

# Implicit Discontinuous Galerkin Method for the Boltzmann Equation

Wei Su, Peng Wang, Yonghao Zhang, Lei Wu\*

*James Weir Fluids Laboratory, Department of Mechanical and Aerospace Engineering,  
University of Strathclyde, G1 1XJ Glasgow, United Kingdom*

---

## Abstract

An implicit high-order discontinuous Galerkin (DG) method is developed to find steady-state solution of rarefied gas flow described by the Boltzmann equation with full collision operator. In the physical space, velocity distribution function is approximated by the piecewise polynomials of degree up to 4, while in the velocity space the fast spectral method is incorporated into the DG discretization to evaluate the collision operator. A specific polynomial approximation for the collision operator is proposed to reduce the computational complexity of the fast spectral method by  $K$  times, where for two-dimensional problems  $K$  is 15 when DG with 4th-order polynomials are used on triangular mesh. Based on the first-order upwind scheme, a sweeping technique is employed to solve the local linear equations resulting from the DG discretization sequentially over spatial elements. This technique can preserve stability of the scheme and requires no nonlinear limiter in solving hypersonic rarefied gas flow when the flow is fully resolved. Moreover, without assembling large sparse linear system, the computational cost in terms of memory and CPU time can be significantly reduced. Five different one/two-dimensional tests including low-speed microscale flows and hypersonic rarefied gas flows are used to verify the proposed approach. Our results show that, DG schemes of different order of approximating polynomial require the same number of iterative steps to obtain the steady-state solution with the same order of accuracy; and the higher order the scheme, fewer spatial elements thus less CPU time is needed. Besides, our method can be faster than the finite difference solver by about one order of magnitude. The produced solutions can be used as benchmark data for assessing the accuracy of other gas kinetic solvers for the Boltzmann equation and gas kinetic models that simplify the Boltzmann collision operator.

*Keywords:* discontinuous Galerkin, high-order discretization, Boltzmann equation, fast spectral method, implicit scheme

---

\*Corresponding author: lei.wu.100@strath.ac.uk

## 1. Introduction

In gas kinetic theory, the motion of molecules in dilute gas is mathematically described by the one-particle velocity distribution function and macroscopic flow properties are derived from its velocity moments [1]. Nowadays, this theory has been used for the description of transport phenomena in a wide range of scientific disciplines and applications such as the aerothermal dynamics in aerospace engineering, fusion processes in nuclear science, natural gas recovery and extraction in unconventional gas industry, freeze drying techniques in pharmaceutical and food manufactures, electron transport in semiconductor devices, and physics of diffuse matter in interstellar medium, just to name a few. In the Boltzmann's description, the variation of velocity distribution function comes from the linear streaming in the phase space, and the nonlinear interaction due to binary collisions. Thus, the distribution function is a seven dimensional variable, with three in the physical space, three in the molecular velocity space, and one in the temporal space. Meanwhile, the nonlinear collision operator is a fivefold operator with three dimensions in the velocity space and two dimensions in a unit sphere (i.e. solid angle).

The multi-dimensional structure of the Boltzmann equation poses a real challenge to its numerical solution [2]. Historically, two major categories of approaches have been developed. One is the stochastic approach, which uses simulation particles to represent a large number of real molecules and mimic the molecular behaviors. The prevail one is the direct simulation Monte Carlo (DSMC) method developed by Bird [3]. During the simulation, particles move through the spatial space in a realistic manner with respect to the time, while intermolecular collisions and molecule-surface interactions are calculated in probabilistic manners. The other category is the deterministic approach, which adopts a numerical quadrature to approximate the integration with respect to the molecular velocity on a set of fixed discrete points [4, 5]. As a result of discretization in the velocity space, the original kinetic equation is represented as a set of linear hyperbolic equations with nonlinear source terms that couple all the equations. To solve the resulting system, the usual schemes of traditional computational fluid dynamic (CFD) techniques for hyperbolic conservation laws can be applied straightforwardly for the streaming term. Some hybrid stochastic-deterministic approaches have also been proposed to solve the Boltzmann equation [6, 7, 8]. Note that the majority of methods are based on the splitting technique, where the streaming and collision are treated separately.

In this paper, we focus on the deterministic method to solve the Boltzmann equation, which requires proper treatment of the linear streaming operator and nonlinear collision operator. The finite difference method (FDM), finite volume method and finite element method have been successfully employed to approximate the derivatives with respect to the spatial ordinates [9, 10, 11, 12, 13, 14]. The advantage of these methods is that they have been well developed to achieve high order spatial and temporal accuracy. However, they might lose robustness and produce nonphysical solution, when the velocity distribution function has large variations and/or the kinetic equations become stiff. Another category of schemes is the semi-Lagrangian [15, 16, 17] and Lagrangian methods [18, 19], which is designed to ensure positivity of the solutions. The basic idea of the schemes is to solve the streaming for

the successive time steps by following the characteristics, i.e. molecular trajectories. The semi-Lagrangian methods still utilize fixed computational grid, but evaluate solution at the points that can be transported by the molecular velocity onto the computational grid within a time step. The Lagrangian methods update solution according to streaming without using a spatial mesh. Instead, the calculation reduces to a single manipulation for each discrete velocity.

For evaluating collision term, the most simple and widely used way is to replace the complicated collision operator by a relatively simple kinetic model, such as the Bhatnagar-Gross-Krook (BGK) model [20], ellipsoidal statistical BGK model [21], and Shakhov model [22], which describes the relaxation of distribution function to the local equilibrium distributions obtained from macroscopic flow properties. Otherwise, the full Boltzmann collision operator should be calculated. The attempts to directly solve the full Boltzmann collision operator started from the late 1980s. Goldstein et al. constructed a discrete collision mechanics on the velocity nodes, which can preserve the main physical properties of the collision operator [23]. However, a large amount of discrete velocities are required, since post-collision velocities must fall on the grid points. The computational cost is of the order  $O(\bar{N}^7)$  ( $\bar{N}$  is the number of points in each velocity direction), and the nominal accuracy is less than first order in the velocity space [24]. Improvement by using an interpolation to map the post-collision velocities onto the velocity grid makes the performance of the scheme comparable to or even faster than DSMC in normal shock wave simulation [25]. The kinetic theory group in Kyoto introduced another family of methods evaluating collision in the velocity space [26, 27, 28, 29], in which the distribution function is expanded in terms of basis functions, while the collision operator is computed by the product of the expansion coefficients obtained at the discrete velocities and the numerical kernels that are the collision operators for the basis functions. The numerical kernels are pre-computed by numerical integration, which are restricted to the hard-sphere model and distribution function with cylindrical symmetry. Note that there are other schemes such as the projection method that evaluates the collision operator over a set of collision pairs with different velocities, aim distances and reflect angles [30, 31], as well as method based on nodal-discontinuous Galerkin discretization of the collision operator and a bi-linear convolution of the Galerkin projection [32].

Instead of directly calculating the collision integral on discrete velocities, there is another route to approximate collision in frequency domain using Fourier transform techniques. These methods can present accuracy of typical spectral approaches. Besides, they can reduce computational cost through fast spectral algorithm. The pioneering work was introduced by Bobylev for Maxwell molecules [33]. Then, several spectral methods were developed, which have computational cost of the order up to  $O(\bar{N}^6)$  [34, 35]. Their computational cost can be reduced to  $O(\bar{N}^3 \log \bar{N})$  for distribution function possessing cylindrical symmetry, when the fast Fourier transform (FFT) and Hankel transform are employed [36]. However, the accuracy is only of  $O(\bar{N}^{-1/2})$ . Based on the Carleman representation, an algorithm was developed for hard-sphere molecules to achieve accuracy of  $O(\bar{N}^{-2})$ , where the integration over the unit sphere is separated from the one over the velocity space [37]. By employing generalized Radon and X-ray transform, its computational cost is of  $O(\bar{N}^6 \log \bar{N})$ . The

algorithm for variable hard-sphere molecules of accuracy  $O(\bar{N}^{-2})$  was also proposed with complexity of  $O(\bar{N}^6)$  [38]. The fast spectral method (FSM) that is spectrally accurate has been developed since the new century [39, 40]. By means of the Carleman representation, the method is improved with the computational cost reduced to  $O(\bar{M}^2 \bar{N}^3 \log \bar{N})$ , where  $\bar{M}$  is the number of polar and azimuthal angles [41, 42]. This is in general the fastest algorithm to date. To extend the applicability of the FSM, novel anisotropic collision kernels were designed and incorporated, which can deal with all inverse power-law potentials (except the Coulomb potential) as well as the Lennard-Jones potential [43, 44]. Later, the collision kernel for Lennard-Jones potential was fully resolved, however, the computational cost increases to  $O(\bar{M}^2 \bar{N}^4 \log \bar{N})$  [45]. This method has been successfully applied to solve many canonical rarefied gas flows, where the computational efficiency is much higher than the low-variance DSMC method for low-speed flows [44, 46].

The challenge to numerically solve the Boltzmann equation with full collision operator is that the computational cost becomes immediately prohibitive for realistic problems, since: 1) the number of governing equations is large due to discretization in the velocity space; 2) for each equation, the collision operator needs to be evaluated at every spatial grid points or elements (even the Lagrangian methods need spatial mesh for approximation of collision). Therefore, high-order CFD approach is critical to improve efficiency of discretization in the spatial space, thus reduce the computational cost. One of the promising methods for this purpose is the discontinuous Galerkin (DG) method, which was first introduced for the neutron transport equation [47]. The DG method provides advantages including: achieving high-order of accuracy with relatively low effort, easy formulation for arbitrary geometry, straightforward implementation of boundary condition with the same high-order accuracy as in the interior of the computational domain, as well as the efficient implementation for parallelism and adaptive refinement. After combining an explicit high-order Runge-Kutta time marching scheme, the method has great success in solving convection-dominated problems [48, 49]. The explicit DG method has been applied to solve the kinetic model equations [50]. Very recently, it has also been applied to the full Boltzmann equation with the variable soft-sphere collision kernel, in which the collision operator is calculated based on a FSM having a cost at the order of  $O(\bar{M}^2 \bar{N}^4 \log \bar{N})$  [51]. It has been shown that the second-order DG method is 15 times faster than the second-order finite volume scheme [50]. However, higher-order explicit DG scheme is not superior to the lower-order one, mainly due to the fact that the time step restricted by the Courant-Friedrichs-Lowy (CFL) condition becomes extremely small [52]; thus the number of iteration becomes very large in finding steady-state solution for high-order discretization.

Note that the FSM has also been incorporated in the Boltzmann solver based on Lagrangian method for streaming [53]. The solution from this method is currently limited to first-order accuracy in space and time. Besides, the time marching is an explicit scheme, thus the total number of time steps is still enormous to obtain a steady solution. It is also interesting to mention that there is a class of methods, named (discrete) unified gas-kinetic scheme, sharing some properties with the semi-Lagrangian scheme [54, 55, 56, 57], in which the flux transport across spatial cell interface contains the evolution of distribution function along the molecular trajectories within a time step due to both stream and collision

processes. By coupling the evaluations of both streaming and collision, the scale of spatial discretization can be reduced. These methods are first developed based on kinetic model equations. Recently, approximation of the Boltzmann collision operator using the FSM is incorporated to correct the relaxation of distribution functions to the local equilibrium states beyond the continuum flow regime [54].

In this paper, we represent a DG method to solve the full Boltzmann equation, which is devoted to improving the scheme in the following ways:

- Implicit iterative schemes are employed to relax the limitation on time step from the CFL condition. As a result, the superiority of high-order discretization in the DG method can be demonstrated, which is in sharp contrast to the explicit DG where the CFL number is rather small.
- An novel scheme is proposed to reduce the computational complexity when using the FSM to calculate the collision operator, say, by 15 times when using 4th order approximating polynomials on two-dimensional triangular mesh.
- A strategy based on the sweeping technique is introduced, which can avoid solving large sparse linear system, and stabilize the scheme without using any nonlinear limiter when the rarefied gas flow is fully resolved.

The remainder of the paper is organized as follows. In Sec. 2, the Boltzmann equation and the FSM are introduced. In Sec. 3, the implicit DG method is described with details in the formulation of collision operator. A scheme to reduce the complexity of DG discretization for the collision operator is proposed in Sec. 4, while the sweeping strategy to solve the linear systems is described in Sec. 5. In Sec. 6, five different problems including one-dimensional shock wave, two-dimensional hypersonic flow past a square cylinder, lid driven cavity flow and two thermal low-speed microscale flows are simulated to assess the accuracy and efficiency of the proposed scheme. Conclusions are presented in Sec. 7.

## 2. The Boltzmann Equation and the Fast Spectral Method

In kinetic theory, the state of a gas system is described by the one-particle velocity distribution function  $f(t, \mathbf{x}, \mathbf{v})$ , which is a function of the time  $t$ , the spatial position  $\mathbf{x} = (x_1, x_2, x_3)$ , and the molecular velocity  $\mathbf{v} = (v_1, v_2, v_3)$ . Neglecting the external force, the evolution of velocity distribution function for a single-species monatomic gas is governed by the following Boltzmann equation:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} = \mathcal{C}(f, f_*), \quad (1)$$

where  $\mathcal{C}(f, f_*)$  is the Boltzmann collision operator that is usually split into the gain term  $\mathcal{C}_+$  and loss term  $\mathcal{C}_-$ :

$$\mathcal{C}(f, f_*) = \underbrace{\int \int B(\theta, |\mathbf{v} - \mathbf{v}_*|) f(\mathbf{v}'_*) f(\mathbf{v}') d\Omega d\mathbf{v}_*}_{\mathcal{C}_+} - \underbrace{\nu(\mathbf{v}) f(\mathbf{v})}_{\mathcal{C}_-}. \quad (2)$$

with the collision frequency

$$\nu(\mathbf{v}) = \int \int B(\theta, |\mathbf{v} - \mathbf{v}_*|) f(\mathbf{v}_*) d\Omega d\mathbf{v}_*. \quad (3)$$

Note that here  $B(\theta, |\mathbf{v} - \mathbf{v}_*|)$  is the collision kernel;  $\mathbf{v}, \mathbf{v}_*$  are the pre-collision molecular velocities of a collision pair, and  $\mathbf{v}', \mathbf{v}'_*$  are the corresponding post-collision molecular velocities;  $\Omega$  is the unit vector along the relative post-collision velocity  $\mathbf{v}' - \mathbf{v}'_*$ , while  $\theta$  is the deflection angle between the pre- and post-collision relative velocities. For simplicity, the time and spatial position is omitted in writing the distribution function, collision operator, and collision frequency.

