_1, \qquad \hat{J}_n=\frac{1}{n}(\theta^W-\theta)\hat{J}_{n-2}-\hat{H}_n, \qquad n\geq 2. \tag{4.7b}$$

For $P_{h,\alpha}$, the definition is

$$P_{h,\alpha}=\frac{\theta^W}{2}P_{g,\alpha}. \tag{4.8}$$

Particularly, $g_{e_1}^b$ can be written as

$$g_{e_1}^b=\frac{2\chi}{2-\chi}\sqrt{\frac{\pi}{2\theta}}(2\rho\theta-g_{2e_1})\Big((u_1^W-u_1)\sqrt{\frac{\pi}{2\theta^W}}+1\Big).$$

Using (3.5), we can obtain the macroscopic velocity on the surface $u_1^s = u_1^W + g_{e_1}^b/\rho$. The slip velocity denotes by

$$u_1^s - u_1^W = g_{e_1}^b/\rho = \frac{2\chi}{2-\chi}\sqrt{\frac{\pi}{2\theta}}\Big(2\theta - \frac{g_{2e_1}}{\rho}\Big)\Big((u_1^W - u_1)\sqrt{\frac{\pi}{2\theta^W}} + 1\Big). \tag{4.9}$$

For the temperature, (4.2) indicates $g_{2e_i}^b = g_{2e_i}$, $i = 1,2$ and $h_0^b = h_0$. Using (3.5), we can obtain the temperature on the surface

$$\theta^s = \frac{\theta + 2\theta^W - (u_1^s - u_1^W)^2}{3}.$$

The temperature-jump relation reads

$$\theta^s - \theta^W = \frac{\theta - \theta^W - (u_1^s - u_1^W)^2}{3}. \tag{4.10}$$

As discussed in Section 2, the boundary condition should be provided for $(\boldsymbol{\xi} - \boldsymbol{u}^W)\cdot\boldsymbol{n} < 0$, hence the number of boundary condition provided in (4.2) is the admissible α with α_1 odd. On the other hand, in Section 3.4, we prove that the moment system (3.32) is globally hyperbolic and the structure of the characteristic speed is fully clarified. Consequently, the number of characteristics waves traveling inward to the domain can be counted, and we have the following proposition.

Proposition 4.1. On the boundary point $\boldsymbol{x}$ with $\boldsymbol{n} = (1,0)$, the number of boundary conditions provided in (4.2) equals the number of characteristics waves traveling inward to the domain.

Proof. First, consider the number of characteristics waves traveling inward to the domain, i.e., the number of eigenvalues of $\mathbf{A}_1$ which are less than u_1. Theorem 3.1 shows that the characteristic polynomial of $\mathbf{A}_1$ is

$$|\lambda\mathbf{I} - \mathbf{A}_1| = \prod_{m=1}^{M_g+1} \theta^{m/2} He_m\Big(\frac{\lambda - u_1}{\sqrt{\theta}}\Big)\cdot\prod_{m=1}^{M_k+1} \theta^{m/2} He_m\Big(\frac{\lambda - u_1}{\sqrt{\theta}}\Big),$$

hence the number of the eigenvalues of $\mathbf{A}_1$ less than u_1 equals to the number of negative roots of

$$\prod_{m=1}^{M_g+1} He_m(x)\cdot\prod_{m=1}^{M_k+1} He_m(x). \tag{4.11}$$

It is well known that the roots of Hermite polynomials are symmetrically distributed with respect to the origin, so the number of negative roots of $He_n(x)$ is $\lfloor n/2\rfloor$, and hence the number of negative roots of (4.11) is

$$\sum_{m=1}^{M_g+1}\lfloor m/2\rfloor + \sum_{m=1}^{M_k+1}\lfloor m/2\rfloor. \tag{4.12}$$

Let us count the number of boundary conditions provided in (4.2), which is the number of admissible α with α_1 odd. Let

$$\mathcal{S}_m=\{\alpha\in\mathbb{N}^2||\alpha|\le m, \text{ and } \alpha_1 \text{ is odd}\},\quad m\in\mathbb{N},$$

and then the number of boundary conditions provided in (4.2) is $\#\mathcal{S}_{M_g}+\#\mathcal{S}_{M_k}$. Since

$$\begin{aligned}\mathcal{S}_m-\mathcal{S}_{m-1}&=\{\alpha\in\mathbb{N}^2||\alpha|=m, \text{ and } \alpha_1 \text{ is odd}\}\\&=\{\alpha\in\mathbb{N}^2|\alpha_1+\alpha_2=m, \text{ and } \alpha_1 \text{ is odd}\},\end{aligned}$$

we have $\#(\mathcal{S}_m-\mathcal{S}_{m-1})=\lfloor m/2\rfloor$. Noting $\#\mathcal{S}_0=0$, we have

$$\#\mathcal{S}_{M_g}+\#\mathcal{S}_{M_k}=\sum_{m=1}^{M_g+1}\lfloor m/2\rfloor+\sum_{m=1}^{M_k+1}\lfloor m/2\rfloor. \tag{4.13}$$

Comparing (4.12) to (4.13), we have proved the proposition. □

For hyperbolic equations, to keep well-posedness of the equations, the number of boundary conditions must equal to the number of characteristics waves traveling inward to the domain. Hence, the proposition indicates that (4.2) provides the DRHME with correct number of boundary conditions.

5 Numerical scheme

In this section, we study the numerical scheme to solve the dimension-reduced hyperbolic moment system. The framework of the scheme is similar to that of the NRxx method introduced in [8].

5.1 Framework of numerical scheme

Let $\mathcal{T}_h$ be a rectangular mesh in $\mathbb{R}^2$, with all grid lines parallel with the axes, and identify each cell by two indices i,j, i.e., for a fixed $\boldsymbol{x}_0\in\mathbb{R}^2$ and $\Delta x_d>0$, $d=1,2$,

$$\mathcal{T}_h=\{T_{i,j}=\boldsymbol{x}_0+[i\Delta x_1,(i+1)\Delta x_1]\times[j\Delta x_2,(j+1)\Delta x_2]|i,j\in\mathbb{N}\}.$$

Let $\phi^n_{i,j}(\boldsymbol{\xi})$, $\phi=g,h,k$ approximate the distribution functions in the cell $T_{i,j}$ at time t_n.

To solve the regularized moment system, we split (3.40) into two steps:

$$\text{Convection part:}\qquad \frac{\partial \boldsymbol{w}}{\partial t}+\sum_{j=1}^{2}\mathbf{B}_j\frac{\partial \boldsymbol{w}}{\partial x_j}=0, \tag{5.1a}$$

$$\text{Collision part:}\qquad \frac{\partial \boldsymbol{w}}{\partial t}=\frac{1}{\tau}\boldsymbol{Q}\boldsymbol{w}. \tag{5.1b}$$

Then the numerical procedure is as follows.

1. Let $n=0$ and $t=0$, and give the initial value of $\boldsymbol{w}_{i,j}^n$ for all i, j.
2. Calculate the time step length Δt_n by CFL condition.
3. Solve the convection part by one time step using the finite volume method, and denote the result by $\boldsymbol{w}_{i,j}^{n,*}$.
4. Use $\boldsymbol{w}_{i,j}^{n,*}$ as the initial values, solve the collision part by one step, and denote the result by $\boldsymbol{w}_{i,j}^{n+1}$.
5. Let $t\leftarrow t+\Delta t_n$ and $n\leftarrow n+1$; then go to Step 2.

5.2 Convection part

Since the hyperbolic regularization in Section 3.4 modifies some terms in the equations involving the time derivative of $g_\alpha|_{|\alpha|=M}$ and $h_\alpha|_{|\alpha|=M-2}$, these equations may not be written into a conservation form. In [7], an effective numerical scheme based on the DLM theory [23] is proposed, which we adopt for the DRHME.

The moment equations (3.26) and (3.31) can be reformulated as [3,4,7]

$$\frac{\partial \boldsymbol{U}}{\partial t}+\sum_{d=1}^{2}\frac{\partial \boldsymbol{F}_d(\boldsymbol{U})}{\partial x_d}+\sum_{d=1}^{2}\boldsymbol{R}_d(\boldsymbol{U})\frac{\partial \boldsymbol{U}}{\partial x_d}=0, \tag{5.2}$$

where $\partial_t\boldsymbol{U}+\sum_{d=1}^{2}\partial_{x_d}\boldsymbol{F}_d(\boldsymbol{U})=0$ is a Grad type moment system, and $\sum_{d=1}^{2}\boldsymbol{R}_d(\boldsymbol{U})\partial_{x_d}\boldsymbol{U}$ is the regularization part proposed in Section 3.4, which contains only a few terms. The details of $\boldsymbol{U}$ and $\boldsymbol{F}_d(\boldsymbol{U})$ with $d=1,2$ will be given in the next subsection. It is clear that an invertible transformation between $\boldsymbol{w}$ and $\boldsymbol{U}$ are in need in solving the convection part. Here we give the numerical scheme for (5.2), and leave the transformations to the next subsection.