The velocity distribution function is defined such that  $f(t, \mathbf{x}, \mathbf{v}) d\mathbf{x} d\mathbf{v}$  is the number of gas molecules in the phase-space volume  $d\mathbf{x} d\mathbf{v}$ . All macroscopic quantities, such as mass density  $\rho$ , bulk velocity  $\mathbf{u} = (u_1, u_2, u_3)$ , temperature  $T$ , pressure tension  $\mathbf{P}$  and heat flux  $\mathbf{Q} = (Q_1, Q_2, Q_3)$  can then be calculated via velocity moments of the distribution function. For simplicity, we use non-dimensional variables hereafter:  $\mathbf{x}$  is normalized by a characteristic flow length  $H$ ,  $T$  is normalized by a reference temperature  $T_0$ ,  $\rho$  is normalized by the average density  $\rho_0$  at  $T_0$ ,  $\mathbf{v}$  and  $\mathbf{u}$  are normalized by the most probable speed  $v_m = \sqrt{2k_B T_0 / m}$  with  $k_B$  and  $m$  being the Boltzmann constant and molecular mass, respectively,  $t$  is normalized by  $H/v_m$ ,  $f$  is normalized by  $\rho_0/m v_m^3$ ,  $\mathbf{P}$  is normalized by  $\rho_0 k_B T_0 / m$ , and  $Q_i$  is normalized by  $\rho_0 k_B T_0 v_m / m$ . Therefore, we have

$$\begin{aligned} \rho &= \int f d\mathbf{v}, \quad \mathbf{u} = \frac{1}{\rho} \int \mathbf{v} f d\mathbf{v}, \quad T = \frac{2}{3\rho} \int |\mathbf{v} - \mathbf{u}|^2 f d\mathbf{v}, \\ \mathbf{P} &= 2 \int (\mathbf{v} - \mathbf{u}) \otimes (\mathbf{v} - \mathbf{u}) f d\mathbf{v}, \quad \mathbf{Q} = \int (\mathbf{v} - \mathbf{u}) |\mathbf{v} - \mathbf{u}|^2 f d\mathbf{v}. \end{aligned} \quad (4)$$

The collision kernel  $B(\theta, |\mathbf{v} - \mathbf{v}_*|)$ , depending on the modules of the pre-collision relative velocity and the deflection angle, is only determined when a certain intermolecular potential is given [1]. The detailed structure of the collision kernel is usually very complicated, except that of the ideal hard-sphere molecule. In the history, both for the analytical and numerical convenience, specific simplification is adopted with the aim to recover the correct transport coefficients, which results in various molecular models that are widely used in the DSMC method. The key to these models is that transport coefficients such as shear viscosity, thermal conductivity, and diffuse coefficient are recovered over the temperature range considered. In this paper, the collision kernel is modeled as [42]:

$$B(\theta, |\mathbf{v} - \mathbf{v}_*|) = \frac{5|\mathbf{v} - \mathbf{v}_*|^{2(1-\omega)}}{2^{7-\omega} \Gamma\left(\frac{5-2\omega}{2}\right) Kn} \sin^{1-2\omega}\left(\frac{\theta}{2}\right), \quad (5)$$

where  $\Gamma$  is the Gamma function,  $\omega$  is the viscosity index (i.e. the shear viscosity  $\mu$  of the gas is proportional to  $T^\omega$ ) and  $Kn$  is the unconfined Knudsen number given at the reference condition:

$$Kn = \frac{\mu(T = T_0)}{\rho_0 H} \sqrt{\frac{m\pi}{2k_B T_0}}. \quad (6)$$

It is noted that the specific form (5) introduced by Mouhot and Pareschi enables the development of Carleman-representation-based FSM to deterministically compute the collision operator. It has the ability to mimic the growth trend of collision kernel when decreasing the deflection angle and recover correct values of shear viscosity, however it cannot deal with general forms of soft potentials. By introduce another free-parameter into the collision kernel, the authors have extended the applicability of FSM to all inverse power law potentials (except the Coulomb potential), thus to recover the correct value of diffusion coefficient [43, 44]. We also mention that more general collision models including the Lennard-Jones potential has been incorporated into the FSM [45, 51]. For general collision kernel, the computational cost will be one order of magnitude higher than that of Eq. (5); therefore, in this paper Eq. (5) is adopted to demonstrate efficiency and accuracy of the DG method on the spatial discretization. As a matter of fact, if viscosity index is chosen appropriately, the collision kernel can yield accurate results when compared to that of the realistic Lennard-Jones potential [44, 45].

### 2.1. The fast spectral method

The Boltzmann collision operator (2) is a five-fold integral with three dimensions in the molecular velocity space and two dimensions in a unit sphere. In this paper, the FSM is applied to evaluate the collision operator, details of which can be found in [42, 43, 45].

Firstly, the distribution function is periodized on a truncated domain  $\mathcal{D} = [-L, L]^3$  and expanded into Fourier series with  $N_1 \times N_2 \times N_3$  components:

$$f(t, \mathbf{x}, \mathbf{v}) = \sum_{j=-N/2}^{N/2-1} \bar{f}^j(t, \mathbf{x}) \exp(\imath \boldsymbol{\xi}^j \cdot \mathbf{v}), \quad (7)$$

$$\bar{f}^j(t, \mathbf{x}) = \frac{1}{(2L)^3} \int_{\mathcal{D}} f(t, \mathbf{x}, \mathbf{v}) \exp(-\imath \boldsymbol{\xi}^j \cdot \mathbf{v}) d\mathbf{v}, \quad (8)$$

where  $L$  is the maximum truncated velocity,  $\imath$  is the imaginary unit,  $\bar{f}^j$  is the spectrum of the velocity distribution function,  $\boldsymbol{\xi}^j = j\pi/L$  is the discrete frequency with  $j = (j_1, j_2, j_3)$  and  $N = (N_1, N_2, N_3)$  denoting the index and total number of frequencies. In order to take the advantage of FFT, the discretized frequency components are equally spaced.

Then, the gain term in collision operator and the collision frequency are evaluated through expanding in Fourier series:

$$\mathcal{C}_+ = \sum_{j=-N/2}^{N/2-1} \bar{\mathcal{C}}_+^j \exp(\imath \boldsymbol{\xi}^j \cdot \mathbf{v}), \quad \nu = \sum_{j=-N/2}^{N/2-1} \bar{\nu}^j \exp(\imath \boldsymbol{\xi}^j \cdot \mathbf{v}), \quad (9)$$

where the  $j$ -th Fourier modes of the gain term in Eq. (2) and collision frequency (3) are calculated from the spectrum  $\bar{f}$  as follows [43, 44]:

$$\bar{\mathcal{C}}_+^j = \sum_{\substack{l+m=j \\ l,m=-N/2}}^{N/2-1} \bar{f}^l \bar{f}^m \beta(l, m), \quad \bar{\nu}^j = \bar{f}^j \beta(j, j). \quad (10)$$

Here,  $\beta$  is the collision kernel mode, whose  $(l, m)$ -th component is approximated through  $M_{\text{qua}}$ -point numerical quadrature in spherical coordinates as:

$$\beta(l, m) \simeq \frac{20 \sum_{p,q=1}^{M_{\text{qua}}} \sin(\theta_p) \Psi\left(\sqrt{|\boldsymbol{\xi}^m|^2 - (\boldsymbol{\xi}^m \cdot \mathbf{e}_{p,q})^2}\right) \Phi(\boldsymbol{\xi}^l \cdot \mathbf{e}_{p,q}) \varpi_p \varpi_q}{2^{7-\omega} \Gamma\left(\frac{5-2\omega}{2}\right) Kn}, \quad (11)$$

where  $\theta_p$  ( $\phi_q$ ) and  $\varpi_p$  ( $\varpi_q$ ) are the  $p$  ( $q$ )-th point and weight of the quadrature rule, respectively, for  $\theta, \phi \in [0, \pi]$ , and  $\mathbf{e}_{p,q} = (\sin \theta_p \cos \phi_q, \sin \theta_p \sin \phi_q, \cos \theta_p)$ . The functions  $\Psi$  and  $\Phi$  are  $\Psi(a) = 2\pi \int_0^R \rho^{1-\gamma} J_0(\rho a) d\rho$  and  $\Phi(a) = 2 \int_0^R \rho^{2(1-\omega)+\gamma} \cos(\rho a) d\rho$ , where  $J_0$  is the zeroth-order Bessel function, and  $R$  is the radius of the sphere to support the distribution function, which is chosen approximately as  $R = 2\sqrt{2}L/(2 + \sqrt{2})$  to avoid the aliasing error [43]. Note that by estimating through numerical quadrature, frequencies  $\boldsymbol{\xi}^m$  and  $\boldsymbol{\xi}^l$  appear in two different functions in the final form of  $\beta(l, m)$ , thus Eq. (10) can be calculated by FFT-based convolution [43].

### 3. Implicit Discontinuous Galerkin Method

To obtain the stationary solution of the Boltzmann equation, the following implicit iterative scheme is usually applied:

$$\bar{\nu} f^{(t+1)} + \mathbf{v} \cdot \frac{\partial f^{(t+1)}}{\partial \mathbf{x}} = \bar{\nu} f^{(t)} + \mathcal{C}(f^{(t)}, f_*^{(t)}), \quad (12)$$

where the superscripts  $(t)$  and  $(t+1)$  represent two consecutive iteration steps. The iteration is terminated when the convergence to the steady solution is achieved. The parameter  $\bar{\nu}$  is a positive constant which is the reciprocal time step in the backward-Euler method and highly influences convergence property of the iterative scheme: too large (small)  $\bar{\nu}$  results in slow convergence (numerical instability). Usually, to strike a balance between efficiency and stability of the iteration,  $\bar{\nu}$  is chosen to be the order of mean collision frequency  $\int \nu(\mathbf{v}) f(\mathbf{v}) d\mathbf{v}$ . Therefore, a safe choice of  $\bar{\nu}$  is the minimum mean collision frequency in the whole computational domain. However, one needs a good estimation for the minimum  $\bar{\nu}$  before calculation.

Another way to find the steady-state solution is to neglect the derivative of distribution function with respect to the time, yielding  $\mathbf{v} \cdot \partial f / \partial \mathbf{x} = \mathcal{C}$ . Then, the collision frequency and gain term of the Boltzmann collision operator are evaluated based on the approximation of distribution at the iteration step  $t$ , while other terms are solved at the next iteration step by:

$$\nu^{(t)}(\mathbf{x}, \mathbf{v}) f^{(t+1)}(\mathbf{x}, \mathbf{v}) + \mathbf{v} \cdot \frac{\partial f^{(t+1)}(\mathbf{x}, \mathbf{v})}{\partial \mathbf{x}} = \mathcal{C}_+^{(t)}(\mathbf{x}, \mathbf{v}), \quad (13)$$

In the following sections, we will denote the iterative scheme (12) with mean collision frequency as ‘ITR-MEAN’ and the iterative scheme (13) with local collision frequency as ‘ITR-LOC’. The two iteration schemes can lead to different computational complexity and convergence history, which will be discussed in Sec. 5. For conciseness, we will omit the index of iteration step in the following unless necessary.

### 3.1. DG formulation for the Boltzmann equation

Now we present the DG method to find the steady-state solutions of rarefied gas flows described by (12) and (13). Let  $\Delta \in \mathbb{R}^d$  be a computational domain in the  $d$ -dimensional spatial space with boundary  $\partial\Delta$ . Then, the domain is partitioned into  $M_{\text{el}}$  disjoint regular elements  $\Delta_i$ . The DG method provides an approximate solution to the velocity distribution function  $f$  on each element  $\Delta_i$  in some piecewise finite element spaces  $\mathcal{V}$  of the following form:

$$\mathcal{V} = \{\varphi_r(\mathbf{x}) : \varphi_r|_{\Delta_i} \in \mathcal{P}^k(\Delta_i), r = 1, \dots, K, \forall \Delta_i \in \Delta\}, \quad (14)$$

where  $\mathcal{P}^k$  denotes the space of  $k$ -th order polynomials, thus we have

$$f(\mathbf{x}, \mathbf{v}) = \sum_{r=1}^K \varphi_r(\mathbf{x}) F_r(\mathbf{v}), \quad (15)$$

with  $F_r$  being the degree of freedom for the distribution function. In general, the degrees of freedom are unknowns for which the equations are being solved. Together with the basis functions  $\varphi_r$ , they give the final polynomial estimation of  $f$  within a spatial element  $\Delta_i$ . The number of degree of freedom,  $K$ , dependents on the shape of element employed. For example,  $K = k + 1$  for line segament in one-dimensional (1D) problem, and  $K = (k + 1)(k + 2)/2$  for triangular element in two-dimensional (2D) problem.

In order to determine  $F_r$ , standard techniques of finite element formulations are applied to obtain the weak formulation of the governing system. Introducing  $(\cdot)$  and  $\langle \cdot \rangle$  as  $(a, b)_{\Delta_i} = \int_{\Delta_i} (a \cdot b) d\mathbf{x}$  and  $\langle a, b \rangle_{\partial\Delta_i} = \int_{\partial\Delta_i} (a \cdot b) d\Upsilon$  to denote operators on the element  $\Delta_i$  and its boundary  $\partial\Delta_i$ , respectively, we find the approximation of distribution function satisfies the following equation (take the ITR-MEAN scheme (12) as an example):

$$-(\nabla \varphi_s, \mathbf{v} f)_{\Delta_i} + \langle \varphi_s, \hat{\mathbf{H}} \cdot \mathbf{n} \rangle_{\partial\Delta_i} + (\varphi_s, \bar{\nu} f)_{\Delta_i} = (\varphi_s, \mathcal{C})_{\Delta_i} + (\varphi_s, \bar{\nu} f)_{\Delta_i}, \quad (16)$$

where  $s = 1, \dots, K$ ,  $\mathbf{n}$  is the outward unit normal vector, and  $\hat{\mathbf{H}}$  is the numerical flux that depends on the solutions from both sides of  $\partial\Delta_i$ , since the solution of  $f$  is discontinuous there. We define the numerical flux from the first-order upwind principle as:

$$\hat{\mathbf{H}} \cdot \mathbf{n} = \frac{1}{2} \mathbf{v} \cdot \mathbf{n} (f + f_{\text{ext}}) + \frac{1}{2} |\mathbf{v} \cdot \mathbf{n}| (f - f_{\text{ext}}). \quad (17)$$

with  $f_{\text{ext}}$  being the distribution from a neighboring element that shares the boundary  $\partial\Delta_i$  with  $\Delta_i$ . If  $\partial\Delta_i$  is at the boundary of computational domain, i.e.  $\partial\Delta_i \cap \partial\Delta \neq 0$ ,  $f_{\text{ext}}$  is evaluated using the given boundary condition.

Now, we focus on the formulation of  $(\varphi_s, \mathcal{C})_{\Delta_i} = (\varphi_s, \mathcal{C}_+)_{\Delta_i} - (\varphi_s, \nu f)_{\Delta_i}$  in Eq. (16). Inserting the polynomial expansion of distribution function (15) into Eq. (8), the  $j$ -th spectrum component of the distribution function can be rewritten in the following polynomial form:

$$\bar{f}^j(\mathbf{x}) = \sum_{r=1}^K \varphi_r(\mathbf{x}) \bar{F}_r^j, \quad (18)$$

where  $\bar{F}_r^j = \frac{1}{(2L)^3} \int_{\mathcal{D}} F_r(\mathbf{v}) \exp(-i\boldsymbol{\xi}^j \cdot \mathbf{v}) d\mathbf{v}$  is the spectrum of the degree of freedom.

With some algebraic calculations from Eqs. (9) and (10), the DG discretization of the gain term of the Boltzmann collision operator and the collision frequency are expressed as

$$\mathcal{C}_+ = \sum_{p=1}^K \sum_{r=1}^K \varphi_p \varphi_r \Xi_{p,r}, \quad \nu = \sum_{p=1}^K \varphi_p \Lambda_p, \quad (19)$$

where

$$\Xi_{p,r} = \sum_{j=-N/2}^{N/2-1} \sum_{\substack{l+m=j \\ l,m=-N/2}}^{\substack{N/2-1}} \bar{F}_p^l \bar{F}_r^m \beta(l, m) \exp(i\boldsymbol{\xi}^j \cdot \mathbf{v}), \quad \Lambda_p = \sum_{j=-N/2}^{N/2-1} \bar{F}_p^j \beta(j, j) \exp(i\boldsymbol{\xi}^j \cdot \mathbf{v}). \quad (20)$$

Finally, we obtain that

$$(\varphi_s, \mathcal{C}_+)_{\Delta_i} = \sum_{p=1}^K \sum_{r=1}^K (\varphi_s, \varphi_p \varphi_r)_{\Delta_i} \Xi_{p,r}, \quad (21)$$

$$(\varphi_s, \nu f)_{\Delta_i} = \sum_{p=1}^K \sum_{r=1}^K (\varphi_s, \varphi_p \varphi_r)_{\Delta_i} \Lambda_p F_r, \quad (22)$$

### 3.2. Discretization in the molecular velocity space

In order to obtain the macroscopic flow properties (4) and the spectrum (8), integrals with respect to the velocity space should be calculated. Numerically, the truncated but continuous velocity domain  $\mathcal{D}$  needs to be represented by  $M := (M_1, M_2, M_3)$  discrete points  $\mathbf{v}^{j'}$  and the integrals are approximated by certain quadrature rules, e.g.  $\rho = \sum_{j'=1}^M f(\mathbf{x}, \mathbf{v}^{j'}) w^{j'}$  with  $w^{j'}$  being the quadrature weight for the corresponding discretized velocity points  $\mathbf{v}^{j'}$ . The discrete velocities are not necessarily equidistant, especially for low-speed microflows with large Knudsen numbers, where the distribution function varies rapidly around  $\mathbf{v} = 0$  due to gas-wall interaction and nonuniform velocity points with refinement in this area is more efficient to capture the variation of  $f$  [58]. However, it should be emphasized that the FFT-based convolution could be efficiently employed only when the frequency space is uniformly discretized, though the number of frequency components can be smaller than that of velocity grid points due to the spectral accuracy of the FSM [44].

As a consequence, we need to approximate the distribution function at each discrete velocity point by solving  $M_1 \times M_2 \times M_3 \times K$  equations on each element  $\Delta_i$  (take the ITR-MEAN scheme (12) as an example):

$$-\left(\nabla \varphi_s, \mathbf{v}^{j'} f^{j'}\right)_{\Delta_i} + \langle \varphi_s, \hat{\mathbf{H}}^{j'} \cdot \mathbf{n} \rangle_{\partial \Delta_i} + \left(\varphi_s, \bar{\nu} f^{j'}\right)_{\Delta_i} = \left(\varphi_s, \mathcal{C}^{j'}\right)_{\Delta_i} + \left(\varphi_s, \bar{\nu} f^{j'}\right)_{\Delta_i}, \quad (23)$$

where  $f^{j'} = f(\mathbf{x}, \mathbf{v}^{j'})$ ,  $\hat{\mathbf{H}}^{j'} = \hat{\mathbf{H}}(f^{j'}, f_{\text{ext}}^{j'})$ , and  $\mathcal{C}^{j'} = \mathcal{C}(f^{j'}, f_*^{j'})$  denote the corresponding variables at each discrete velocities. The resulting governing equations can be re-written into matrix form as:

$$\mathbf{A}^{i,j'} \mathbf{F}_i^{j'} + \mathbf{B}^{\text{ext},j'} = \mathbf{S}^{i,j'}, \quad (24)$$

where  $\mathbf{F}_i^{j'} = [F_1(\mathbf{v}^{j'}), \dots, F_r(\mathbf{v}^{j'}), \dots]^T$  are the unknowns, i.e. the vector of degrees of freedom of  $f^{j'}$  on  $\Delta_i$ . Other coefficient matrices are given in the Appendix.

The strategy to solve the linear systems that are coupled through numerical fluxes over all spatial elements will be described in Sec. 5.

### 3.3. Boundary condition

At the boundary of computational domain, to determine the flux for an element  $\Delta_i$ , the distribution function obtained from the exterior of the element,  $f_{\text{ext}}$ , is described by a given boundary condition  $b^{j'}$ . In this paper, the diffuse boundary condition is used at solid surface. Suppose the solid wall moves with a constant speed  $\mathbf{u}_w$ , and has a temperature  $T_w$  that can either be a constant or vary along the wall, the distribution function for reflected molecules [i.e. when  $(\mathbf{v}^{j'} - \mathbf{u}_w) \cdot \mathbf{n}_w \leq 0$ ,  $\mathbf{n}_w$  is the outward unit normal vector of the solid surface] is given by the equilibrium distribution:

$$b^{j'} = \frac{\rho_w}{(\pi T_w)^{3/2}} \exp\left(-\frac{|\mathbf{v}^{j'} - \mathbf{u}_w|^2}{T_w}\right), \quad (25)$$

where,  $\rho_w$  is defined by:

$$\sum_{(\mathbf{v}^{j'} - \mathbf{u}_w) \cdot \mathbf{n}_w < 0} (\mathbf{v}^{j'} - \mathbf{u}_w) \cdot \mathbf{n}_w f^{j'} + \sum_{(\mathbf{v}^{j'} - \mathbf{u}_w) \cdot \mathbf{n}_w \leq 0} (\mathbf{v}^{j'} - \mathbf{u}_w) \cdot \mathbf{n}_w b^{j'} = 0 \quad (26)$$

such that the mass flux across wall is equal to zero. Implementation of other types of boundary conditions such as symmetry boundary, far-pressure inlet/outlet boundaries, and supersonic inlet/outlet boundaries can be found in Ref. [50].

## 4. Reduction of the Computational Complexity in DG Formalism

The major computational cost to solve the system (24) arises from two parts: 1) evaluating collision operator and 2) solving linear equations. In this section, we analysis the computational complexity for evaluation of the collision operator and left the one for solution of linear systems in Sec. 5. For simplification, we assume that equidistant discrete molecular velocities and frequencies are employed with  $M = N$  and  $N_1 = N_2 = N_3 = \bar{N}$ . Then, at each iterative step, equipped with the FFT-based convolution, the computational complexity is  $O(K^2 M_{\text{el}} M_{\text{qua}}^2 \bar{N}^3 \log \bar{N} + K^3 M_{\text{el}} \bar{N}^3)$ , in which the first term arises in the calculation of  $\Xi_{p,r}$  and  $\Lambda_r$  in Eq. (20), while the second term is for conducting the loops in Eqs. (21) and (22).

Now, we propose an approach to reduce the cost in evaluating Boltzmann collision operator. In the following discussion, we will omit the index of discrete molecular velocities  $j'$ . The approach may be described heuristically in the following manner. If we choose the supporting polynomials as nodal shape functions:

$$\varphi_r(\mathbf{x}_p) = \begin{cases} 0, & \text{if } r \neq p, \\ 1, & \text{if } r = p, \end{cases} \quad (27)$$

where  $\mathbf{x}_p$  is the nodal points for interpolation, the degree of freedom  $F_r$  is actually the nodal value of distribution function, say  $f(\mathbf{x}_r)$ . We assume that the distribution of  $\mathcal{C}$  within an element might as well be estimated by the nodal approximation:

$$\mathcal{C} \simeq \tilde{\mathcal{C}} = \sum_{r=1}^K \varphi_r \left( \tilde{\Xi}_r - \Lambda_r F_r \right), \quad (28)$$

where  $\tilde{\Xi}_r$  and  $\Lambda_r$  are the nodal values of the collision gain term and collision frequency, respectively. The nodal value  $\tilde{\Xi}_r$  is estimated from  $F_r$  as  $\tilde{\Xi}_r = \Xi_{r,r}$ . As a result, the computational cost of  $(\varphi_s, \mathcal{C})_{\Delta_i}$  in Eqs. (21) and (22) is reduced to  $O(K M_{el} M_{qua}^2 \bar{N}^3 \log \bar{N} + K^2 M_{el} \bar{N}^3)$ , that is, by  $K$  times; this is considerable especially when high-order approximation polynomials are employed. For instance, nominally, 14 times less cost is expected when  $k = 4$  in 2D problems on triangular mesh.

It is interesting to note that, in the recent paper where an explicit DG Boltzmann solver has been developed, the singular value decomposition is proposed to reduce the computational cost [51]. The singular value decomposition is pre-computed to the  $K \times K$  matrix for  $(\varphi_s, \varphi_p \varphi_r)$ . Thus, the computational cost for the loops in Eqs. (21) and (22) can be reduced to  $O(K^2 M_{el} \bar{N}^3)$ . However, the computational complexity for the collision operators by FSM [Eq. (20)] remains unchanged and always consumes the majority of CPU time; thus this saving may not be in the order of magnitude.

Note that the introduced error of using Eq. (28) is proportional to:

$$\left( \varphi_s, |\mathcal{C} - \tilde{\mathcal{C}}| \right)_{\Delta_i} \propto \left( \varphi_s, |\bar{f}^l \bar{f}^m - \sum_{r=1}^K \varphi_r \bar{F}_r^l \bar{F}_r^m| \right)_{\Delta_i} + \left( \varphi_s, |\bar{f}^j \bar{f}^j - \sum_{r=1}^K \varphi_r \bar{F}_r^j \bar{F}_r^j| \right)_{\Delta_i}, \quad (29)$$

which is small when the variation of distribution function within a spatial element is not significant. In Sec. 6, we are going to validate this approximation numerically. The scheme with full calculation of collision terms (21) and (22) is labeled as ‘DG-FULL’, while the one using reduced calculation (28) is labeled as ‘DG-RED’.

## 5. Sweeping technique to solve the linear systems

Now, we present the strategy to solve the linear systems resulting from the DG discretization. In the linear equations (24) on each spatial element  $\Delta_i$ , the unknown distribution function on neighboring element appears in  $\mathbf{B}^{ext,j'}$ , the usual treatment in implicit DG

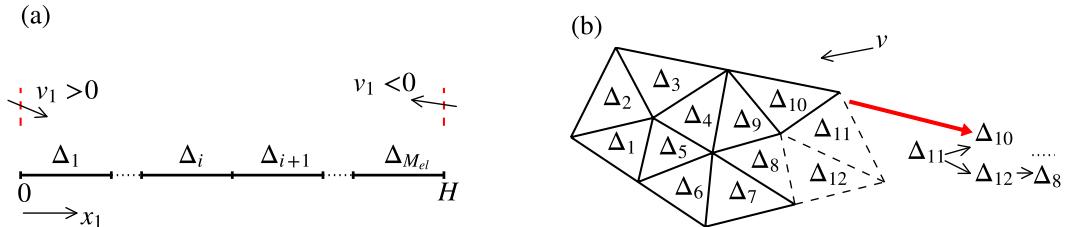


Figure 1: Schematic demonstration for determination of the spatial element ordering with respect to a given molecular velocity. (a) line mesh for 1D problem. (b) triangular mesh for 2D problem.

is assembling the linear systems over all spatial elements and solving a large sparse linear equation to determine the unknowns simultaneously. However, this requires huge memory and is prohibitively expensive in solving the Boltzmann equation, since we have to solve not one but a large number, e.g. several ten thousands, of large sparse linear systems. The matrix-free technique might be useful to improve the scheme [59]. In this paper, a more intuitive and simpler strategy is adopted.

Due to the fact that the upwind flux is applied, it is important to notice that only the distribution function on neighboring elements in the upwind side appears in  $\mathbf{B}^{\text{ext},j'}$ . Thus, the solution of  $f^{j'}$  on  $\Delta_i$  can be obtained by solving the small linear system (24), once  $f_{\text{ext}}$  on the upwind side is known, or it is equipped with prescribed boundary conditions. Hence, starting from the element at the inflow boundary of computational domain, we can obtain the solution of  $f^{j'}$  sequentially for all elements.

This sweeping technique, which requires no assembling large sparse linear system, relies on finding an ordering of the spatial elements, which is determined by the characteristic wind direction (that is, the direction of molecular velocity). The key to ensure feasible implementation of the sweeping technique in solving the Boltzmann equation is that, the discrete molecular velocity is fixed in the governing equations. Hence, we can find and store the spatial element ordering for each discrete velocity immediately after discretization and before the first iteration.

For a given discrete molecular velocity  $\mathbf{v}^{j'}$ , the topological ordering is easily found in 1D cases. As shown in Figure 1(a), a 1D computational domain  $[0, H]$  is parallel to the  $x_1$  axis. When  $v_1^{j'} > 0$ , starting from the boundary at  $x_1 = 0$ , the spatial ordering is of ascending order in index  $i$ , while when  $v_1^{j'} < 0$  the spatial ordering has a descending order in  $i$  starting from the boundary at  $x_1 = H$ . For higher dimensional problems, we assume that the spatial grid is paved with convex elements and the element ordering is acyclic. A simple topological sorting algorithm is applied: gradually removing elements that have no incoming flux from elements left in the computational domain, placing them in the ordering, until no element remains. Figure 1(b) illustrates the schematic demonstration for the sorting procedure, where the ordering starts from the element  $\Delta_{11}$  since it has only one inflow boundary located at the boundary of the computational domain. After removing  $\Delta_{11}$ , either element  $\Delta_{10}$  or  $\Delta_{12}$  will be put into the ordering, because there is no flux flowing from the elements left in the computational domain to these two elements. Note that the sequence of  $\Delta_{11}$  and  $\Delta_{12}$  in the ordering is interchangeable, since they do not share any common interface. The pseudo-code of the algorithm can be found in Ref. [60] (Algorithm 3.2.2).

In Sec. 4, we have mentioned that one of the majority consumptions in computational resources is to solve the linear systems. On the basis of the sweeping technique, if we use *LU*-fabrication-based direct solver to solve the linear equations, the computational complexity is  $O(2/3K^3M_{\text{el}}\bar{N}^3 + 2K^2M_{\text{el}}\bar{N}^3)$  since we have  $M_{\text{el}}\bar{N}^3$  systems, and each has a coefficient matrix of rank  $K$ . Note that we have assumed that the number of discrete velocities in each direction is  $\bar{N}$ . In the ITR-MEAN scheme (12), the complexity to solve linear equations can be reduced to  $O(2K^2M_{\text{el}}\bar{N}^3)$  due to the fact that the coefficient matrix  $\mathbf{A}^{i,j'}$  remains

Table 1: Flow properties across normal shock waves.

	$Ma = 2.05$		$Ma = 9.0$	
	upstream	downstream	upstream	downstream
$T$	1.0	2.144	1.0	26.185
$\rho$	1.0	2.334	1.0	3.857
$u_1$	1.871	0.802	8.216	2.130

unchanged during all iterations and  $LU$ -decomposition can be calculated and stored before the first iteration. The computational cost for  $LU$ -decomposition is roughly  $K/3$  times that for substitution in solving the linear equations, which becomes large as the grid density and/or the order of approximating polynomial increases. For example, when  $k = 4$  on triangular mesh, the computational complexity of  $LU$ -decomposition is 4 times larger than that of substitution. Therefore, completing  $LU$ -decomposition before iteration and only executing substitution during iteration can further save CPU time.

## 6. Numerical Results and Discussions

The DG method with  $k$  up to 4 is applied to solve the Boltzmann equation with full collision operator. The convergence criterion for the iterative schemes described above is that the global relative residual in the flow property  $\mathcal{Q}$  between two successive iteration steps:

$$R_{\mathcal{Q}} = \frac{|\int_{\Delta} \mathcal{Q}^{(t+1)} - \mathcal{Q}^{(t)} d\mathbf{x}|}{|\int_{\Delta} \mathcal{Q}^{(t)} d\mathbf{x}|}, \quad (30)$$

is less than a threshold value  $\epsilon$ .

The following tests are performed in double precision on a workstation with Intel Xeon-E5-2680 processors and 132 GB RAM. During iteration, we call the routines in Intel Math Kernel Library (MKL) to conduct  $LU$ -fabrication and solve linear equations. For the calculation of collision kernel  $\beta(l, m)$ , the trapezoidal rule is applied and we set  $M_{\text{qua}} = 5$  in Eq. (11) that is adequate to maintain the spectral accuracy of the FSM [43]. Due to the fact that we only consider 1D and 2D flows, symmetry of the distribution function in the third ( $v_3$ ) direction allows us to reduce the computational cost of Eq. (11) by half, that is,  $\theta$  can be limited to the range of  $[0, \pi/2]$ ; more details can be found in Ref. [44].