Following [7], we propose the numerical scheme

$$\begin{aligned}\boldsymbol{U}_{i,j}^{n,*}=\boldsymbol{U}_{i,j}^{n}&-\frac{\Delta t_n}{\Delta x_1}(\hat{\boldsymbol{F}}_{i+1/2,j}^{n}-\hat{\boldsymbol{F}}_{i-1/2,j}^{n})-\frac{\Delta t_n}{\Delta x_1}(\hat{\boldsymbol{R}}_{i+1/2,j}^{n-}-\hat{\boldsymbol{R}}_{i-1/2,j}^{n+})\\&-\frac{\Delta t_n}{\Delta x_2}(\hat{\boldsymbol{F}}_{i,j+1/2}^{n}-\hat{\boldsymbol{F}}_{i,j-1/2}^{n})-\frac{\Delta t_n}{\Delta x_2}(\hat{\boldsymbol{R}}_{i,j+1/2}^{n-}-\hat{\boldsymbol{R}}_{i,j-1/2}^{n+}),\end{aligned} \tag{5.3}$$

where $\hat{\boldsymbol{F}}_{i+1/2,j}^{n}$ is the HLL numerical flux defined by

$$\hat{\boldsymbol{F}}_{i+1/2,j}^{n}=\begin{cases}\boldsymbol{F}_1(\boldsymbol{U}_{i,j}^{n}), & \lambda_{i+1/2,j}^{L}\geq 0,\\ \dfrac{\lambda_{i+1/2,j}^{R}\boldsymbol{F}_1(\boldsymbol{U}_{i,j}^{n})-\lambda_{i+1/2,j}^{L}\boldsymbol{F}_1(\boldsymbol{U}_{i+1,j}^{n})}{\lambda_{i+1/2,j}^{R}-\lambda_{i+1/2,j}^{L}}\\ \quad+\dfrac{\lambda_{i+1/2,j}^{L}\lambda_{i+1/2,j}^{R}(\boldsymbol{U}_{i+1,j}^{n}-\boldsymbol{U}_{i,j}^{n})}{\lambda_{i+1/2,j}^{R}-\lambda_{i+1/2,j}^{L}}, & \lambda_{i+1/2,j}^{L}<0<\lambda_{i+1/2,j}^{R},\\ \boldsymbol{F}_1(\boldsymbol{U}_{i+1,j}^{n}), & \lambda_{i+1/2,j}^{R}\leq 0,\end{cases} \tag{5.4}$$

and $\hat{\boldsymbol{F}}^n_{i,j+1/2}$ is defined in the same form. The numerical flux for the nonconservative part $\hat{\boldsymbol{R}}^{n\pm}_{i+1/2,j}$ is defined by

$$\hat{\boldsymbol{R}}^{n-}_{i+1/2,j}=\begin{cases}0, & \lambda^L_{i+1/2,j}\geq 0,\\ -\dfrac{\lambda^L_{i+1/2,j}}{\lambda^R_{i+1/2,j}-\lambda^L_{i+1/2,j}}\boldsymbol{g}^n_{i+1/2,j}, & \lambda^L_{i+1/2,j}<0<\lambda^R_{i+1/2,j},\\ \boldsymbol{g}^n_{i+1/2,j}, & \lambda^R_{i+1/2,j}\leq 0,\end{cases} \tag{5.5}$$

and

$$\hat{\boldsymbol{R}}^{n+}_{i+1/2,j}=\begin{cases}-\boldsymbol{g}^n_{i+1/2,j}, & \lambda^L_{i+1/2,j}\geq 0,\\ -\dfrac{\lambda^R_{i+1/2,j}}{\lambda^R_{i+1/2,j}-\lambda^L_{i+1/2,j}}\boldsymbol{g}^n_{i+1/2,j}, & \lambda^L_{i+1/2,j}<0<\lambda^R_{i+1/2,j},\\ 0, & \lambda^R_{i+1/2,j}\leq 0,\end{cases} \tag{5.6}$$

where

$$\boldsymbol{g}^n_{i+1/2,j}=\int_0^1 \boldsymbol{R}(\boldsymbol{\Phi}(s;\boldsymbol{q}^n_{i,j},\boldsymbol{q}^n_{i+1,j}))\frac{\partial\boldsymbol{\Phi}}{\partial s}(s;\boldsymbol{q}^n_{i,j},\boldsymbol{q}^n_{i+1,j})\mathrm{d}s, \tag{5.7}$$

$\boldsymbol{\Phi}(s;\cdot,\cdot)$ is a path to connect the two states (see [23] for details). It was pointed out in [3] that the integral path $\boldsymbol{\Phi}(s;\cdot,\cdot)$ seems not essential in our problem. The numerical flux $\hat{\boldsymbol{R}}^{n\pm}_{i,j+1/2}$ is defined in the same form as $\hat{\boldsymbol{R}}^{n\pm}_{i+1/2,j}$. $\lambda^L_{i+1/2,j}$ ($\lambda^R_{i+1/2,j}$) is the maximal (minimal) characteristic speed on both sides. Theorem 3.1 tells us that

$$\begin{aligned}\lambda^R_{i+1/2,j}&=\max\left((u_1)_{i,j}+\mathrm{C}_{M+1}\sqrt{\theta_{i,j}},(u_1)_{i+1,j}+\mathrm{C}_{M+1}\sqrt{\theta_{i+1,j}}\right),\\ \lambda^L_{i+1/2,j}&=\min\left((u_1)_{i,j}-\mathrm{C}_{M+1}\sqrt{\theta_{i,j}},(u_1)_{i+1,j}-\mathrm{C}_{M+1}\sqrt{\theta_{i+1,j}}\right),\end{aligned}$$

where C_{M+1} is the maximal roots of the $M+1$ degree Hermite polynomial $He_{M+1}(x)$. Analogously, we have

$$\begin{aligned}\lambda^R_{i,j+1/2}&=\max\left((u_2)_{i,j}+\mathrm{C}_{M+1}\sqrt{\theta_{i,j}},(u_2)_{i,j+1}+\mathrm{C}_{M+1}\sqrt{\theta_{i,j+1}}\right),\\ \lambda^L_{i,j+1/2}&=\min\left((u_2)_{i,j}-\mathrm{C}_{M+1}\sqrt{\theta_{i,j}},(u_2)_{i,j+1}-\mathrm{C}_{M+1}\sqrt{\theta_{i,j+1}}\right).\end{aligned}$$

The time step Δt is given by the CFL condition:

$$\max_{i,j}\left(\frac{\Delta t}{\Delta x_1}\left(|(u_1)_{i,j}|+\mathrm{C}_{M+1}\sqrt{\theta_{i,j}}\right)+\frac{\Delta t}{\Delta x_2}\left(|(u_2)_{i,j}|+\mathrm{C}_{M+1}\sqrt{\theta_{i,j}}\right)\right)<1. \tag{5.8}$$

5.3 Projection

In this discretisation, the variables $\boldsymbol{U}$ and $\boldsymbol{F}_d(\boldsymbol{U})$ with $d=1,2$ are constructed. In [5], a viable method was proposed to transform between $\boldsymbol{w}$ and $\boldsymbol{U}$. In this subsection, we first give the definition of $\boldsymbol{U}$ and $\boldsymbol{F}_d(\boldsymbol{U})$, and then introduce an even conciser and cheaper projection algorithm to implement the transformation between $\boldsymbol{w}$ and $\boldsymbol{U}$.

Let $\boldsymbol{U}^{(g)}\in\mathbb{R}^{N_g}$, and let $\boldsymbol{U}^{(h)}\in\mathbb{R}^{N_k}$ and denote $U_\alpha^{(g)}$ and $U_\alpha^{(h)}$ the $\mathcal{N}(\alpha)$-th entry of $\boldsymbol{U}^{(g)}$ and $\boldsymbol{U}^{(h)}$, respectively. When we update the solution $\boldsymbol{w}_{i,j}^n$ on the (i,j)-th grid for the next time step, $\boldsymbol{U}^{(g)}$ and $\boldsymbol{U}^{(h)}$ are chosen as

$$U_\alpha^{(g)}=\sum_{|\beta|\le M} g_\beta H_{i,j}^n(\alpha,\beta),\quad U_\alpha^{(h)}=\sum_{|\beta|\le M-2} h_\beta H_{i,j}^n(\alpha,\beta), \tag{5.9}$$

where

$$H_{i,j}^n(\alpha,\beta)=\frac{2\pi\cdot(\theta_{i,j}^n)^{1+|\alpha|}}{\alpha_1!\alpha_2!}\int_{\mathbb{R}^2}\mathcal{H}_{\theta,\beta}\Big(\frac{\boldsymbol{\xi}-\boldsymbol{u}}{\sqrt{\theta}}\Big)\mathcal{H}_{\theta_{i,j}^n,\alpha}\Big(\frac{\boldsymbol{\xi}-\boldsymbol{u}_{i,j}^n}{\sqrt{\theta_{i,j}^n}}\Big)\exp\Big(\frac{|\boldsymbol{\xi}-\boldsymbol{u}_{i,j}^n|^2}{2\theta_{i,j}^n}\Big)\mathrm{d}\boldsymbol{\xi}.$$

The vector $\boldsymbol{U}$ in (5.2) is

$$\boldsymbol{U}=\begin{pmatrix}\boldsymbol{U}^{(g)}\\ \boldsymbol{U}^{(h)}\end{pmatrix}.$$

Clearly, the value of $\boldsymbol{U}$ depends on $\boldsymbol{u}_{i,j}^n$ and $\theta_{i,j}^n$, so the transformation between $\boldsymbol{w}$ and $\boldsymbol{U}$ is dependent on $\boldsymbol{u}_{i,j}^n$ and $\theta_{i,j}^n$, too. For $d=1,2$, define

$$F_{d,\alpha}^{(g)}=\sum_{|\beta|\le M}\Big[\theta g_{\beta-e_d}+u_d g_\beta+(\alpha_d+1)g_{\beta+e_d}\Big]H_{i,j}^n(\alpha,\beta),\qquad |\alpha|\le M, \tag{5.10a}$$

$$F_{d,\alpha}^{(h)}=\sum_{|\beta|\le M-2}\Big[\theta h_{\beta-e_d}+u_d h_\beta+(\alpha_d+1)h_{\beta+e_d}\Big]H_{i,j}^n(\alpha,\beta),\qquad |\alpha|\le M-2, \tag{5.10b}$$

then $\boldsymbol{F}_d(\boldsymbol{U})$ equals

$$\boldsymbol{F}_d(\boldsymbol{U})=\begin{pmatrix}\boldsymbol{F}_d^{(g)}\\ \boldsymbol{F}_d^{(h)}\end{pmatrix}.$$