### 6.1. 1D normal shock wave

We first simulate the normal shock wave problem to assess the proposed method for the steady-state solution of the Boltzmann equation. Due to the absence of boundary effects, the flow is ideal to test the accuracy of DG discretization for streaming and the FSM approximation for the Boltzmann collision operator in capturing highly non-equilibrium, especially to validate the scheme with reduced DG calculation as described in Sec. 4. The argon gas is considered with Mach numbers  $Ma = 2.05$  and  $Ma = 9.0$ . We use the same parameters as those in Alsmeyer's experiments [61]: the upstream density  $\rho_0 = 1.067 \times 10^{-4} \text{ kg/m}^3$  and temperature  $T_0 = 300 \text{ K}$ , corresponding to the mean free path and collision frequency of

hard sphere molecules as  $\bar{\lambda} = 1.098 \times 10^{-3}$  m and  $\bar{\nu} = 3.633 \times 10^5$  s<sup>-1</sup>, respectively. For all the DG results, the length scale is normalized with  $H = \bar{\lambda}$  resulting in  $Kn = 5\pi/16$ . The 1D computational domains  $\Delta$  in the  $x_1$  direction are  $[-20, 20]$  and  $[-30, 30]$  for  $Ma = 2.05$  and  $Ma = 9.0$  cases, respectively, which are partitioned by line elements with uniform length. The dimensionless up/downstream conditions normalized by the upstream properties are listed in Table 1. Initially, the domains  $x_1 \leq 0$  and  $x_1 > 0$  are setup by the equilibrium distributions at upstream and downstream conditions, respectively. The implicit iteration scheme (13) with local collision frequency, i.e. ITR-LOC is applied. Iteration is terminated when  $\max\{R_T, R_n, R_{|u_1|}\} < 10^{-5}$ . When  $Ma = 2.05$ , the truncated velocity domain  $[-8, 8]^3$  is divided into  $32^3$  uniform points, while when  $Ma = 9.0$ , the velocity domain  $[-30, 30]^3$  is divided into  $96 \times 64 \times 64$  uniform points. The same number of uniform frequencies are used for approximation of the collision operator.

Numerical tests show that by using the sweeping technique, the implicit DG method is stable without any limiter in solving the 1D normal shock structure. Figure 2 illustrates the DG results of normalized flow velocity, density and temperature, compared with the DSMC results and experimental data [61]. The DSMC results presented here are computed using the code developed and verified in [62]. In order to ensure accuracy of the DSMC method, the cell sizes and time steps are set to be  $\sim 0.13\bar{\lambda}$  and  $\sim 0.12/\bar{\nu}$ , respectively. The average number of molecules per spatial cell is about 50. About 30,000 iterations are needed to reach the steady-state solutions. To obtain smooth results, macroscopic flow properties are sampled over another 100,000 steps. For comparison, the viscosity index in both methods are set as  $\omega = 0.81$ . The DG results are obtained using 4<sup>th</sup>-order approximating polynomial on 16 elements, which agree well with those of DSMC simulation (the profiles from the DG-FULL scheme are not shown, since they overlap with the ones of DG-RED). We also compare the DG solutions for density with the experimental data. For  $Ma = 2.05$ , the agreement is good, although slight discrepancy can be observed in the downstream side of the shock wave. For  $Ma = 9.0$  where the non-equilibrium effect is strong, the DG solutions agrees well with the DSMC ones. However, disagreement between the DG (DSMC) solution and experimental one enlarges, where the variation of density is steeper in experiment. Actually, the profiles in high Mach number flow are more sensitive with respect to the value of viscosity index  $\omega$ . The works in [63, 64] suggest that to set  $\omega$  being around 0.7, the Boltzmann solver or DSMC can produce result closed to the experimental one. Hence, we include the DG-RED solution with  $\omega = 0.72$  (dash lines in Figure 2(b)), and obtain an improved agreement.

To further validate the DG-RED scheme, we compare the marginal distribution functions  $\int f dv_2 dv_3$  at different locations of the shock wave with those obtained using the DG-FULL scheme. The profiles are plotted in Figure 3. To ensure accuracy of the DG-FULL results, we have doubled the number of discrete velocity and frequency points in the longitudinal direction. It is demonstrated that in low Mach number flow, the distribution functions are closed to the corresponding equilibrium (Gaussian) distribution. As Mach number increases, the distribution functions within the shock wave structure greatly deviate from the equilibrium states. The comparison shows that, even for highly non-equilibrium flow, the DG-RED scheme can produce correct solutions, so that the numerical error brought by the reduced

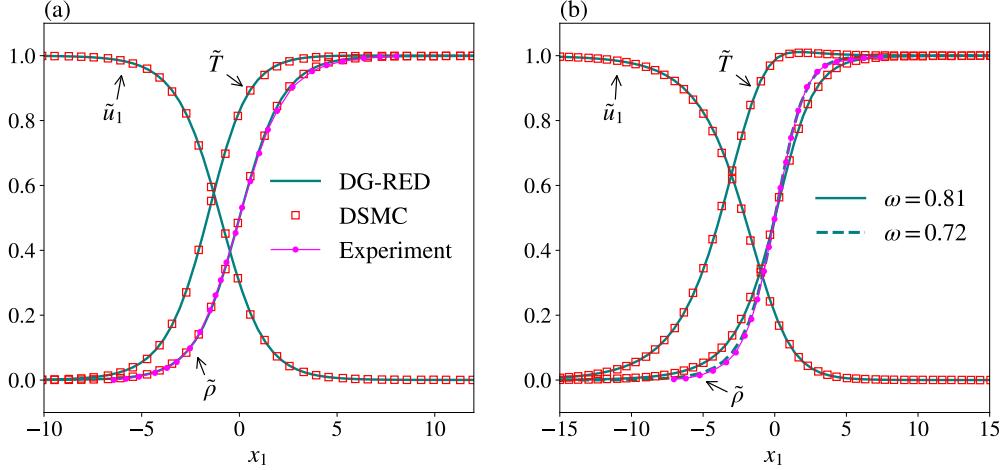


Figure 2: Profiles of normalized flow velocity  $\tilde{u}_1 = \frac{u_1 - u_{1,R}}{u_{1,L} - u_{1,R}}$ , density  $\tilde{\rho} = \frac{\rho - \rho_L}{\rho_R - \rho_L}$  and temperature  $\tilde{T} = \frac{T - T_L}{T_R - T_L}$  for normal shock wave of argon gas at (a)  $Ma = 2.05$  and (b)  $Ma = 9.0$ . The subscripts ‘L’ and ‘R’ denote the properties in upstream and downstream, respectively. The DG-RED solutions are obtained with  $k = 4$ ,  $M_{el} = 16$ . The ITR-LOC scheme (13) is applied for implicit iteration.

calculation of collision operator is negligible.

Another important property of a shock wave with  $Ma > \sqrt{9/5}$  in a monatomic gas is the overshoot of temperature associated with the longitudinal component of thermal velocities,  $T_x$ , which could be larger than the gas temperature behind the front of shock due to the non-equilibrium in translational energies of longitudinal and transversal directions. The analytical form of  $T_x$  is related to the density  $\rho$  as [65]:

$$T_{x,an} = \frac{1}{3} \left[ \frac{(5Ma^2 + 3)}{\rho} - 5 \left( \frac{Ma}{\rho} \right)^2 \right]. \quad (31)$$

Based on  $T_x$ , we compare the convergence behavior of DG-RED and DG-FULL schemes with respect to various orders of approximating polynomials  $k$  and numbers of spatial elements  $M_{el}$ . The relative  $L_2$  error of  $T_x$  that is evaluated as

$$\mathcal{E} = \frac{\int_{\Delta} (T_x - T_{x,an})^2 dx_1}{\int_{\Delta} T_{x,an}^2 dx_1}, \quad (32)$$

the number of iteration steps and the total CPU time are listed in Table 2.

All tests are done on single processor, and the internal parallelism for MKL functions is not activated. It is shown that for each  $k$ , as the number of elements increases, errors of  $T_x$  gradually converges to 0.016% and 0.036% for Mach numbers of 2.05 and 9.0, respectively. The higher order approximating polynomials, the fewer elements needed to obtain the converged results. The numbers of iterative steps to reach the steady-state solutions also converge to fixed values of around 201 and 225 for Mach numbers of 2.05 and 9.0, respectively. Therefore, compared to the lower-order scheme, the higher-order discretization consumes less CPU time to obtain solution with the same order of accuracy. For example,

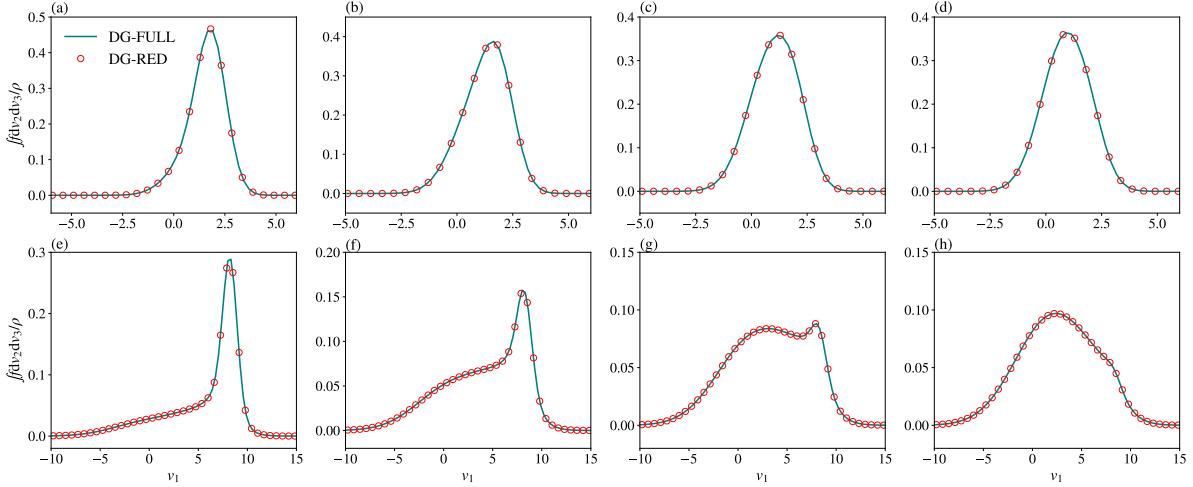


Figure 3: Comparison of the marginal distribution functions  $\int f dv_2 dv_3 / \rho$  from the DG-RED and the DG-FULL schemes: first row is ones for  $Ma = 2.05$  presented at (a)  $\rho = 1.197$ , (b)  $\rho = 1.468$ , (c)  $\rho = 1.766$  and (d)  $\rho = 1.991$ ; second row is ones for  $Ma = 9.0$  presented at (e)  $\rho = 1.423$ , (f)  $\rho = 1.943$ , (g)  $\rho = 2.543$  and (h)  $\rho = 3.123$ . For the DG-FULL, the molecular velocity domains  $[-8, 8]^3$  and  $[-30, 30]^3$  are uniformly divided into  $64 \times 32 \times 32$  and  $192 \times 64 \times 64$  points for  $Ma = 2.05$  and  $9.0$ , respectively. The ITR-LOC scheme (13) is applied for implicit iteration.

for  $Ma = 2.05$ , the DG-FULL scheme with  $k = 4$  cost about 30% less CPU time to produce solution with  $\mathcal{E} = 0.016\%$  on the mesh of 16 segments, compared to the one with  $k = 3$  that obtains the same accurate result on 32 segments.

It is found that the DG-RED scheme can preserve these convergence properties. That is, by using the same order of approximating polynomials on the same mesh, the DG-RED and DG-FULL require the same number of iterative step to obtain solutions of the same order of accuracy. However, the DG-RED can significantly save the computational cost in terms of CPU time. The higher degree of approximating polynomials, the more the saving. For example, for  $Ma = 9.0$ , to obtained solution of  $\mathcal{E} = 0.036\%$ , both the schemes need 64, 32 and 16 spatial elements for  $k = 2, 3$  and 4, respectively, and the CPU time consumed by the DG-RED is about 50%, 41%, and 36% of that by the DG-FULL.

## 6.2. Hypersonic flow past a square cylinder

Now we consider a 2D high-speed flow. the DG-RED scheme of  $k = 4$  is applied to compute hypersonic flow past a square cylinder having a dimension of  $1 \times 1$  and a constant wall temperature of  $T_w = 1.0$ . The free stream has dimensionless temperature and density of  $T_0 = 1.0$  and  $\rho_0 = 1.0$ . The Mach number and Knudsen number in the free stream are set as 5.0 and 0.13, respectively. As shown in Figure 4(a), the argon gas of viscosity index  $\omega = 0.81$  moves from left to right along the  $x_1$  direction. The computational domain is chosen with extension up to 1.95, 7.5 and 5.5 away from the cylinder in the upwind, downstream and  $x_2$  direction, respectively. Due to symmetry, only half of the flow field is considered. The boundary conditions and triangular mesh are also illustrated in Figure 4(a). Besides the full-diffuse solid surfaces, the lower boundaries parallel to  $x_1$  is

Table 2: Comparisons between the DG-FULL (with FULL calculation of  $(\varphi_s, \mathcal{C})_{\Delta_i}$ ) and the DG-RED (with reduced calculation of  $(\varphi_s, \mathcal{C})_{\Delta_i}$ ) in terms of the relative  $L_2$  error ( $\mathcal{E}$ ) of longitudinal temperature  $T_x$  (compared with the analytical result), the number of iterations (Itr denotes the number of iteration steps to reach the convergence criterion  $\max\{R_T, R_\rho, R_{|u_1|\}} < 10^{-5}$ ), and the CPU time  $t_c$ . Normal shock wave is considered. The ITR-LOC scheme (13) is applied for implicit iteration.

k	Ma = 2.05						Ma = 9.0							
	DG-FULL			DG-RED			DG-FULL			DG-RED				
	$M_{el}$	$\mathcal{E} \times 10^{-2}$	Itr	$t_c, [h]$	$\mathcal{E} \times 10^{-2}$	Itr	$M_{el}$	$\mathcal{E} \times 10^{-2}$	Itr	$t_c, [h]$	$\mathcal{E} \times 10^{-2}$	Itr		
1	4	14.656	208*	0.02	7.734	204**	0.01	8	12.91	201	0.64	5.814	278	0.61
	8	3.450	194	0.03	2.159	219	0.02	16	1.668	222	1.32	1.553	252	1.08
	16	0.643	199	0.06	0.622	207	0.03	32	0.396	225	2.57	0.381	238	1.96
	32	0.158	201	0.11	0.152	203	0.06	64	0.088	225	5.19	0.087	229	3.72
2	4	3.343	189	0.03	1.629	203	0.01	8	0.824	225	1.39	0.361	231	0.75
	8	0.202	199	0.06	0.187	198	0.02	16	0.239	225	2.73	0.247	226	1.49
	16	0.097	200	0.13	0.097	200	0.04	32	0.049	225	5.75	0.049	225	2.86
	32	0.020	201	0.24	0.020	201	0.08	64	0.036	225	11.24	0.036	225	5.52
3	4	0.221	190	0.05	0.350	201	0.02	8	0.550	225	2.58	0.595	235	1.06
	8	0.216	198	0.09	0.219	199	0.03	16	0.066	225	5.06	0.065	226	1.98
	16	0.022	200	0.21	0.022	200	0.06	32	0.036	225	9.38	0.036	225	3.88
	32	0.016	201	0.41	0.016	201	0.11	64	0.036	225	16.22	0.036	225	7.49
4	4	0.557	189	0.09	0.584	188	0.02	8	0.236	225	3.89	0.233	226	1.30
	8	0.060	198	0.15	0.061	198	0.04	16	0.037	225	7.35	0.037	225	2.53
	16	0.016	200	0.30	0.016	201	0.07	32	0.036	225	13.15	0.036	225	4.85
	32	0.016	202	0.59	0.016	201	0.14	64	0.036	225	25.85	0.036	225	9.55

Due to round-off errors, \*this case only converged to residual of about  $2 \times 10^{-5}$  and \*\*this case only converged to residual of about  $3 \times 10^{-5}$ .

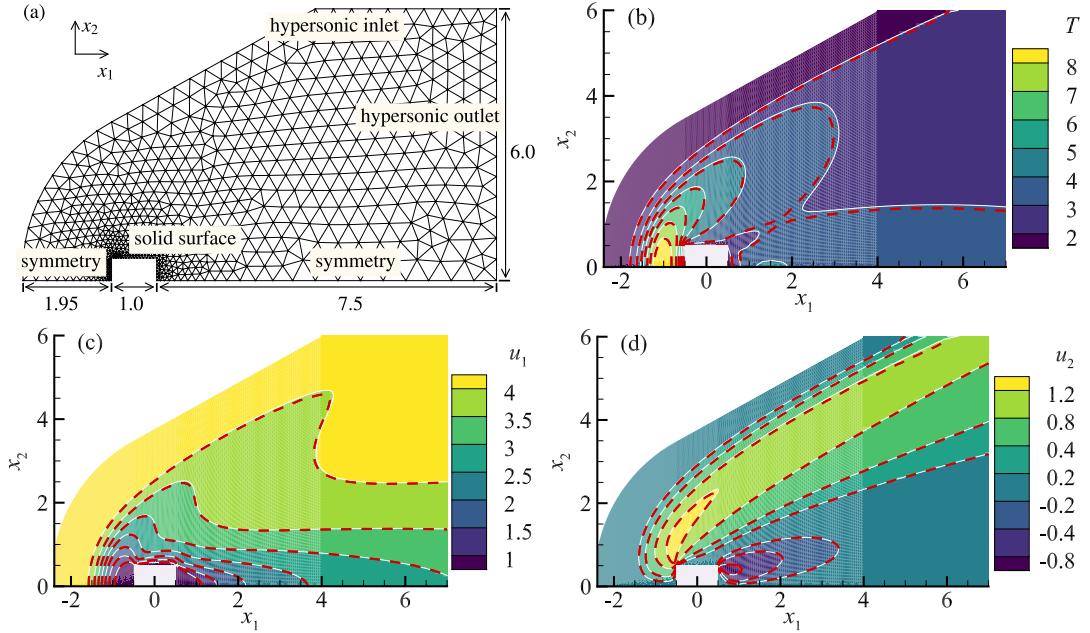


Figure 4: Hypersonic flow of  $Ma = 5$  and  $Kn = 0.13$  past a square cylinder. (a) Schematic for computational domain, boundary condition and unstructured triangular mesh. (b) Temperature contours. (c) Horizontal velocity contours. (d) Vertical velocity contours. The white solid contour lines with background illustrate the solution from the DG-RED of  $k = 4$ , where the molecular velocity domain  $[-13, 13]^3$  is discretized by  $48 \times 48 \times 48$  uniform grid points. The red dashed contour lines are the DSMC results in Ref. [66].

symmetric boundaries, while other boundaries are set as hypersonic inlet/outlet boundaries where the distribution function is the equilibrium distribution at free-stream condition. 1490 unstructured triangles are employed to discretize the computational domain, with refinement near the solid surfaces. The truncated molecular velocity space  $[-13, 13]^3$  are discretized by  $48 \times 48 \times 48$  uniform points, and the same number of uniform frequencies are used for evaluation of collision terms. The flow field is initialized by the free-stream condition, and the ITR-LOC scheme (13) is applied which costs about 346 steps to reach the convergence criterion of  $\max\{R_T, R_\rho, R_{|\mathbf{u}|}\} < 5 \times 10^{-5}$ . The test is run on 28 processors using OpenMP for parallelism and consumes about 24.6 hours of wall time.

At the very beginning of iteration, strong discontinuity appears in the upwind side of the square cylinder due intense stagnation effect of gas flow, and the DG scheme can generate spurious oscillation which may make the approximated distribution functions negative. As a consequence, the loss term will become the gain term and the iteration will lead to unphysical blowup solutions. To tackle this problem, instead of using any nonlinear limiters as one usually does, we take absolute values to the negative degrees of freedom after solving the linear systems at each iterative step. Numerical test shows that this simple treatment does not destroy accuracy of the DG discretization but does guarantee its stability.

Contours of temperature, horizontal velocity and vertical velocity are illustrated in Figure 4(b)-(d). The white lines with background are the DG-RED solution, while the red dashed contour lines are the DSMC results in Ref. [66]. Note that the Knudsen num-

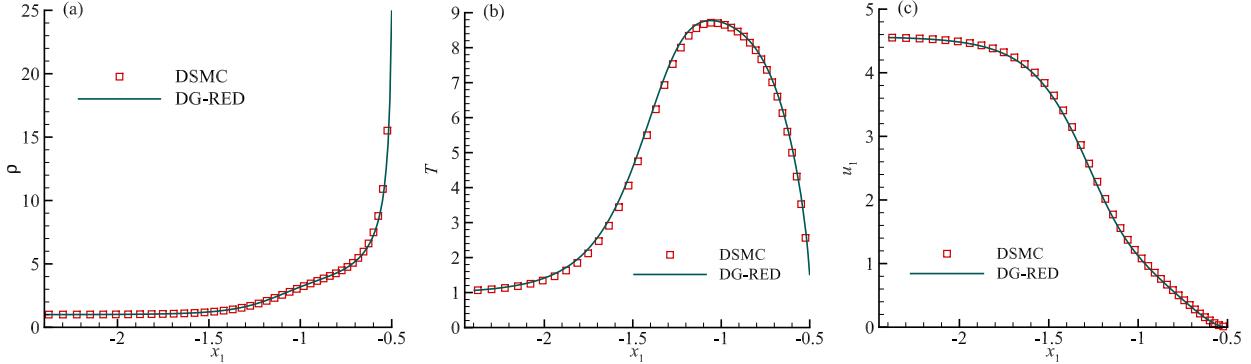


Figure 5: Hypersonic flow of  $Ma = 5$  and  $Kn = 0.13$  past a square cylinder. Profiles of (a) density (b) temperature and (c) horizontal velocity along the symmetric line in front of the stagnation point. Solid lines are the solutions from the DG-RED of  $k = 4$ , where the molecular velocity domain  $[-13, 13]^3$  is discretized by  $48 \times 48 \times 48$  uniform grid points. Symbols are the DSMC results in Ref. [66].

ber in [66] is  $2(7 - 2\omega)(5 - 2\omega)/15\pi$  times the unconfined Knudsen number in this paper. Comparison between the DG-RED solutions and DSMC ones on the distributions of density, temperature and horizontal velocity along the symmetric line in the front of the stagnation point are shown in Figure 5. It is found that due to the stagnation effect from the static cylinder to the gas flow, the flow density increases about 25 times within 10 (free-streaming) mean free paths when approaching to the cylinder, and the bulk horizontal velocity drops to zero. Since the isothermal wall condition is applied, the flow temperature first increases to its maximum value of 8.7 at about 5 mean free paths away from the stagnation point and then decreases to 1.45 at the solid wall.

Figure 6 illustrates the distributions of normal stress  $P_n$  and shear stress  $P_t$  along the surfaces of the square cylinder, where  $P_n = \mathbf{n}_w \cdot \mathbf{P} \cdot \mathbf{n}_w$  and  $P_t = \mathbf{t}_w \cdot \mathbf{P} \cdot \mathbf{t}_w$  with  $\mathbf{n}_w$  and  $\mathbf{t}_w$  denoting the outward unite normal vector and tangential vector of the solid surface, respectively. The largest  $P_n$  is at the surface in the upwind side where the normal momentum flux is large, while the shear stress gradually increases along that surface as the bulk vertical velocity increases. Both the  $P_n$  and  $P_t$  vary slightly along the top surface and the lateral surface in the weak. Figures 4 to 6 demonstrate the good agreement between the DG and DSMC results.

### 6.3. 2D lid-driven cavity flow

By comparing with the DSMC results, a 2D low-speed flow in a square cavity driven by the top lid is used to compare performances of the DG Boltzmann solvers and a Boltzmann solver using the second-order FDM to approximate derivatives in the spatial space [45]. The wall temperature is set as the reference temperature  $T_0 = 273$  K. The velocity of the driven lid is 50 m/s. The flow gas is argon with a viscosity index of 0.81. The gas flow is initialized to be rest at  $T_0$  with  $Kn = 1$ , where the characteristic length  $H$  is chosen to be the side length of the square cavity. The computational configuration for DSMC can be found in [67].

For deterministic solutions, the truncated molecular domain is selected as  $[-6, 6]^3$ . The DG and FDM solvers utilize the same FSM to evaluate collision terms in frequency domain,

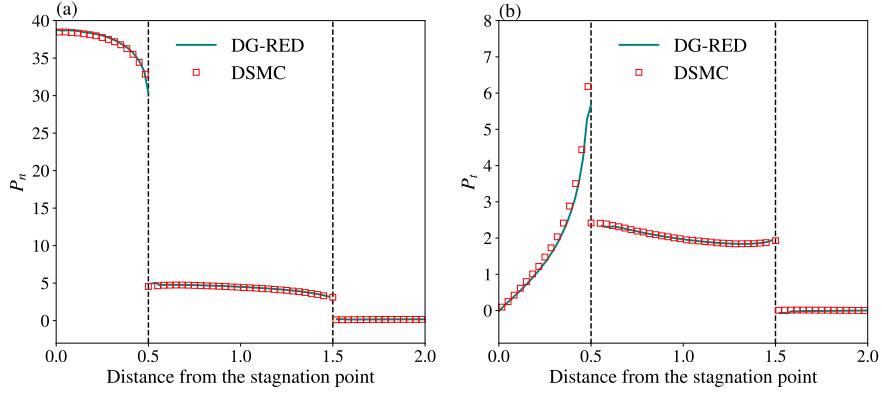


Figure 6: Hypersonic flow of  $Ma = 5$  and  $Kn = 0.13$  past a square cylinder. The distributions of (a) normal stress and (b) shear stress along the surface of cylinder. The horizontal axis represents the distance along the surface of the square, starting from the stagnation point in a counter-clockwise direction. Solid lines are the DG-RED solutions and symbols are the DSMC results in Ref. [66].

which is discretized with  $32 \times 32 \times 24$  equidistant frequencies. For discretization in the molecular velocity, non-uniform points are used for  $v_1$  and  $v_2$ , while uniform discrete velocities are used in the third direction. The non-uniform discretization with refinement around  $v_{1(2)} = 0$  is efficient to calculate low-speed flows especially at large Knudsen numbers, where the distribution function changes rapidly within a narrow area around the origin in the  $v_1$  and  $v_2$  directions [58]. For spatial discretization, uniform triangular mesh is used in the DG method, as shown in Figure 7(a), while the FDM uses equidistant grid points in the  $x_1$  and  $x_2$  directions. Determination on the numbers of spatial elements and discrete velocities is a trivial task. General speaking, flows with small values of  $Kn$  need relatively large number of spatial elements to ensure that the artificial diffusion is much smaller than the physical viscosity that is small in near-continuum flows, while highly rarefied flows require a large number of discrete velocities to resolve significant variations and/or discontinuities in the distribution function. Moreover, the spatial and velocity grids have ‘contrary’ effects, where finite discretization of the velocity space tends to capture discontinuities, whereas limited spatial discretization tends to smooth flow field due to artificial diffusion. Incompatible spatial and velocity grids can lead to emergence of the so called ‘ray effect’, which causes deterministic solution oscillating around its mean value [68, 69]. To overcome this shortcoming, the velocity grid should be fine enough so that error induced by the ray effect is small, which can be compensated by the error of numerical diffusion [69].

Temperature contours from the DG-RED for  $k = 4$  and  $M_{el} = 72$  (highly resolved in the spatial space) are compared with the DSMC results in Figure 7. Results in Figure 7(b) and (c) are obtained with  $36 \times 36 \times 24$  and  $108 \times 108 \times 24$  velocities, respectively. It is observed that relative coarser velocity grid produces temperature contour with violent fluctuations, and refinement in the velocity discretization can largely improve the accuracy. Besides, DG solver with higher order of approximating polynomial is more likely to suffer the ray effect. This is mainly due to the fact that, compared to lower-order scheme, higher-order scheme can obtain more accurate result on same spatial grid so that the numerical diffusion

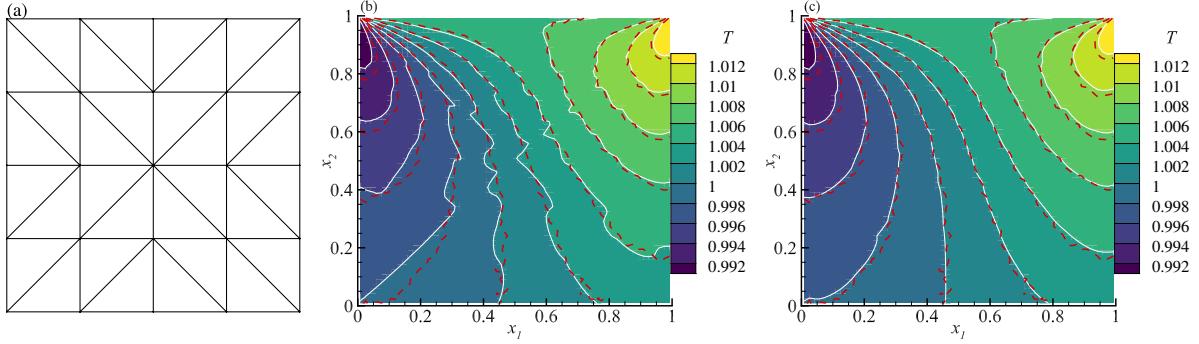


Figure 7: Comparison of the DG-RED and the DSMC on square cavity flow at  $Kn = 1.0$  driven by moving lid with speed of  $U_0 = 0.148$ . The DG-RED solutions are obtain with  $k = 4$  on 72 uniform triangles. (a) typical triangular mesh; (b) and (c): temperature contours when the molecular velocity domain  $[-6, 6]^3$  is discretized by  $36 \times 36 \times 24$  and  $108 \times 108 \times 24$  grid points in the DG-RED, respectively. White solid lines with background indicate solution of the DG-RED, while red dashed lines are the DSMC result. The ITR-LOC scheme (13) is applied for implicit iteration.

is relatively smaller which can not smear the ray effect.

Further comparison on the results of DG-RED and DSMC are illustrated in Figure 8 in terms of horizontal (vertical) flow velocity along selected vertical (horizontal) lines. The DG-RED results possess good agreement with those of the DSMC.