Let

$$\mathcal{R}_\alpha^{(g)}=\sum_{l=1}^2(\alpha_l+1)\Big[\sum_{d=1}^2 g_{\alpha-e_d+e_l}\frac{\partial u_d}{\partial x_l}+\frac{1}{2}\Big(\sum_{d=1}^2 g_{\alpha-2e_d+e_l}\Big)\frac{\partial\theta}{\partial x_l}\Big],\qquad |\alpha|=M,$$

$$\mathcal{R}_\alpha^{(h)}=\sum_{l=1}^2(\alpha_l+1)\Big[\sum_{d=1}^2 k_{\alpha-e_d+e_l}\frac{\partial u_d}{\partial x_l}+\frac{1}{2}\Big(g_{\alpha+e_l}+\sum_{d=1}^2 k_{\alpha-2e_d+e_l}\Big)\frac{\partial\theta}{\partial x_l}\Big],\qquad |\alpha|=M-2,$$

and then (5.1a) can be written as

$$\frac{\partial U_\alpha^{(\phi)}}{\partial t}+\sum_{d=1}^{2}\frac{\partial F_{d,\alpha}^{(\phi)}}{\partial x_d}-(1-\delta_{|\alpha|,M_\phi})\mathcal{R}_\alpha^{(\phi)}=0,\quad |\alpha|\le M_\phi \text{ with } \phi=g,h. \tag{5.11}$$

As in [7], one may verify directly that (5.1a) and (5.11) are equivalent for any $\boldsymbol{u}_{i,j}^n$ and $\theta_{i,j}^n$.

Until now, the finite volume scheme in the last subsection is totally complete. We remark that $\boldsymbol{U}$ in Eq. (5.11) is only applied on the (i,j)-th grid cell and on the (i',j')-th grid cell, $\boldsymbol{U}$ dependent on $\boldsymbol{u}_{i',j'}^n$ and $\theta_{i',j'}^n$ is applied.

The transformation as (5.9) may be costly, and an efficient one has been proposed in [5]. Here we introduce a more concise and easier algorithm to implement the transformation. For convenience, we use ϕ to replace g and h in the following discussion, and the result is valid for both variables.

Define two reduced distribution functions

$$\phi(\boldsymbol{\xi})=\sum_{|\alpha|\le M_\phi}\phi_\alpha\mathcal{H}_{\theta,\alpha}\Big(\frac{\boldsymbol{\xi}-\boldsymbol{u}}{\sqrt{\theta}}\Big),\quad \phi_{i,j}^n(\boldsymbol{\xi})=\sum_{|\alpha|\le M_\phi}U_\alpha^{(\phi)}\mathcal{H}_{\theta_{i,j}^n,\alpha}\Big(\frac{\boldsymbol{\xi}-\boldsymbol{u}_{i,j}^n}{\sqrt{\theta_{i,j}^n}}\Big). \tag{5.12}$$

By the definition of $U_\alpha^{(\phi)}$, it is clear that for any polynomial $p(\boldsymbol{\xi})$ with its degree not more than M_ϕ, the following equality holds:

$$\int_{\mathbb{R}^2}p(\boldsymbol{\xi})\phi(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}=\int_{\mathbb{R}^2}p(\boldsymbol{\xi})\phi_{i,j}^n(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}. \tag{5.13}$$

Actually, it is given by simply expanding $p(\boldsymbol{\xi})$ at the basis functions $\mathcal{H}_{\theta_{i,j}^n,\alpha}\exp\big(\frac{(\boldsymbol{\xi}-\boldsymbol{u}_{i,j}^n)^2}{2\theta_{i,j}^n}\big)$.

Then we have the following theorem.

Theorem 5.1. *For any $\boldsymbol{u}_{i,j}^n$ and $\theta_{i,j}^n$, if $\{X_\alpha^{(\phi)}\}_{|\alpha|\le M_\phi}$ satisfies*

$$\begin{cases}\dfrac{\mathrm{d}X_\alpha^{(\phi)}}{\mathrm{d}\tau}=\displaystyle\sum_{d=1}^{2}\Big[(u_d-(u_d)_{i,j}^n)X_{\alpha-e_d}^{(\phi)}+\frac{1}{2}(\theta-\theta_{i,j}^n)X_{\alpha-2e_d}^{(\phi)}\Big],\\ X_\alpha^{(\phi)}(0)=\phi_\alpha,\quad |\alpha|\le M_\phi,\quad \tau\in[0,1],\end{cases} \tag{5.14}$$

then $U_\alpha^{(\phi)}=X_\alpha^{(\phi)}(1)$ holds for any $|\alpha|\le M_\phi$. On the other hand, if $\{X_\alpha^{(\phi)}\}_{|\alpha|\le M_\phi}$ satisfies

$$\begin{cases}\dfrac{\mathrm{d}X_\alpha^{(\phi)}}{\mathrm{d}\tau}=\displaystyle\sum_{d=1}^{2}\Big[(u_d-(u_d)_{i,j}^n)X_{\alpha-e_d}^{(\phi)}+\frac{1}{2}(\theta-\theta_{i,j}^n)X_{\alpha-2e_d}^{(\phi)}\Big],\\ X_\alpha^{(\phi)}(1)=U_\alpha^{(\phi)},\quad |\alpha|\le M_\phi,\quad \tau\in[0,1],\end{cases} \tag{5.15}$$

then $\phi_\alpha=X_\alpha^{(\phi)}(0)$ holds for any $|\alpha|\le M_\phi$.

Before the proof of the theorem, we define the polynomials

$$\overline{He}_n(x)=\exp\Big(-\frac{x^2}{2}\Big)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\exp\Big(\frac{x^2}{2}\Big),$$

which satisfy the following properties:

- recurrence relations:

$$\begin{aligned}&\overline{He}_0(x)=1,\qquad \overline{He}_1(x)=x,\\&\overline{He}_{n+1}(x)=x\overline{He}_n(x)+n\overline{He}_{n-1}(x),\quad n\geq 1;\end{aligned}$$

- differential relations:

$$\frac{\mathrm{d}\overline{He}_n(x)}{\mathrm{d}x}=n\overline{He}_{n-1}(x);$$

- conjugate relations:

$$\sum_{m=0}^{n}\frac{\overline{He}_m(x)}{m!}\frac{He_{n-m}(y)}{(n-m)!}=\frac{(x+y)^n}{n!}.$$

For 2-dimensional case, let $\overline{He}_\alpha(\boldsymbol{x})=\prod_{d=1}^{2}\overline{He}_{\alpha_d}(x_d)$. Define an auxiliary function

$$F_\alpha(\tau)=\sum_{\alpha-\beta\geq 0}X_\beta(\tau)\frac{\overline{He}_{\alpha-\beta}\big(\frac{\boldsymbol{u}(\tau)}{\sqrt{\theta(\tau)}}\big)}{(\alpha-\beta)!}\theta(\tau)^{|\alpha-\beta|/2},\quad \tau\in[0,1],\tag{5.16}$$

where $\alpha-\beta\geq 0$ stands for $(\alpha-\beta)_d\geq 0$, $d=1,2$, $X_\beta(\tau)$ is same as that in Theorem 5.1, and $\boldsymbol{u}(\tau)$ and $\theta(\tau)$ is defined as

$$\boldsymbol{u}(\tau)=\boldsymbol{u}+\tau(\boldsymbol{u}_{i,j}^n-\boldsymbol{u}),\quad \theta(\tau)=\theta+\tau(\theta_{i,j}^n-\theta).$$

Then we have the following result:

Lemma 5.1. *If the first equation of* (5.14) *is valid, then*

$$\frac{\mathrm{d}F_\alpha}{\mathrm{d}\tau}=0,\quad \tau\in[0,1]\tag{5.17}$$

holds for all $|\alpha|\leq M_\phi$. *Particularly, we have* $F_\alpha(0)=F_\alpha(1)$, $|\alpha|\leq M_\phi$.

Proof. For convenience, let $\boldsymbol{c}=\boldsymbol{u}(\tau)/\sqrt{\theta(\tau)}$. Direct calculation yields

$$\begin{aligned}&\frac{\mathrm{d}\overline{He}_{\alpha-\beta}(\boldsymbol{c})\theta(\tau)^{|\alpha-\beta|/2}}{\mathrm{d}\tau}\\&=\sum_{d=1}^{D}(\alpha_d-\beta_d)\left[\overline{He}_{\alpha-\beta-e_d}\frac{1}{\sqrt{\theta(\tau)}}\frac{\mathrm{d}u_d(\tau)}{\mathrm{d}\tau}+\frac{\alpha_d-\beta_d-1}{2\theta(\tau)}\frac{\mathrm{d}\theta(\tau)}{\mathrm{d}\tau}\overline{He}_{\alpha-\beta-2e_d}\right].\end{aligned}\tag{5.18}$$