In Table 3, we list the relative  $L_2$  error of velocity magnitude  $|\mathbf{u}|$ , the number of iterations to reach the convergence criterion  $\max\{R_T, R_n, R_u\} < 10^{-5}$ , as well as the total CPU time cost for the DG-FULL with the ITR-LOC, the DG-RED with the ITR-LOC and the DG-RED with the ITR-MEAN. In the ITR-MEAN iterative scheme (12), the mean collision frequency is set as  $\bar{\nu} = 1.4$  for this specific flow. For all cases, the molecular velocity domain  $[-6, 6]^3$  is discretized by  $72 \times 72 \times 24$  grid points. The errors are calculated in reference to the DSMC results, which are obtained at  $60 \times 60$  equidistant points  $\mathbf{x}_p$  in the computational domain. The errors are evaluated as

$$\mathcal{E} = \sqrt{\sum (|\mathbf{u}(\mathbf{x}_p)|_{\text{DG}} - |\mathbf{u}(\mathbf{x}_p)|_{\text{DSMC}})^2 / \sum |\mathbf{u}(\mathbf{x}_p)|_{\text{DSMC}}^2}, \quad (33)$$

where the DG solution at any point  $\mathbf{x}_p$  can be easily obtained through polynomial approximation. All tests are done on single processor. It is shown that, for each order of approximating polynomials, the three schemes can produce solution with the same accuracy on the same spatial mesh. The iterative scheme using local collision frequency can obtain the steady-state solution within 21 steps, no matter which DG calculation (DG-FULL or DG-RED) is applied. Thus, due to the reduction of computational complexity in calculation of the Boltzmann collision operator, the DG-RED cost less CPU time than the DG-FULL. Equipped with the chosen mean collision frequency, the ITR-MEAN iterative scheme (12) uses 17 steps to reach the steady-state solution. Since it does not require  $LU$ -decomposition during iterations, scheme combining the DG-RED and the ITR-MEAN can further reduce the computational cost. For example, to obtain solution of error in velocity magnitude equal to 0.014 with  $k = 4$  and  $M_{\text{el}} = 18$ , the DG-RED plus the ITR-MEAN costs about 50% and

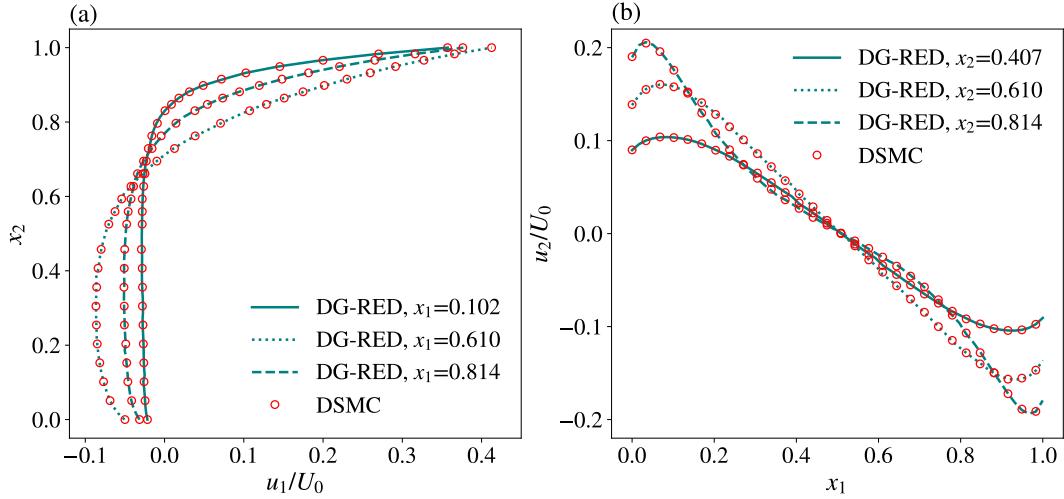


Figure 8: Comparison of the DG-RED and the DSMC on square cavity flow at  $Kn = 1.0$  driven by moving lid with speed of  $U_0 = 0.148$ . (a) normalized horizontal flow velocity  $u_1/U_0$  along vertical lines at different locations; (b) normalized vertical flow velocity  $u_2/U_0$  along horizontal lines at different locations. The DG-RED solutions are obtained with  $k = 4$  and 72 triangles. The molecular velocity domain  $[-6, 6]^3$  is discretized by  $108 \times 108 \times 24$  grid points. The ITR-LOC scheme (13) is applied for implicit iteration.

92% less CPU time than that of the DG-RED with the ITR-LOC and the DG-FULL with the ITR-LOC, respectively.

We also list the error of velocity magnitude, the number of iterations and the CPU time for the FDM in Table 4. Uniformly distributed points are employed to discretize the spatial space. Thus, the computational domain is partitioned by rectangular elements and flow properties are evaluated at the vertices of rectangles. To estimate the error of velocity magnitude,  $\mathbf{u}(\mathbf{x}_p)$  may not associated to a discrete grid point, then it is obtained through linear interpolation using the four values at vertices of the grid cell that  $\mathbf{x}_p$  locates in. The FDM solver also uses 21 steps to obtain steady-state solutions, since the ITR-LOC iterative scheme (13) is employed. For comparison of the DG and the FDM, we find that the DG discretization is more efficient. For instance, the FDM predicts solution with error in  $|\mathbf{u}|$  of 0.015 on the spatial grid with  $71 \times 71$  grid points, while the DG scheme achieves solution with the same order of accuracy on 50 and 18 triangles for  $k = 3$  and 4, respectively. However, the DG method with  $k = 3$  and full calculation in collision terms cost more CPU time than the FDM. This is because, although the computational complex for the Boltzmann collision operator in the DG-FULL with  $k = 3$  and  $M_{\text{el}} = 50$  ( $\propto M_{\text{el}} K^2$ ) and in the FDM with  $M_p = 71 \times 71$  ( $\propto M_p$ ) is similar, the DG scheme requires additional time to solve linear equations. As a consequence, only the DG-RED scheme can preserve the efficiency of DG in terms of CPU time. Equipped with the ITR-LOC iteration (13), to obtain solution with error in  $|\mathbf{u}|$  of 0.015, the DG-RED solvers of  $k = 3$  and 4 are about 4 and 7 times faster than the FDM. The ITR-MEAN iteration (12) can further boost its efficiency, now the DG-RED solvers of  $k = 3$  and 4 can be 6 and 13 times faster than the FDM. Although higher-order FDM could achieve better efficiency, it needs much more computational effort

Table 3: Comparisons between the DG-FULL with the ITR-LOC iteration, and the DG-RED with the ITR-LOC as well as the ITR-MEAN in the lid-driven square cavity flow with  $Kn = 1.0$ , in terms of the relative  $L_2$  error  $\mathcal{E}$  (33), the number of iterations (Itr denotes the number of iteration steps to reach the convergence criterion  $\max\{R_T, R_\rho, R_{|\mathbf{u}|}\} < 10^{-5}$ ), and the CPU time  $t_c$ . The molecular velocity domain  $[-6, 6]^3$  is discretized by  $72 \times 72 \times 24$  grid points.

$k$	$M_{el}$	DG-FULL + ITR-LOC			DG-RED + ITR-LOC			DG-RED + ITR-MEAN		
		$\mathcal{E}$	Itr	$t_c, [h]$	$\mathcal{E}$	Itr	$t_c, [h]$	$\mathcal{E}$	Itr	$t_c, [h]$
$1$	32	0.102	21	0.044	0.102	21	0.024	0.102	17	0.018
	50	0.080	21	0.068	0.080	21	0.038	0.080	17	0.029
	72	0.065	21	0.100	0.065	21	0.056	0.065	17	0.042
	98	0.054	21	0.258	0.054	21	0.077	0.054	17	0.059
$2$	32	0.039	21	0.154	0.039	21	0.058	0.039	17	0.043
	50	0.029	21	0.244	0.029	21	0.080	0.029	17	0.059
	72	0.023	21	0.457	0.023	21	0.125	0.023	17	0.088
	98	0.019	21	0.895	0.019	21	0.173	0.019	17	0.126
$3$	18	0.025	21	0.222	0.025	21	0.056	0.025	17	0.033
	32	0.019	21	0.551	0.019	21	0.120	0.019	17	0.082
	50	0.014	21	0.950	0.014	21	0.180	0.014	17	0.117
	72	0.012	21	1.078	0.012	21	0.272	0.012	17	0.190
$4$	8	0.024	21	0.219	0.024	21	0.044	0.024	17	0.023
	18	0.014	21	0.696	0.014	21	0.111	0.014	17	0.057
	32	0.011	21	1.321	0.011	21	0.234	0.011	17	0.143
	50	0.008	21	2.013	0.008	21	0.349	0.008	17	0.224

since stencils involving large numbers of points are required. Also, it has difficulty to treat complex geometries.

#### 6.4. 2D flow induced by a hot micro-beam in a rectangular chamber

We then consider the performance of the DG method in simulation of low-speed rarefied gas flow inside micro-channel. As depicted in Fig. 9(a), we consider a 2D rarefied gas flow induced by a hot micro-beam with a thickness of  $2 \mu\text{m}$  and a width of  $4 \mu\text{m}$ , which is encompassed in a cold rectangular chamber with a dimension of  $10 \times 8 \mu\text{m}^2$  and a wall temperature of 500 K. The beam with a temperature of 300 K is placed  $1 \mu\text{m}$  away from the left and bottom walls of the enclosure. Gas is filled between the beam and chamber. Unlike the continuum flow where the flow velocity is zero and the temperature is governed by the Fourier's heat conduction law, at rarefied conditions, the temperature inhomogeneity induces anisotropic momentum transfer that in turn produces pressure gradient and bulk gas flow. Due to the asymmetric geometry, momentum fluxes impinging on the beam surface are unbalanced, giving rise to a net Knudsen force [70, 71], which can be exploited for microstructure actuation and gas sensing [72]. Previous researches have shown that the thermal edge flow occurring near the boundary with sharp curvatures plays a critical roles in the formation of Knudsen force [73, 74].

Table 4: Performance of the FDM combining the ITR-LOC iteration (13) for solution of lid-driven square cavity flow at  $Kn = 1.0$ .  $M_p$  is the number of discrete points in the spatial space.  $\mathcal{E}$  is the relative  $L_2$  error of velocity magnitude  $|\mathbf{u}|$  compared with the DSMC results. Itr denotes the number of iteration steps to reach the convergence criterion  $\max\{R_T, R_\rho, R_{|\mathbf{u}|}\} < 10^{-5}$ .  $t_c$  is the total CPU time. The molecular velocity domain  $[-6, 6]^3$  is discretized by  $72 \times 72 \times 24$  non-uniform grid points.

$M_p$	$\mathcal{E}$	Itr	$t_c, [h]$	$M_p$	$\mathcal{E}$	Itr	$t_c, [h]$
$31^2$	0.052	21*	0.159	$61^2$	0.018	21	0.611
$41^2$	0.046	21	0.282	$71^2$	0.015	21	0.845
$51^2$	0.028	21	0.433	$81^2$	0.016	21	1.065

\*This case only converged to residual of about  $1.2 \times 10^{-5}$  due to round-off errors.

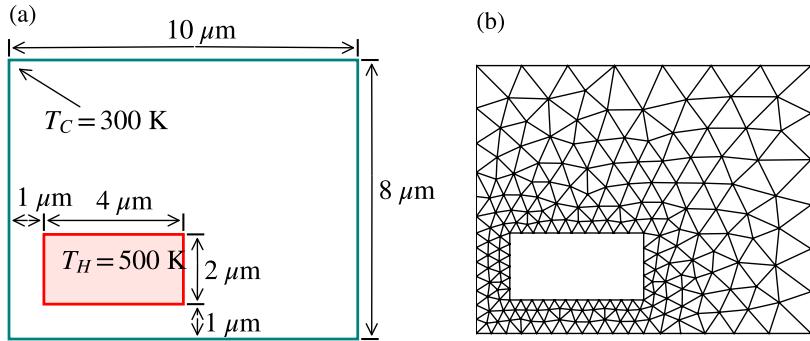


Figure 9: The micro gas flow around heated beam in a rectangular chamber. (a) geometry and (b) schematic of the triangular mesh.

The DG-RED with  $k = 4$  is applied to solve the rarefied gas flows using the ITR-LOC scheme (13). The truncated molecular velocity is set as  $[-6, 6]^3$ . 96 non-uniform velocity points are used to discretize  $v_1$  and  $v_2$ , while 24 uniform points are used for  $v_3$ . For evaluation of collision terms,  $32 \times 32 \times 24$  equidistant frequencies are employed. Fig. 9(b) illustrates the schematic of the unstructured triangular mesh, where more triangles are placed near the micro-beam. We first consider flows at  $Kn = 0.13, 1.30$  and  $12.96$ . The Knudsen numbers are calculated using  $T_0 = 400$  K and  $H = 1$  μm. The total iterative steps and the CPU time to obtain the steady-state solutions vary for flows. For the same spatial and velocity discretization, the smaller the Knudsen number, the more iterative steps thus more CPU time are required. To obtain the solution of  $Kn = 12.96$  on 881 triangles, 84 steps are needed to reach the convergence criterion of  $\max\{R_T, R_\rho, R_{|\mathbf{u}|}\} < 10^{-5}$ , which cost 4.3 hours on 12 processors (OpenMP for parallelism).

Figure 10 shows the temperature contours and streamlines. It is observed that noticeable curls that originate at the corners of the beam emerge in the temperature contour lines at highly rarefied condition ( $Kn = 12.96$ ). However, in small Knudsen number flow, sufficient intermolecular collisions gradually smooth these curls when they propagate to the chamber. When the Knudsen number is small, at each surface of the beam, the thermal edge flows drive gas molecules from the corners to the surface centers and form a relatively high pressure region therein. Then, the high pressure results in the appearance of Poiseuille flows that

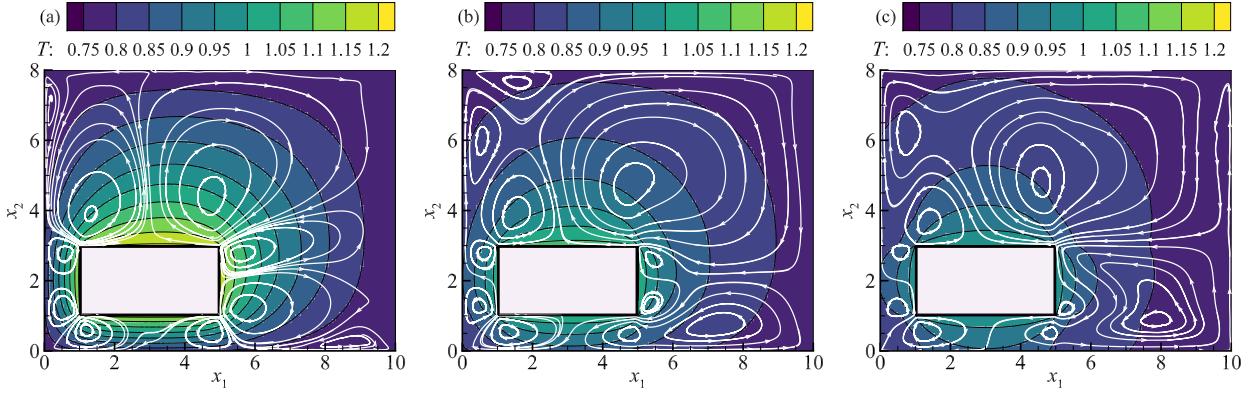


Figure 10: Temperature contours and streamlines in micro flow of argon gas with  $\omega = 0.81$ . (a)  $Kn = 0.13$ ; (b)  $Kn = 1.30$ ; (c)  $Kn = 12.96$ . The molecular velocity domain  $[-6, 6]^3$  is discretized by  $96 \times 96 \times 24$  non-uniform grid points. 1290 triangles are used for flows of  $Kn = 0.13$  and  $Kn = 1.30$ , while 881 triangles for case of  $Kn = 12.96$ .

promote gas flowing to the chamber. Due to the confinement of chamber walls, gas molecules finally return to the corners of the beam. Hence, eight localized vortices are observed in the flow field. When  $Kn$  increases to 1.30, 3 more vortices are developed with one in the lower-right corner of the chamber and two in the upper-left corner of the chamber. As the degree of rarefaction further increases, the vortex in the lower-right corner of the chamber gradually dissolves the localized vortices near the right and bottom sides of the beam, which forms a large counter-clockwise vortex. Besides, the vortices in the region above the beam also start to melt together.

Figure 11 illustrates the normal stress (pressure)  $P_n$  and the magnitude of heat flux  $|\mathbf{Q}|$  distributed on the surfaces of the hot beam, where  $P_n$  is calculated as  $P_n = \mathbf{n}_w \cdot \mathbf{P} \cdot \mathbf{n}_w$  with  $\mathbf{n}_w$  denoting the outward unit normal vector of the beam surfaces. The DSMC solutions in Ref. [75] are included for comparison, where good agreement can be observed. It can be seen that the more rarefied flow the larger  $P_n$ . This is due to the fact that momentum fluxes are enhanced when fewer intermolecular collisions are involved. Moreover, heat transfer is also strengthened by the non-equilibrium effect. The unbalance of  $P_n$  on the surfaces mainly contributes to arising of the Knudsen force (the shear stress component is smaller than the normal one by two orders of magnitude). It is observed that  $P_n$  on the top (right) surface of the beam is greater than that on the bottom (left) surface, thus both the horizontal and vertical components of the resultant force point to the negative directions of axes.

The resultant force  $\mathcal{F}$  acting on the hot beam and total heat  $\mathcal{H}$  releasing from the hot beam at  $Kn$  ranging from 0.2 to 10 are plotted in Fig. 12, where results for argon molecules with  $\omega = 0.81$ , hard-sphere molecules with  $\omega = 0.5$  and Maxwell molecules with  $\omega = 1.0$  are compared. The force and heat are calculated from integration as

$$[\mathcal{F}_{x_1}, \mathcal{F}_{x_2}]^T = - \oint_{\partial\Delta_b} \mathbf{P} \cdot \mathbf{n}_w d\Upsilon, \quad \mathcal{H} = \oint_{\partial\Delta_b} \mathbf{Q} d\Upsilon, \quad (34)$$

where,  $\partial\Delta_b$  represents the surfaces of the beam. It is observed that the magnitude of Knud-

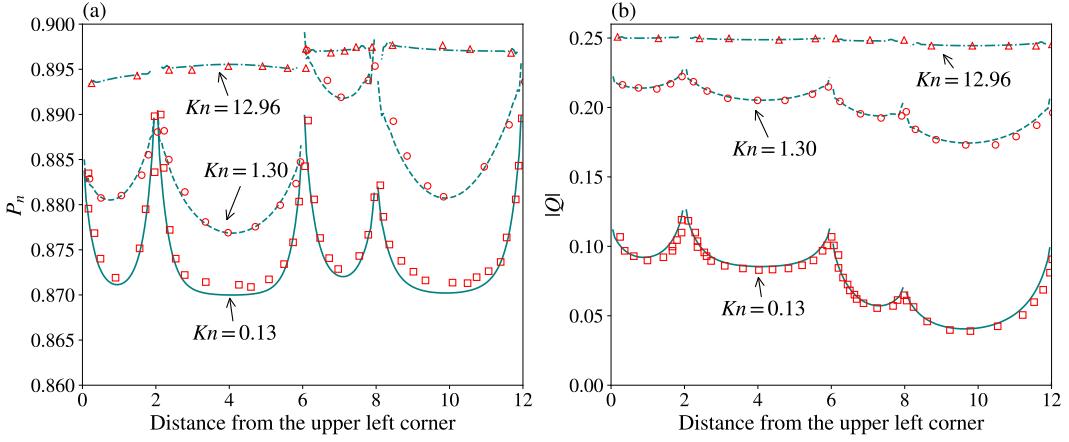


Figure 11: (a) normal stress and (b) magnitude of heat flux along the surface of hot beam. Lines are the DG-RED solutions and symbols are the DSMC result in Ref. [75]. The horizontal axis represents the distance starting from the left-upper corner in a counter-clockwise direction. The gas is argon with  $\omega = 0.81$ .

sen force first rises and then falls against the Knudsen number. The maximum magnitude of Knudsen force occurs around  $Kn = 2.0$ . The total heat always increases with increasing  $Kn$ . The variation of Knudsen force can be ascribed to the development and competition of the localized thermal flows described above. When  $Kn$  is small, i.e. the non-equilibrium effect is light, the variation of pressure on each beam surface is small and about the same magnitude, hence the Knudsen force is weak. As the Knudsen number increases, the strength of local flows are enhanced, and the more spacious spaces on the top and right of the beam allow formations of bigger vortices, which drive more gas molecules from the upper- and lower-right corners of the chamber to the center of the right surface of the beam, causing the pressure there to be larger than that near the left beam surface. On the other hand, the counter-clockwise vortex originating from the lower-right corner of the chamber penetrates into the bottom of the beam and efficiently takes gas molecules away from there. This causes the pressure near the bottom surface of the beam to be lower than that on its top surface. Therefore, the magnitudes of the horizontal and vertical components of Knudsen force both become larger. As the Knudsen number further increases, the thermal flows are further strengthened. The large vortex on the top surface of the beam starts to swallow the small vortices near the upper-left corner of the chamber, while the large vortex at the lower-right corner of the chamber begins to dissolve the small vortices on the right surface of the beam. The formations of two giant vortices release some pressure on the top and right surfaces of the beam, thus the magnitude of Knudsen force falls down. The profiles of  $\mathcal{F}$  and  $\mathcal{H}$  for  $\omega = 0.81$  always lie between the ones for  $\omega = 0.5$  and  $\omega = 1$ .

### 6.5. 2D thermal cavity flow

We also test the thermal cavity flow induced by temperature gradients at wall, which was recently used to verify an explicit DG Boltzmann solver by comparing with DSMC results [51]. In this section, we intend to provide accurate results for this flow that may serve as benchmark solutions, when the Knudsen numbers are  $Kn = 0.1, 0.5$  and  $1$ .

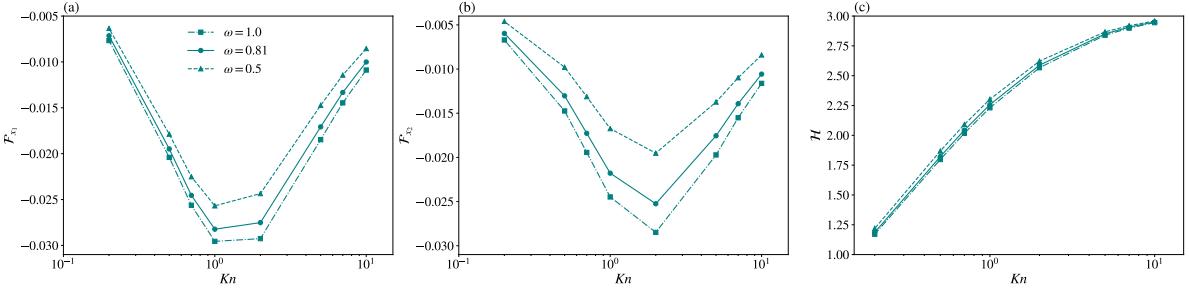


Figure 12: Resultant (a)-(b) force  $[\mathcal{F}_{x_1}, \mathcal{F}_{x_2}]^T = -\oint_{\partial\Omega_b} \mathbf{P} \cdot \mathbf{n}_w d\Upsilon$  acting on the hot beam; (c) total heat  $\mathcal{H} = \oint_{\partial\Omega_b} \mathbf{Q} d\Upsilon$  releasing by the hot beam. Solutions for argon of  $\omega = 0.81$ , hard-sphere molecules of  $\omega = 0.5$  and Maxwell molecules of  $\omega = 1.0$  at  $Kn$  ranging from 0.2 to 10 are compared. The scheme of DG-RED of  $k = 4$  combining with the ITR-LOC iteration (13) is applied. The molecular velocity domain  $[-6, 6]^3$  is discretized by  $96 \times 96 \times 24$  non-uniform grid points. 1290 triangles are used for all Knudsen numbers.

The computational domain is  $1 \times 1$  square partitioned by structured triangular mesh as shown in Fig. 7(a). The left and right walls are maintained at constant temperature  $T_C$ , while the bottom and top walls have varied temperature given by:

$$T(x_1, x_2 = 0 \text{ or } 1) = \begin{cases} 2(T_H - T_C)x_1 + T_C, & x_1 \leq 0.5, \\ -2(T_H - T_C)x_1 + 2T_H - T_C, & x_1 > 0.5, \end{cases} \quad (35)$$

where  $T_C$  and  $T_H$  are set as 263 K and 283 K, respectively.

The argon gas with viscosity index  $\omega = 0.81$  is initialized at the reference temperature of  $T_0 = 273$  K. For all cases, the molecular velocity domain is chosen as  $[-6, 6]^3$ , which is discretized by  $72 \times 72$  non-uniform points in the  $v_1$  and  $v_2$  directions, and 24 uniform points in the  $v_3$  direction. The corresponding frequency space, however, are discretized by  $32 \times 32 \times 24$  equidistant frequencies for evaluation of the collision operator.

For verification of the DG results, the FDM results serve as reference solutions. In order to ensure accuracy of the FDM,  $201 \times 201$  equidistant grid points are employed for the spatial discretization. Further refinement of both the velocity and spatial girds would only improve the solution by a magnitude no more than 0.5%. The DG-RED scheme with  $k = 4$  is used to solve the flows on 72 triangles. The ITR-LOC iteration (13) is applied. Figure 13 illustrates the dimensionless temperature and shear stress contours, as well as the streamlines for flow at  $Kn = 0.5$ . The DG-RED steady-state solution presented here costs about 34 iterative steps and 0.96 hour CPU time on single processor. It is observed that high flow temperatures occur near the centers of bottom and top walls due to heating from the walls, while low temperatures appear in the four corners. The tangential temperature gradients near the walls lead to the thermal creep flows, where gas molecules along the bottom and top walls move from the colder regions towards the hotter ones. Due to the confinement of vertical walls, 4 vortexes are generated: two at the lower left and upper right quarters rotate counter-clockwise and the other two rotate clockwise. As a consequence, the maximum shear stresses appear at the centers of clockwise vortices, while the minimum ones occur at the centers of counter-clockwise vortices. The flow patterns at  $Kn = 0.1$  and  $Kn = 1.0$  are similar.

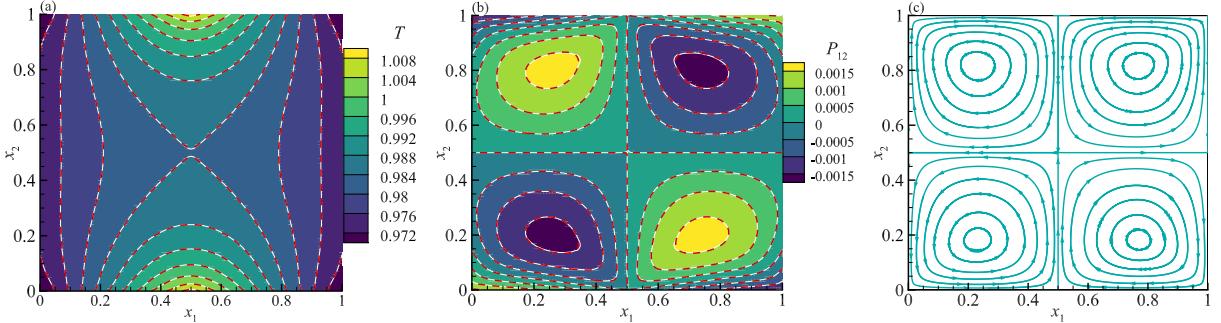


Figure 13: Comparison of the DG-RED and the FDM on the thermal cavity flow induced by the temperature gradients at wall when  $Kn = 0.5$ . Contours of the temperature  $T$  and shear stress  $P_{12}$  are shown in (a) and (b), respectively, where the solid lines with background are solutions of the DG-RED, while the red dashed lines are the FDM results. (c) Stream lines. The DG-RED solutions are obtained with  $k = 4$  and 72 uniform triangles. The molecular velocity domain  $[-6, 6]^3$  is discretized by  $72 \times 72 \times 24$  non-uniform grid points.

Figure 14 illustrates the variations of temperature  $T$ , shear stress  $P_{12}$ , horizontal (vertical) heat fluxes  $Q_1$  ( $Q_2$ ) and horizontal (vertical) flow velocities  $u_1$  ( $u_2$ ) along selected horizontal and vertical lines for rarefied gas flow when  $Kn = 0.1$ ; those for  $Kn = 0.5$  and  $Kn = 1$  are plotted in Figs. 15 and 16, respectively. Due to the symmetry of flow field, results are only shown within the lower left quarter of the computational domain. It is found that from the regions near solid walls to the flow field center, the gas temperature increases along horizontal lines, while decreases along vertical lines. However, along both the horizontal and vertical directions, the shear stress first drops to the local minimum values then rises back to zero. The variations of horizontal heat flux are similar as those of shear stress, while the changes of the vertical component of heat flux are in accordance with the variations of gas temperature. The variations of bulk flow velocity are more complicated. Along the vertical lines, the horizontal velocity  $u_1$  first increases to the local peaked values and then falls to the minimums. Along the horizontal lines near the bottom wall,  $u_1$  is positive and has a local maximum at  $x_1 = 0.25$ , while in the regions away from the bottom wall,  $u_1$  becomes negative and has a local minimum at  $x_1 = 0.25$ . Similarly, near the left lateral wall, the vertical velocity  $u_2$  is negative and gradually changes its sign and reaches the local maximal values when approaching to the field center along the horizontal lines. For all flow properties, agreement between the DG-RED and the FDM results is pretty good. It is also interesting to note that, as the degree of rarefaction increases, the maximum values of temperature decrease since the intensity of gas-gas/gas-wall interactions becomes weaker. On the other hand, the maximum value in magnitudes of heat fluxes  $|\mathbf{Q}|$  occurring near the centers of the bottom and top walls becomes larger, due to the larger temperature jump in high rarefied gas.

## 7. Conclusions

In summary, we have developed a high-order discontinuous Galerkin discretization to solve the Boltzmann equation with full collision operator. The proposed numerical scheme

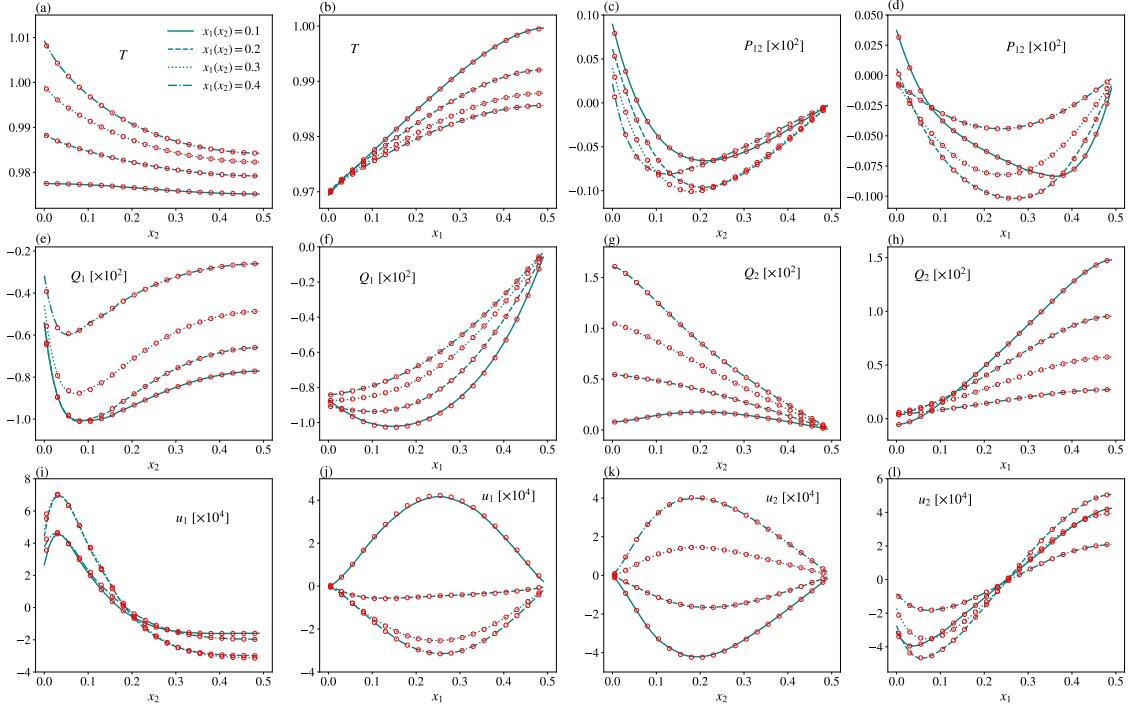


Figure 14: Comparison of the DG-RED (solid lines) and the FDM (circles) on thermal cavity flow induced by temperature gradients at wall when  $Kn = 0.1$ . The first and third columns are flow properties along vertical lines at  $x_1 = 0.1, 0.2, 0.3$  and  $0.4$ , while the second and forth columns are flow properties along horizontal lines at  $x_2 = 0.1, 0.2, 0.3$  and  $0.4$ . The DG-RED solutions are obtained with  $k = 4$  and 72 uniform triangles.

is based on the classical discrete velocity method. At each discrete velocity grid points, the velocity distribution function is approximated in the piecewise polynomial spaces of degree up to 4 in the spatial space. Concerning the resolution of the Boltzmann collision operator, we rely on the Carleman-representation-based Fourier techniques, which can preserve mass and momentum and energy with spectral accuracy. Due to incorporation of the DG discretization and the fast spectral method, the computational complexity to evaluate the collision operator is of the order of  $O(K^2 M_{el} M_{qua} \bar{N}^3 \log \bar{N} + 2K^3 M_{el} \bar{N}^3)$ , which can be significantly large when high order approximating polynomial is used. Based on the nodal DG approximating, we have proposed a reduced DG discretization for the collision operator, which can reduce the computational complexity by  $K$  times of magnitude.