Thus,

$$\begin{aligned}\frac{\mathrm{d}F_\alpha(\tau)}{\mathrm{d}\tau}=&\sum_{\alpha-\beta\geq 0}\frac{1}{(\alpha-\beta)!}\Bigg[\frac{\mathrm{d}X_\beta(\tau)}{\mathrm{d}\tau}\overline{He}_{\alpha-\beta}(\boldsymbol{c})\theta(\tau)^{|\alpha-\beta|/2}\\&+X_\alpha(\tau)\sum_{d=1}^{D}(\alpha_d-\beta_d)\left(\overline{He}_{\alpha-\beta-e_d}\frac{1}{\sqrt{\theta(\tau)}}\frac{\mathrm{d}u_d(\tau)}{\mathrm{d}\tau}+\frac{\alpha_d-\beta_d-1}{2\theta(\tau)}\frac{\mathrm{d}\theta(\tau)}{\mathrm{d}\tau}\overline{He}_{\alpha-\beta-2e_d}\right)\Bigg]\\=&\sum_{\alpha-\beta\geq 0}\frac{\theta(\tau)^{|\alpha-\beta|/2}}{(\alpha-\beta)!}\overline{He}_{\alpha-\beta}(\boldsymbol{c})\left(\frac{\mathrm{d}X_\beta(\tau)}{\mathrm{d}\tau}+\sum_{d=1}^{D}\left(\frac{\mathrm{d}u_d(\tau)}{\mathrm{d}\tau}X_{\alpha-e_d}+\frac{1}{2}\frac{\mathrm{d}\theta(\tau)}{\mathrm{d}\tau}X_{\alpha-2e_d}\right)\right).\end{aligned}$$

Since the first equation of (5.14) is valid, we have $F_\alpha(\tau)/\tau=0$. This completes the proof. □

Proof of Theorem 5.1. We only prove the first part of the theorem, and the second part can be proved analogously. Substituting the definition of ϕ_α into (5.16), we obtain

$$\begin{aligned}F_\alpha(0)=&\sum_{\alpha-\beta\geq 0}\phi_\beta\frac{\overline{He}_{\alpha-\beta}\left(\frac{\boldsymbol{u}}{\sqrt{\theta}}\right)}{(\alpha-\beta)!}\theta^{|\alpha-\beta|/2}\\=&\int_{\mathbb{R}^2}\phi\theta^{(|\alpha|+D)/2}\sum_{\alpha-\beta\geq 0}\frac{\overline{He}_{\alpha-\beta}(\boldsymbol{c})}{(\alpha-\beta)!}\frac{He_\beta(\boldsymbol{v})}{\beta!}\mathrm{d}\boldsymbol{v}\\=&\int_{\mathbb{R}^2}\phi\boldsymbol{\xi}^\alpha\mathrm{d}\boldsymbol{\xi}.\end{aligned}$$

Similarly, if $U_\alpha^{(\phi)}=X_\alpha^{(\phi)}(1)$ holds, then $F_\alpha(1)=\int_{\mathbb{R}^2}\phi_{i,j}^n\boldsymbol{\xi}^\alpha\mathrm{d}\boldsymbol{\xi}$. Actually, Lemma 5.1 shows that $F_\alpha(1)=\int_{\mathbb{R}^2}\phi\boldsymbol{\xi}^\alpha\mathrm{d}\boldsymbol{\xi}$. Considering (5.13), we have $F_\alpha(1)=\int_{\mathbb{R}^2}\phi_{i,j}^n\boldsymbol{\xi}^\alpha\mathrm{d}\boldsymbol{\xi}$. Hence, $U_\alpha^{(\phi)}=X_\alpha^{(\phi)}(1)$ is one solution of (5.14).

On the other hand, the ordinary differential equations (ODEs) (5.14) is a constant coefficient system, thus $U_\alpha^{(\phi)}=X_\alpha^{(\phi)}(1)$ is the unique solution of (5.14). □

Similarly as in the definition of $\boldsymbol{U}$, let $X^{(\phi)}$ be the $\mathcal{N}(\alpha)$-th entry of $\boldsymbol{X}^{(\phi)}$ and $\boldsymbol{X}=(\boldsymbol{X}^{(g),T},\boldsymbol{X}^{(h),T})^T$. Then the ODEs (5.14) and (5.15) can be written as

$$\frac{\mathrm{d}\boldsymbol{X}(\tau)}{\mathrm{d}\tau}=\boldsymbol{H}\boldsymbol{X}(\tau),\quad \tau\in[0,1]. \tag{5.19}$$

With (5.14) and (5.15), it is clear that $\boldsymbol{H}$ is a constant coefficient sparse nilpotent matrix. So we need only to solve an ODE system with a constant coefficient matrix to implement the projection.

Let us now discuss the technical details in solving the ODE system (5.14) and (5.15). By setting $\alpha=0$, we have

$$\frac{\mathrm{d}X_0^{(\phi)}}{\mathrm{d}\tau}=0,$$

so $X_0^{(\phi)}$ is a constant. Using mathematical induction, immediately we can prove that $X_\alpha^{(\phi)}$ is an $|\alpha|$-th degree polynomial of τ (cf. Appendix for the proof), so the ODE system (5.14) and (5.15) can be solved analytically. For any $|\alpha|=m$, since $X_\alpha^{(\phi)}$ is an m-th degree polynomial of τ, the computational cost in solving (5.19) is $\mathcal{O}(m)$. Noting that there are $m+1$ different α satisfying $|\alpha|=m$, $\mathcal{O}(m^2)$ operations are needed to calculate all of the $X_\alpha^{(\phi)}$ with $|\alpha|=m$. Hence, the transformation from $\boldsymbol{w}$ to $\boldsymbol{U}$ costs $\mathcal{O}(M_g^3)+\mathcal{O}(M_k^3)=\mathcal{O}(M^3)$ operations.

In practice, an accurate computation of (5.14) and (5.15) is not always necessary. Recall that the transformation only takes place when updating the solution on the (i,j)-th grid, and the scheme (5.3) shows that only solutions on the neighboring cells of $T_{i,j}$ are involved. Thus $\boldsymbol{u}$ and θ in (5.14) are only taken as $\boldsymbol{u}_{i\pm1,j\pm1}^n$ and $\theta_{i\pm1,j\pm1}^n$ in the computation. Since the solution is varying mildly in most of the computational domain, this results in $u_d-(u_d)_{i,j}^n=\mathcal{O}(\Delta x)$, $d=1,2$ and $\theta-\theta_{i,j}^n=\mathcal{O}(\Delta x)$, where $\Delta x=\max(\Delta x_1,\Delta x_2)$. Thus we can use an efficient numerical integrator, such as a Runge-Kutta type scheme, to solve (5.19). And it is accurate enough after only a few steps. In this way, the cost of the transformation is decreased to $\mathcal{O}(M^2)$. In particular, from (3.5), the mass, momentum and energy remain unchanged when the algebraic accuracy order of the numerical integrator is greater than 2.

5.4 Collision part

In Section 3.2, we showed that the collision part does not change the macroscopic density ρ, velocity $\boldsymbol{u}$ and temperature θ of gas, hence the collision part can be expanded accurately. Here we follow [5] and [6], to give the expression of collision part in the moment expansion. For neatness, the superscripts n, $n+1$ and the subscripts i, j are omitted, for example, ϕ_α^* and ϕ_α with $\phi=g,h$ are used to represent $(\phi_\alpha)_{i,j}^{n,*}$ and $(\phi_\alpha)_{i,j}^{n+1}$, respectively.

- For the BGK model:

$$
\begin{aligned}
&g_\alpha=g_\alpha^*\exp\Big(-\frac{\Delta t}{\tau}\Big),\quad 2\le|\alpha|\le M,\\
&h_0=\frac{1}{2}g_0^*\theta^*\left(1-\exp\Big(\frac{\Delta t}{\tau}\Big)\right)+h_0^*\exp\Big(-\frac{\Delta t}{\tau}\Big),\\
&h_\alpha=h_\alpha^*\exp\Big(-\frac{\Delta t}{\tau}\Big),\quad 1\le|\alpha|\le M-2.
\end{aligned}
$$

- For the Shakhov model (for $i,j=1,2$):

$$
\begin{aligned}
&g_{e_i+2e_j}=\frac{1}{5}g_{e_i+2e_j}^*\exp\Big(-\frac{\Delta t}{\tau}\Big)+\frac{1}{5}q_i^*\left(\exp\Big(-\frac{\Pr\Delta t}{\tau}\Big)-\exp\Big(-\frac{\Delta t}{\tau}\Big)\right),\\
&g_\alpha=g_\alpha^*\exp\Big(-\frac{\Delta t}{\tau}\Big),\quad |\alpha|=2 \text{ or } 4\le|\alpha|\le M,
\end{aligned}
$$

$$h_0 = h_0^* \exp\left(-\frac{\Delta t}{\tau}\right) + \frac{1}{2} g_0^* \theta^* \left(1 - \exp\left(-\frac{\Delta t}{\tau}\right)\right),$$
$$h_{e_i} = \frac{1}{5} h_{e_i}^* \exp\left(-\frac{\Delta t}{\tau}\right) + \frac{1}{5} q_i^* \left(\exp\left(-\frac{\Pr \Delta t}{\tau}\right) - \exp\left(-\frac{\Delta t}{\tau}\right)\right),$$
$$h_{e_i+2e_j} = h_{e_i+2e_j}^* \exp\left(-\frac{\Delta t}{\tau}\right) + \frac{\theta^* q_i^*}{5} \left(\exp\left(-\frac{\Pr \Delta t}{\tau}\right) - \exp\left(-\frac{\Delta t}{\tau}\right)\right),$$
$$h_\alpha = h_\alpha^* \exp\left(-\frac{\Delta t}{\tau}\right), \quad |\alpha| = 2 \text{ or } 4 \le |\alpha| \le M-2.$$

5.5 Numerical boundary treatment

In Section 4, the kinetic boundary condition for the moment system was considered. In practical computation, one may use the ghost cell technique to construct the boundary condition [6]. We take the right boundary, where the unit outer normal is $\boldsymbol{n} = (1,0)$ as an example, and let N denote the number of grids on the x_1 direction. For the (N,j)-th cell on the right boundary, we construct the distribution functions $g_{N+1,j}^n$ and $h_{N+1,j}^n$ in the ghost cell as

$$(u_1)_{N+1,j}^n = 2u_1^b - (u_1)_{N,j}^n, \qquad (u_2)_{N+1,j}^n = (u_2)_{N,j}^n, \qquad \theta_{N+1,j}^n = \theta_{N,j}^n,$$
$$(g_\alpha)_{N+1,j}^n = 2g_\alpha^b - (g_\alpha)_{N,j}^n, \qquad (h_\alpha)_{N+1,j}^n = 2h_\alpha^b - (h_\alpha)_{N,j}^n.$$

For other boundaries, the distribution functions in the ghost cells can be constructed in the same way. As remarked in [6], we point out that the ghost cell technique and the numerical flux, proposed in Section 5.2, preserve conservation of mass.

6 Numerical examples

Two numerical examples are presented in this section. The first is a one-dimensional example, used to demonstrate the validity of the DRHME. In this example, for different Knudsen numbers we verify the convergence of DRHME as M increases, and compare the CPU time of the DRHME with the case without dimension-reduction. The second example is a two-dimensional cavity flow simulation, to show the validity and computational capabilities of DRHME. As there are few exact rarefied flow solutions in multi-dimensional case, we compare our results with the DSMC. The results of DRHME agree with the solutions of DSMC entirely well.

6.1 Shock tube test

The shock tube test has been used as a convergence test for several times in our previous work [4,5,7,9]. In this paper, this test is used not only to examine the convergence of the

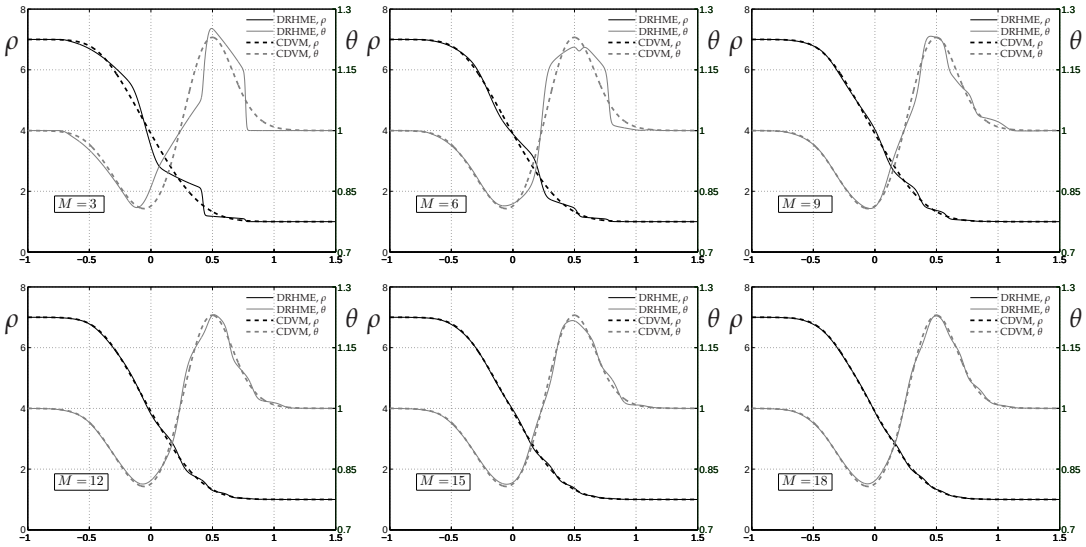


Figure 1: The density and temperature profile for the shock tube test of the BGK model with $Kn=0.5$.

DRHME, but also to compare the CPU time of the DRHME with that of the case without dimension-reduction. The initial value is

$$(\rho,\boldsymbol{u},\theta)=\begin{cases}(7,\mathbf{0},1), & x_1<0,\\ (1,\mathbf{0},1), & x_1>0,\end{cases} \tag{6.1}$$

and the fluid is in local equilibrium everywhere. Consistent with our previous work [4, 5, 7, 9], the BGK model is used, i.e. $\mathrm{Pr}=1$. The relaxation time is taken as

$$\tau=\frac{Kn}{\rho}.$$

Different Knudsen numbers from 0.5 up to 5 are tested.

The first case is $Kn=0.5$. The conservative discrete velocity model (CDVM) in [25] is applied to get a reference solution. For the CDVM, the dimension is also reduced, so that only 1D velocity space is considered. We discretize the 1D velocity space with 500 grid points, when the numerical result proves accurate enough. The computational domain is $x_1\in[-2,2]$ and 4000 grid points are used for spatial discretization for both the DRHME and the CDVM. In Fig. 1, the numerical results from the DRHME with $M=3,6,\cdots,18$ at time $t=0.3$ are given. The hyperbolic structure of the moment equations is observed instantly, and it is clear that the solution of the DRHME converges to that of the CDVM as M increases.

Now we use the shock tube test to compare the CPU time of the DRHME and the hyperbolic moment system (HME) without dimension-reduction (cf. [7] for details). The stopping time is $t=0.1$, and only 1000 spatial grid points are used, with

Table 1: Comparison of CPU time of DRHME and HME with same parameters and same algorithm.

M	DRHME		HME		Ratio of HME to DRHME	
	moment number	CPU time (sec)	moment number	CPU time (sec)	moment number	CPU time
3	13	6.93	20	8.97	1.54	1.30
4	21	11.55	35	18.09	1.67	1.57
5	31	16.07	56	31.82	1.81	1.98
6	43	23.00	84	56.34	1.95	2.45
7	57	31.67	120	89.21	2.11	2.82
8	73	43.53	165	133.17	2.26	3.06
9	91	56.92	220	190.94	2.42	3.35
10	111	72.01	286	265.68	2.58	3.69
11	133	89.08	364	354.19	2.74	3.98
12	157	109.60	455	466.04	2.90	4.25
13	183	131.80	560	605.14	3.06	4.59
14	211	155.23	680	757.10	3.22	4.88
15	241	185.13	816	943.78	3.39	5.10
16	273	217.06	969	1162.07	3.55	5.35
17	307	251.18	1140	1412.77	3.71	5.62
18	343	289.32	1330	1705.85	3.88	5.90

the other parameters the same as in the former test. The computation was carried on a desktop computer with CPU core speed 3GHz. The moment number of the HME is $(M+1)(M+2)(M+3)/6$ [3], while it is $(M+1)(M+2)/2+(M-1)M/2$ for the DRHME. The ratio of the moment number of the HME to that of the DRHME is

$$\frac{(M+1)(M+2)(M+3)/6}{(M+1)(M+2)/2+(M-1)M/2}\approx\frac{M+5}{6}+\frac{1}{M}.$$

Since $M\geq 3$, $(M+5)/6+1/M>1$, so the DRHME has fewer moments than the HME for a fixed M. In Table 1, CPU times of the DRHME and the HME with $M=3,6,\cdots,18$ are given. Clearly, the DRHME saves more computational time with a greater M. In our computation, in order to approximate Boltzmann equation the truncation order M is a fairly large number,-cf. the tests below, where the dimension-reduction of the moment method can provide significant efficiency improvement. One may observe that the ratio of the CPU time of the HME to the DRHME grows faster than that of moment number. Actually, on dividing the last column in Table 1 by the penultimate column, the result approximate a constant 1.5. This is partly attributed to the requirement that the higher dimension moment system needs more additional processing in the computation.

The second case is for $Kn=5$, where the computational domain $[-2,2]$ and 4000 grid points are used for the spatial discretization. This test demonstrates the convergence of the moment method at high Knudsen number. In microflow simulation, the fluid is in the

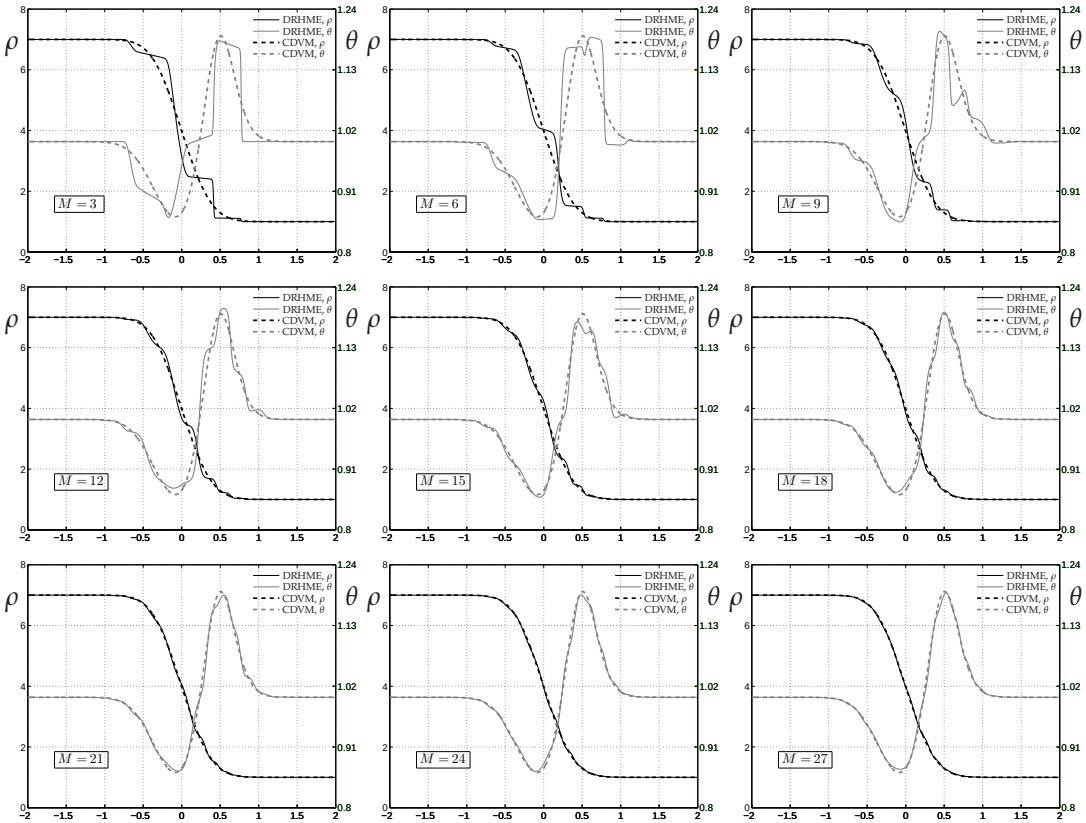
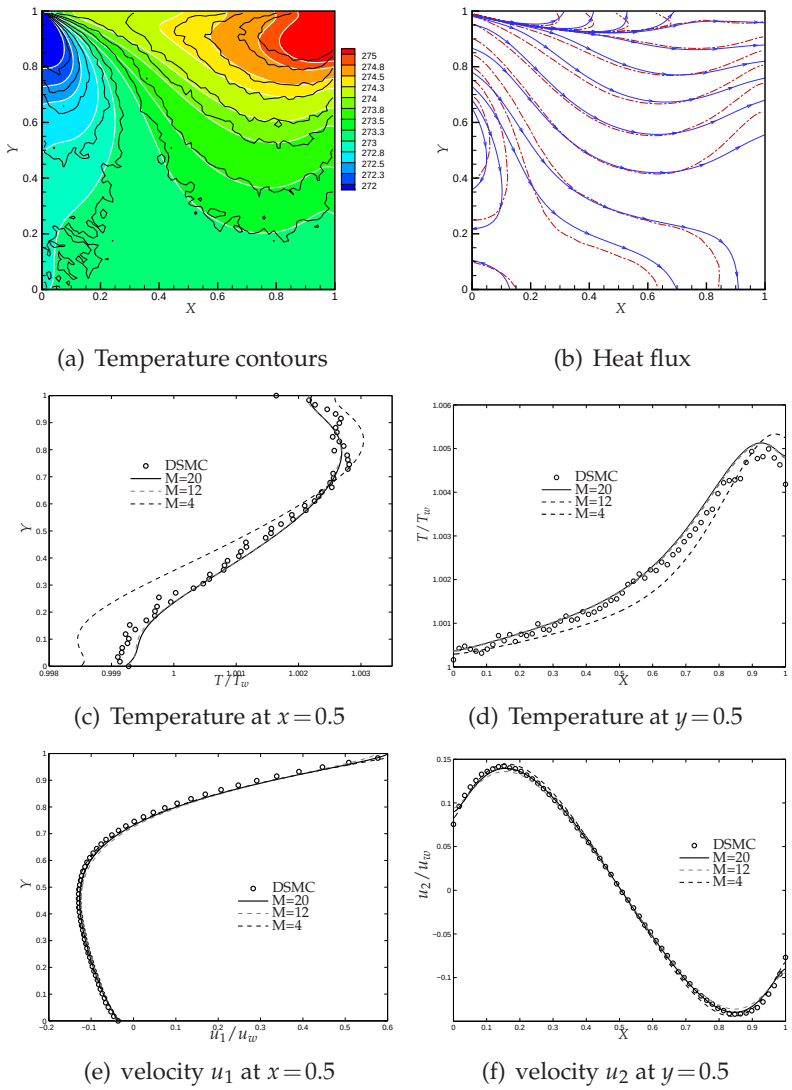


Figure 2: Density and temperature profiles for the shock tube test of the BGK model with $Kn=5$.

transition regime with $Kn>0.1$, so the moment system for high Knudsen number makes sense. The reference solution is again provided by CDVM. Fig. 2 shows the numerical result with $M=3,6\cdots,24$ at time $t=0.3$. It is clear that the numerical solutions of the DRHME converge to the reference solution as M increases. Comparing Figs. 1 and 2, the solution from the DRHME with $Kn=0.5$ converges to the reference solution faster than that when $Kn=5$. The reason is that the exact solution of the Boltzmann equation is further away from the Maxwellian with greater Kn, so greater M will be used in numerical simulation for greater Kn.

6.2 Cavity flow test

In the multi-dimensional case, there are few exact rarefied flow solutions. A common way to demonstrate the validity of a numerical method is to compare its solutions with DSMC's results. This test mainly follows a recent paper [19], where non-equilibrium heat transfer in a cavity using parallel DSMC method at four different Knudsen numbers $Kn=0.1,0.5,1,8$ was investigated.



(a) Temperature contours (b) Heat flux

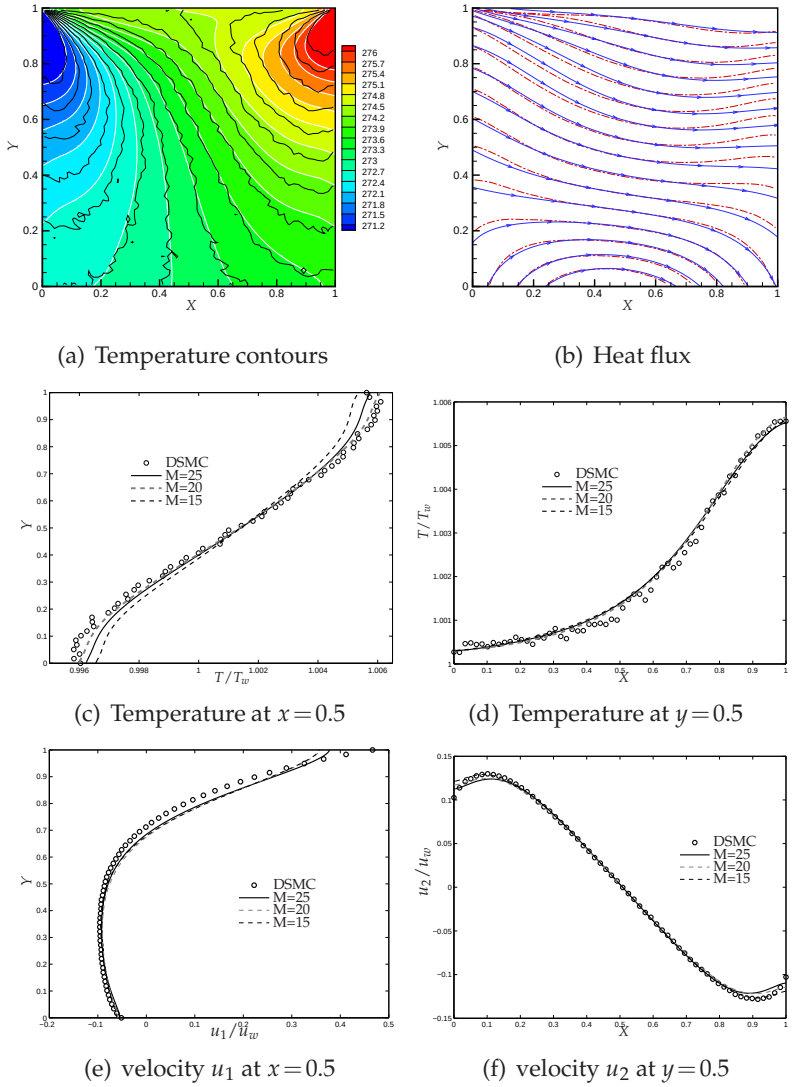
(c) Temperature at $x=0.5$ (d) Temperature at $y=0.5$

(e) velocity u_1 at $x=0.5$ (f) velocity u_2 at $y=0.5$

Figure 3: Cavity at $Kn=0.1$. (a) Temperature $T=\theta T_0$ contours with $M=20$, black lines: DSMC, white lines and background: DRHME. (b) Heat flux with $M=20$, red dash-dot line: DSMC, blue line: DRHME. (c), (d), (e), (f): 1D cross cut of temperatures (above) and velocities (below) along a vertical (left) and a horizontal (right) line crossing the center of the cavity of DRHME with different M and DSMC.

In this test, the gas is assumed to consist of monatomic molecules corresponding to argon (with mass $m=6.63\times10^{-26}$kg). In the DSMC solutions, the variable hard sphere (VHS) collision model has been used, with a reference particle diameter of $d=4.17\times10^{-10}$m. The wall temperature is set to the reference temperature, i.e., $T_w=T_0=273$K. The wall velocity is kept fixed, i.e., $U_w=50m/s$. The Knudsen number variation is achieved by using a vary density.

In order to match with the DSMC VHS model, the Shakhov model is used with Pr=



(a) Temperature contours (b) Heat flux

(c) Temperature at $x=0.5$ (d) Temperature at $y=0.5$

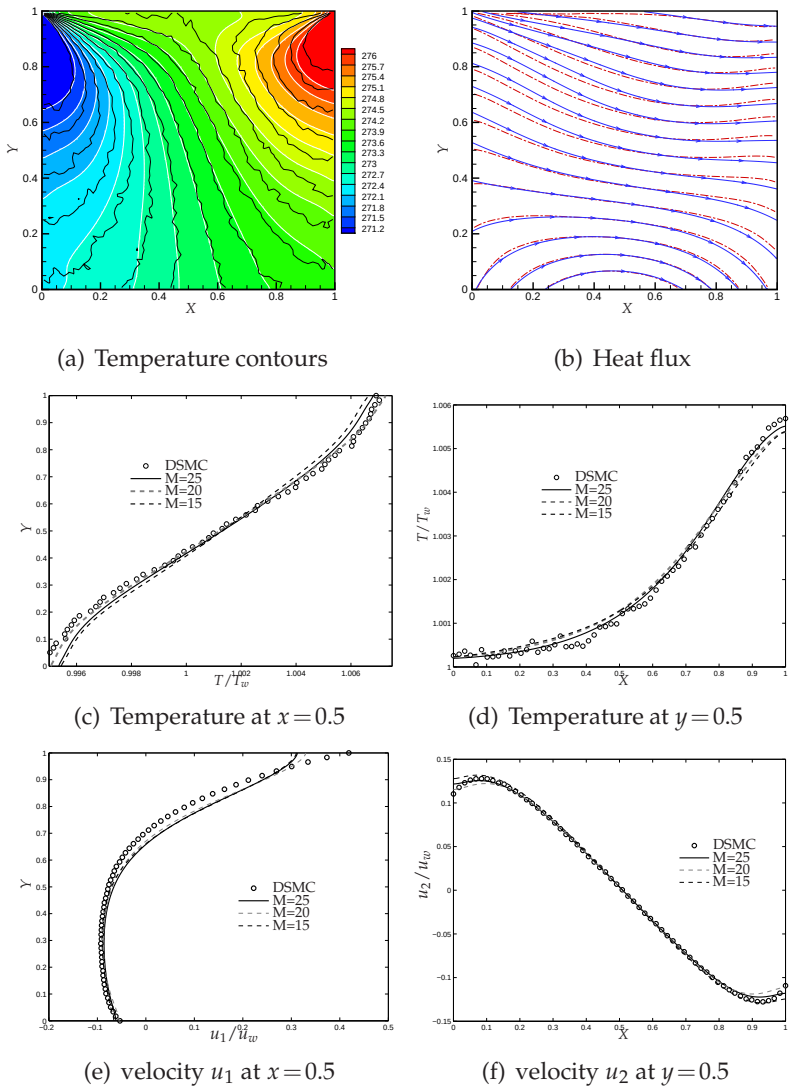
(e) velocity u_1 at $x=0.5$ (f) velocity u_2 at $y=0.5$

Figure 4: Cavity at $Kn=0.5$. (a) Temperature $T=\theta T_0$ contours with $M=25$, black lines: DSMC, white lines and background: DRHME. (b) Heat flux with $M=25$, red dash-dot line: DSMC, blue line: DRHME. (c), (d), (e), (f): 1D cross cut of temperatures (above) and velocities (below) along a vertical (left) and a horizontal (right) line crossing the center of the cavity of DRHME with different M and DSMC.

2/3, and the relaxation time is taken to be

$$\tau=\sqrt{\frac{2}{\pi}}\frac{Kn}{p}\theta^{\omega},$$

where $\omega=0.81$. For the boundary conditions, χ in boundary condition (2.10) is taken as $\chi=1$. In our computation, 200×200 grid points are used for spatial discretization, and different truncation orders are selected at different Knudsen numbers.



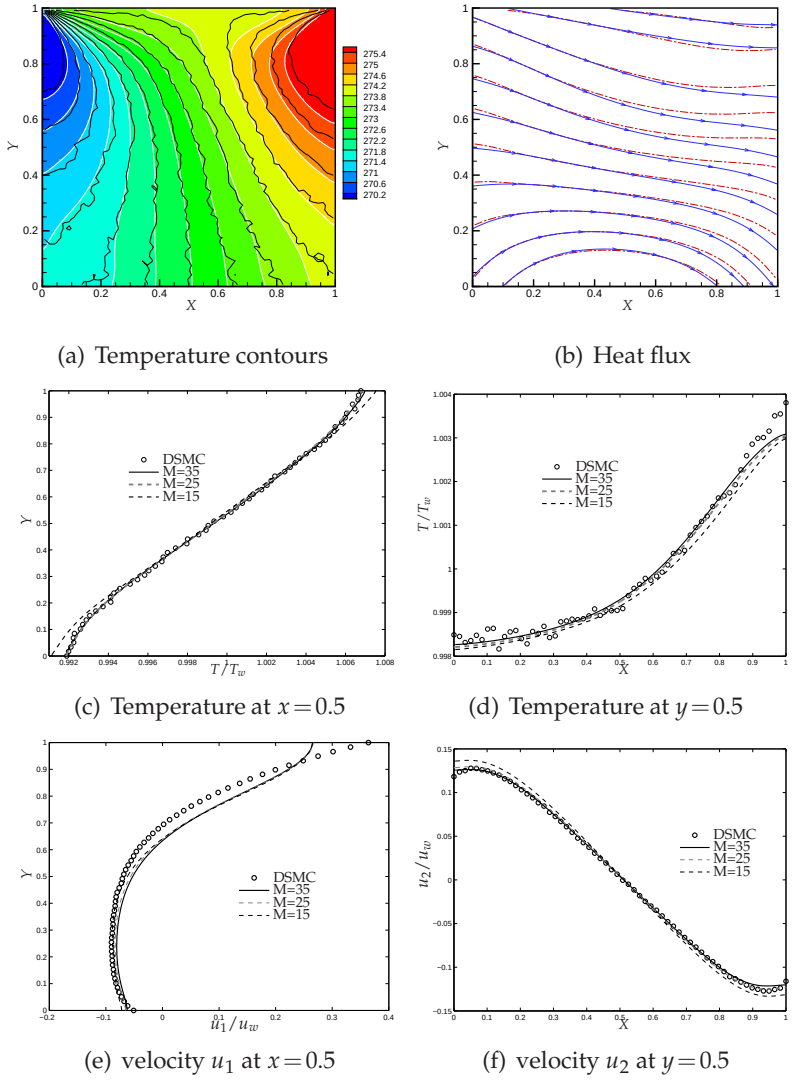
(a) Temperature contours (b) Heat flux

(c) Temperature at $x=0.5$ (d) Temperature at $y=0.5$

(e) velocity u_1 at $x=0.5$ (f) velocity u_2 at $y=0.5$

Figure 5: Cavity at $Kn=1$. (a) Temperature $T=\theta T_0$ contours with $M=25$, black lines: DSMC, white lines and background: DRHME. (b) Heat flux with $M=25$, red dash-dot line: DSMC, blue line: DRHME. (c), (d), (e), (f): 1D cross cut of temperatures (above) and velocities (below) along a vertical (left) and a horizontal (right) line crossing the center of the cavity of DRHME with different M and DSMC.

Thermal patterns at different Knudsen number are plotted in Figs. 3, 4, 5 and 6, showing temperature contours, heat flux, and 1D cross cut of temperatures and velocities along a vertical and a horizontal line crossing the center of the cavity at $Kn=0.1,0.5,1,8$. The solutions from the DSMC and the results from the DRHME are presented, and clearly, they agree well for almost all flow variables at different Knudsen numbers. From the plots of streamlines of heat flux, it is seen that flows are mainly from the cold to the hot region at the different Knudsen number, although there are slight deviations between



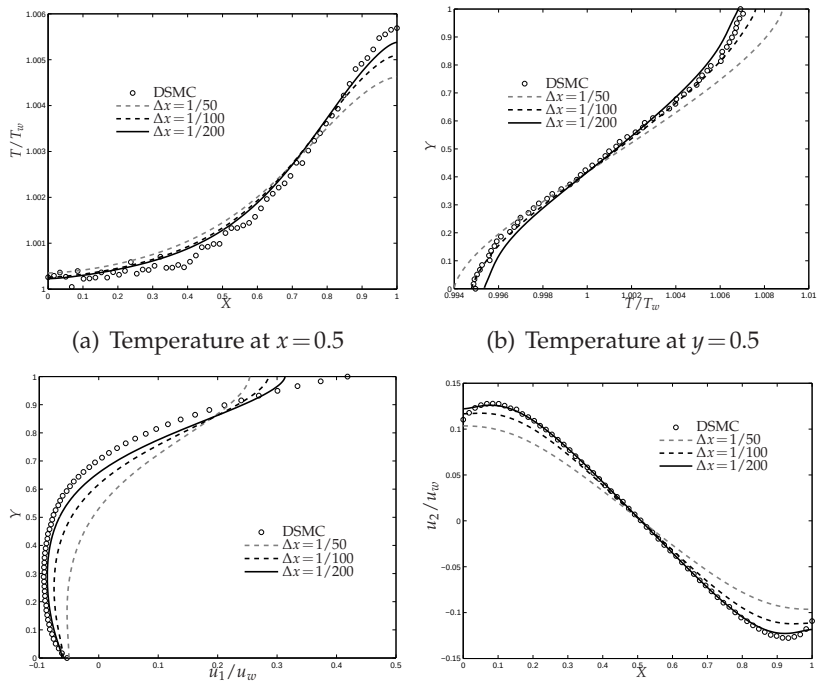
(a) Temperature contours (b) Heat flux

(c) Temperature at $x=0.5$ (d) Temperature at $y=0.5$

(e) velocity u_1 at $x=0.5$ (f) velocity u_2 at $y=0.5$

Figure 6: Cavity at $Kn=8$. (a) Temperature $T=\theta T_0$ contours with $M=35$, black lines: DSMC, white lines and background: DRHME. (b) Heat flux with $M=35$, red dash-dot line: DSMC, blue line: DRHME. (c), (d), (e), (f): 1D cross cut of temperatures (above) and velocities (below) along a vertical (left) and a horizontal (right) line crossing the center of the cavity of DRHME with different M and DSMC.

the results from the DSMC and DRHME close to the right boundary. The gaseous heat transfer direction denotes a counter-gradient heat flux, which implies that the Fourier's law may fail in the transition regime. For non-equilibrium flow, expansion cooling and compression of the gas flow affect the heat transport significantly. There is also good agreement in the temperature profile along vertical and horizontal symmetric lines. As M increases, the results of the DRHME agree with those from the DSMC even better. So do the velocity profiles along vertical and horizontal symmetric lines. On the other hand,



(a) Temperature at $x=0.5$ (b) Temperature at $y=0.5$

(c) velocity u_1 at $x=0.5$ (d) velocity u_2 at $y=0.5$

Figure 7: Cavity at $Kn=1$ and $M=25$ with different mesh: 1D cross cut of temperatures (above) and velocities (below) along a vertical (left) and a horizontal (right) line crossing the center of the cavity of DRHME and DSMC.

it is observed in Figs. 3, 4, 5 and 6 that when Kn is high (i.e., $Kn=1,8$), there is some difference between the results along the symmetric lines of the DRHME and the DSMC close to the boundary, even when the truncation order M is fairly high. This is probably due to excessive dissipation in the spatial discretization. The structure of solutions turns out to be more complex as the Knudsen number increases, when greater M and finer spatial grid have to be used.

Fig. 7 shows a 1D cross cut of temperatures and velocities along a vertical and a horizontal line crossing the center of the cavity at $Kn=1$ and $M=25$, with different mesh. It is observed that the results of the DRHME tends to those from the DSMC as the mesh is refined. With a fine enough mesh, the results of the DRHME agree with the DSMC quite well.

7 Conclusions

In this paper, a dimension-reduced hyperbolic moment method is proposed to solve the dimension reduced Boltzmann equations with both BGK and Shakhov models and the wall boundary conditions are developed. With dimension-reduction, the numerical effi-

ciency is significantly improved. We are now developing a high order scheme based on the method realized in this paper.

Acknowledgments

We wish to thank Dr. Xiao-Jun Gu for providing the DSMC solutions. The research of R. Li was supported in part by the National Basic Research Program of China (2011CB309704) and the National Natural Science Foundation of China (NSFC91330205). The research of Z.-H. Qiao was partly supported by the Hong Kong Research Council GRF grant (PolyU 2021/12P) and the Hong Kong Polytechnic University grant (A-PL61). Y.-W. Fan and Z.-N. Cai were supported by the Hong Kong RGC grant PolyU 2017/10P during their visits to the Hong Kong Polytechnic University.

Appendix: The degree the polynomial $X_\alpha^{(\phi)}$

We assert that $X_\alpha^{(\phi)}$ is a degree $|\alpha|$ polynomial about τ.

Proof. We use mathematical induction on $|\alpha|$. For simplicity, we rewrite Eq. (5.14) as

$$\frac{\mathrm{d}X_\alpha^{(\phi)}}{\mathrm{d}\tau}=\sum_{d=1}^{2}\left[a_d X_{\alpha-e_d}^{(\phi)}+\frac{b_d}{2}X_{\alpha-2e_d}^{(\phi)}\right],\quad |\alpha|\le M_\phi, \tag{A.1}$$

where a_d, b_d, $d=1,2$ are constants.

Basis: Considering the case $|\alpha|=0$, we get

$$\frac{\mathrm{d}X_0^{(\phi)}}{\mathrm{d}\tau}=0,$$

so $X_0^{(\phi)}$ is constant, i.e., degree 0 polynomial about τ.

Inductive step: For any $m\in\mathbb{N}$, for all $|\alpha|\le m$, assume that $X_\alpha^{(\phi)}$ is a degree $|\alpha|$ polynomial about τ, and let us consider the case $|\alpha|=m+1$. Since $|\alpha-e_d|=|\alpha|-1=m$, $X_{\alpha-e_d}^{(\phi)}$ is a degree m polynomial or zero (there is a negative index in $\alpha-e_d$). Analogically, $X_{\alpha-2e_d}^{(\phi)}$ is either a degree $m-1$ polynomial or zero. Hence

$$\sum_{d=1}^{2}\left[a_d X_{\alpha-e_d}^{(\phi)}+\frac{b_d}{2}X_{\alpha-2e_d}^{(\phi)}\right]$$

is a polynomial of degree no more than m. Thus, $X_\alpha^{(\phi)}$ is a degree $m+1$ polynomial about τ.

Since both the basis and the inductive step have been proved, it follows by mathematical induction that $X_\alpha^{(\phi)}$ is a degree $|\alpha|$ polynomial about τ. □

References

[1] 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(3) (1954), 511–525.

[2] G. A. Bird, Molecular Gas Dynamics and the Direct Simulation of Gas Flows, Oxford: Clarendon Press, 1994.

[3] Z. Cai, Y. Fan, and R. Li, Globally hyperbolic regularization of Grad's moment system, Commun. Pure Appl. Math., 2013.

[4] Z. Cai, Y. Fan, and R. Li, Globally hyperbolic regularization of Grad's moment system in one dimensional space, Commun. Math Sci., 11(2) (2013), 547–571.

[5] Z. Cai and R. Li, Numerical regularized moment method of arbitrary order for Boltzmann-BGK equation, SIAM J. Sci. Comput., 32(5) (2010), 2875–2907.

[6] Z. Cai, R. Li, and Z. Qiao, NRxx simulation of microflows with Shakhov model, SIAM J. Sci. Comput., 34(1) (2012), A339–A369.

[7] Z. Cai, R. Li, and Z. Qiao, Globally hyperbolic regularized moment method with applications to microflow simulation, Computer Fluids, 81 (2013), 95–109.

[8] Z. Cai, R. Li, and Y. Wang, An efficient NRxx method for Boltzmann-BGK equation, J. Sci. Comput., 50(1) (2012), 103–119.

[9] Z. Cai, R. Li, and Y. Wang, Numerical regularized moment method for high Mach number flow, Commun. Comput. Phys., 11(5) (2012), 1415–1438.

[10] C. K. Chu, Kinetic-theoretic description of the formation of a shock wave, Phys. Fluids, 8(1) (1965), 12–22.

[11] P. Degond, L. Pareschi, and G. Russo editors, Modeling and Computational Methods for Kinetic Equations, Birkhäuser, 2004.

[12] B. Dubroca and L. Mieussens, A conservative and entropic discrete-velocity model for rarefied polyatomic gases, in CEMRACS 1999 (Orsay), volume 10 of ESAIM Proc., pages 127–139, Paris, 1999, Soc. Math. Appl. Indust.

[13] F. Filbet and S. Jin, A class of asymptotic preserving schemes for kinetic equations and related problems with stiff sources, J. Comput. Phys., 229 (2010), 7625–7648.

[14] H. Grad, On the kinetic theory of rarefied gases, Commun. Pure Appl. Math., 2(4) (1949), 331–407.

[15] X. J. Gu and D. R. Emerson, A computational strategy for the regularized 13 moment equations with enhanced wall-boundary equations, J. Comput. Phys., 255(1) (2007), 263–283.

[16] X-J. Gu, D. R. Emerson, and G.-H. Tang, Kramers' problem and the Knudsen minimum: a theoretical analysis using a linearized 26-moment approach, Continuum Mech. Thermodyn., 21 (2009), 345–360.

[17] J. C. 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), 662–690.

[18] S. Jin and M. Slemrod, Regularization of the Burnett equations via relaxation, J. Stat. Phys., 103(5-6) (2001), 1009–1033.

[19] B. John, X.-J Gu, and D. R. Emerson, Investigation of heat and mass transfer in a lid-driven cavity under nonequilibrium flow conditions, Numer. Heat Transfer, 58 (2010), 287–303.

[20] G. Karniadakis, A. Beskok, and N. Aluru, Microflows: Fundamentals and Simulation, Springer, New York, 2001.

[21] G. Karniadakis, A. Beskok, and N. Aluru, Microflows and Nanoflows: Fundamentals and Simulation, volume 29 of Interdisciplinary Applied Mathematics, Springer, New York,

U.S.A., 2005.
[22] R. Kosik, Numerical Challenges on the Road to NanoTCAD, PhD thesis, Technische Universität Wien, August 2004.
[23] G. Dal Maso, P. G. LeFloch, and F. Murat, Definition and weak stability of nonconservative products, J. Math. Pures Appl., 74(6) (1995), 483–548.
[24] J. Clerk Maxwell, On stresses in rarefied gases arising from inequalities of temperature, Proc. R. Soc. Lond., 27(185-189) (1878), 304–308.
[25] L. Mieussens, Discrete velocity model and implicit scheme for the BGK equation of rarefied gas dynamics, Math. Models Methods Appl. Sci., 10(8) (2000), 1121–1149.
[26] I. Müller and T. Ruggeri, Rational Extended Thermodynamics, Second Edition, volume 37 of Springer Tracts in Natural Philosophy, Springer-Verlag, New York, 1998.
[27] S. Rhebergen, O. Bokhove, and J. J. W. van der Vegt, Discontinuous Galerkin finite element methods for hyperbolic nonconservative partial differential equations, J. Comput. Phys., 227(3) (2008), 1887–1922.
[28] E. M. Shakhov, Generalization of the Krook kinetic relaxation equation, Fluid Dyn., 3(5) (1968), 95–96.
[29] H. Struchtrup and M. Torrilhon, Regularization of Grad's 13 moment equations: Derivation and linear analysis, Phys. Fluids, 15(9) (2003), 2668–2680.
[30] M. Torrilhon, Hyperbolic moment equations in kinetic gas theory based on multi-variate Pearson-IV-distributions, Commun. Comput. Phys., 7(4) (2010), 639–673.
[31] K. Xu and J. C. Huang, A unified gas-kinetic scheme for continuum and rarefied flows, J. Comput. Phys., 229 (2010), 7747–7764.
[32] K. Xu and J. C. Huang, An improved unified gas-kinetic scheme and the study of shock structures, IMA J. Appl. Math., 76 (2011), 698–711.
[33] J. Y. Yang and J. C. Huang, Rarefied flow computations using nonlinear model Boltzmann equations, J. Comput. Phys., 120(2) (1995), 323–339.