Implicit iterative scheme scheme is employed to find the steady-state solution. At each iterative step, the DG discretization results in a system of linear equations for the degrees of freedom of velocity distribution functions on each spatial element. Since the first-order upwind principle is applied to approximate fluxes on the faces of spatial elements, the local linear equations only couple the unknowns on the immediate neighboring elements in the upwind side. Due to the fact that the direction of molecular velocity is fixed after discretization of the molecular velocity space, we have successfully employed the sweeping technique to sequentially solve the local linear systems, which avoids solving large sparse linear systems

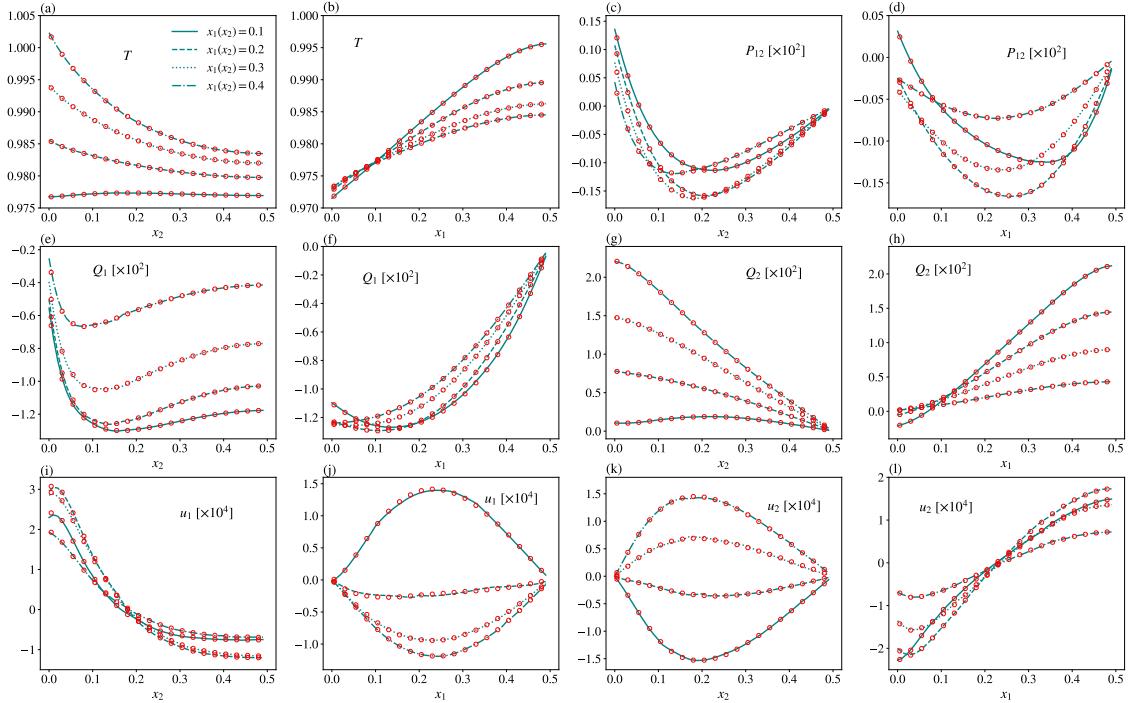


Figure 15: Comparison of the DG-RED (solid lines) and the FDM (circles) on thermal cavity flow induced by temperature gradients at wall when  $Kn = 0.5$ . The first and third columns are flow properties along vertical lines at  $x_1 = 0.1, 0.2, 0.3$  and  $0.4$ , while the second and forth columns are flow properties along horizontal lines at  $x_2 = 0.1, 0.2, 0.3$  and  $0.4$ . The DG-RED solutions are obtained with  $k = 4$  and 72 uniform triangles.

for all elements that are extremely expensive in both memory and CPU time when a large number of discrete velocities are required.

Five different test cases including hypersonic flows, as well as shear-driven and thermal-driven low-speed flows have been presented to show accuracy and capability of the proposed method. Several conclusions are summarized through the performance analysis:

- The implicit iterative scheme has no restriction on time step by CFL condition. The DG schemes with different order of approximating polynomials can obtain steady-state solution of the same order of accuracy within same number of iterative steps. Thus, the higher-order the discretization, the fewer spatial elements thus less CPU time is required.
- Compared to the full DG discretization in the collision operator, the proposed reduced DG approximation preserves the accuracy of the numerical scheme even for highly non-equilibrium flows, and significantly reduces the computational cost. To obtain the results with the same order of accuracy, the higher degree of approximation polynomial, the more the saving of CPU time in the reduced DG approximation.
- Based on the same fast spectral method for the approximation of the Boltzmann collision operator, comparison with the finite difference method shows that the DG

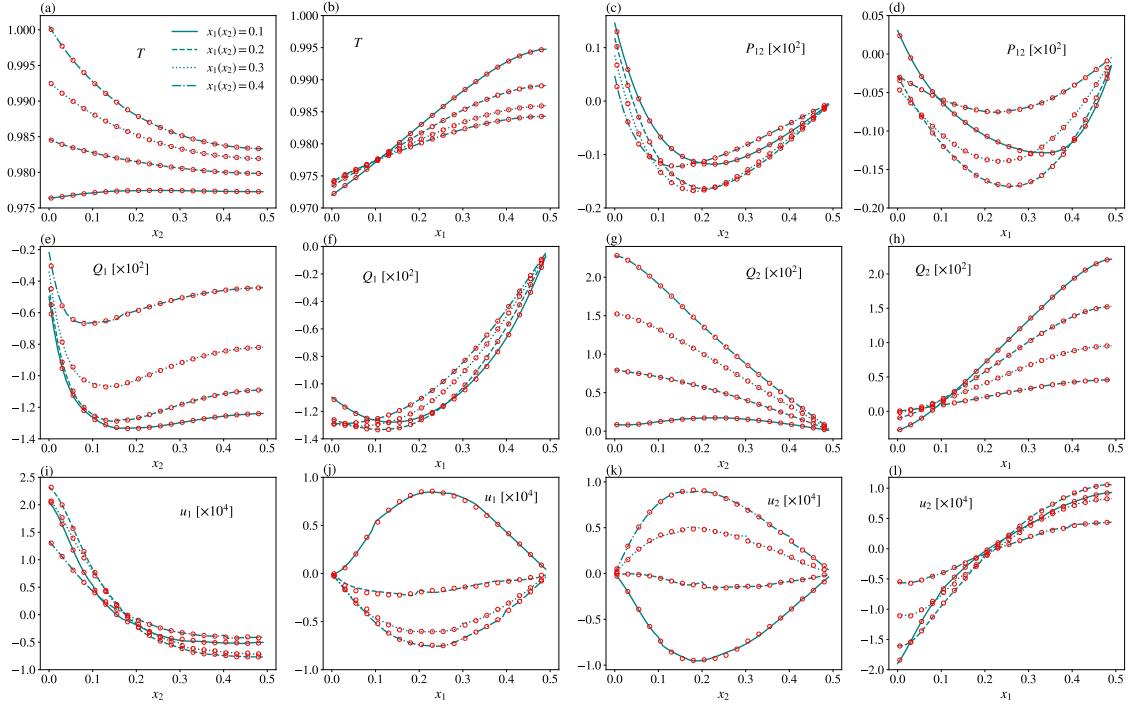


Figure 16: Comparison of the DG-RED (solid lines) and the FDM (circles) on thermal cavity flow induced by temperature gradients at wall when  $Kn = 1.0$ . The first and third columns are flow properties along vertical lines at  $x_1 = 0.1, 0.2, 0.3$  and  $0.4$ , while the second and forth columns are flow properties along horizontal lines at  $x_2 = 0.1, 0.2, 0.3$  and  $0.4$ . The DG-RED solutions are obtained with  $k = 4$  and 72 uniform triangles.

discretization is more efficient. Equipped with the implicit iterative scheme involving global mean collision frequency, the DG scheme can be faster than the finite difference method by one order of magnitude.

- The implicit iterative scheme combining with the sweeping technique to sequentially solve the local linear systems on each spatial element preserves the stability of the DG scheme. Since in rarefied gas flow simulations the shock wave structure are resolved by fine spatial grids, the proposed method can solve hypersonic flows without any nonlinear limiter.

The developed numerical method is straightforward to be extended for the simulation of rarefied gas mixtures, where the velocity distribution function for each species is governed by its own Boltzmann equation. The Boltzmann equations for all constituents are coupled through pairwise collision operators. Thus, the computational complexity in resolving the collision terms via the FSM significantly increases as the number of gas species increases. In such situation, the advantage of using implicit DG method as well as the reduced calculation in collision operator will become more pronounced. Moreover, by incorporating more realistic intermolecular potentials such as the Lennard-Jones potential or even the ab initio potential based on quantum scattering [76], the developed scheme is ready to simulate a wide range

of rarefied gas problems.

## Acknowledgments

This work is founded by the Engineering and Physical Sciences Research Council (EPSRC) in the UK under grant EP/R041938/1.

## Appendix

Here, we present details of the DG formulation for the Boltzmann equation. The linear systems (24) to determine the solution of  $f^{j'}$  on spatial element  $\Delta_i$  are recalled here:

$$\mathbf{A}^{i,j'} \mathbf{F}_i^{j'} + \mathbf{B}^{\text{ext},j'} = \mathbf{S}^{i,j'}, \quad (\text{A.1})$$

for  $i = 1, \dots, M_{\text{el}}$ ,  $j' = 1, \dots, M$ .

We denote  $F_r^{j'} = F_r(\mathbf{v}^{j'})$ ,  $\Lambda_p^{j'} = \Lambda_p(\mathbf{v}^{j'})$  and  $\Xi_{p,r}^{j'} = \Xi_{p,r}(\mathbf{v}^{j'})$  as values of the corresponding variables at each discrete velocity point, and  $\mathbf{F}_i^{j'} = [F_1^{j'}, \dots, F_r^{j'}, \dots]^T$  is the vector of degrees of freedom of  $f^{j'}$  on  $\Delta_i$ . For ITR-LOC scheme, the coefficient matrices are:

$$\mathbf{A}_{sr}^{i,j'} = \frac{1}{2} \left( \mathbf{v}^{j'} \cdot \mathbf{n} + |\mathbf{v}^{j'} \cdot \mathbf{n}| \right) \langle \varphi_s, \varphi_r \rangle_{\partial\Delta_i} - \left( \mathbf{v}^{j'} \cdot \nabla \varphi_s, \varphi_r \right)_{\Delta_i} + \sum_{p=1}^K (\varphi_s, \varphi_p \varphi_r)_{\Delta_i} \Lambda_p^{j'}, \quad (\text{A.2})$$

$$\mathbf{B}_s^{\text{ext},j'} = \begin{cases} \frac{1}{2} (\mathbf{v}^{j'} \cdot \mathbf{n} - |\mathbf{v}^{j'} \cdot \mathbf{n}|) \sum_{r=1}^K \langle \varphi_s, \varphi_r^{\text{ext}} \rangle_{\partial\Delta_i} F_{r,\text{ext}}^{j'}, & \partial\Delta_i \not\subset \partial\Delta \\ \frac{1}{2} (\mathbf{v}^{j'} \cdot \mathbf{n} - |\mathbf{v}^{j'} \cdot \mathbf{n}|) \langle \varphi_s, b^{j'} \rangle_{\partial\Delta_i}, & \partial\Delta_i \subset \partial\Delta \end{cases} \quad (\text{A.3})$$

$$\mathbf{S}_s = \sum_{p=1}^K \sum_{r=1}^K (\varphi_s, \varphi_p \varphi_r)_{\Delta_i} \Xi_{p,r}^{j'} \quad (\text{A.4})$$

where  $\varphi_r^{\text{ext}}$  denotes the supporting polynomials on the neighboring element, from which  $f_{\text{ext}}$  is obtained. For ITR-MEAN scheme, the coefficient matrices become:

$$\mathbf{A}_{sr}^{i,j'} = \frac{1}{2} \left( \mathbf{v}^{j'} \cdot \mathbf{n} + |\mathbf{v}^{j'} \cdot \mathbf{n}| \right) \langle \varphi_s, \varphi_r \rangle_{\partial\Delta_i} - \left( \mathbf{v}^{j'} \cdot \nabla \varphi_s, \varphi_r \right)_{\Delta_i} + (\varphi_s, \varphi_r)_{\Delta_i} \bar{\nu}, \quad (\text{A.5})$$

$$\mathbf{B}_s^{\text{ext},j'} = \begin{cases} \frac{1}{2} (\mathbf{v}^{j'} \cdot \mathbf{n} - |\mathbf{v}^{j'} \cdot \mathbf{n}|) \sum_{r=1}^K \langle \varphi_s, \varphi_r^{\text{ext}} \rangle_{\partial\Delta_i} F_{r,\text{ext}}^{j'}, & \partial\Delta_i \not\subset \partial\Delta \\ \frac{1}{2} (\mathbf{v}^{j'} \cdot \mathbf{n} - |\mathbf{v}^{j'} \cdot \mathbf{n}|) \langle \varphi_s, b^{j'} \rangle_{\partial\Delta_i}, & \partial\Delta_i \subset \partial\Delta \end{cases} \quad (\text{A.6})$$

$$\mathbf{S}_s = \sum_{p=1}^K \sum_{r=1}^K (\varphi_s, \varphi_p \varphi_r)_{\Delta_i} \left( \Xi_{p,r}^{j'} - \Lambda_p^{j'} F_r^{j'} \right) + \sum_{r=1}^K (\varphi_s, \varphi_r)_{\Delta_i} \bar{\nu} F_r^{j'} \quad (\text{A.7})$$

In this paper, nodal shape functions are chosen as the approximating polynomials. Integrals of the shape functions such as  $(\varphi_s, \varphi_r)$ ,  $(\nabla \varphi_s, \varphi_r)$ ,  $(\varphi_s, \varphi_p \varphi_r)$  and  $\langle \varphi_s, \varphi_r \rangle$  can be obtained analytically. To evaluate  $\langle \varphi_s, b^{j'} \rangle$ , the Gaussian rule is applied.

## References

- [1] S. Chapman, T. Cowling, *The Mathematical Theory of Non-uniform Gases*, 3rd Edition, Cambridge University press, New York, 1970.
- [2] G. Dimarco, L. Pareschi, Numerical methods for kinetic equations, *Acta Numerica* 23 (2014) 369–520.
- [3] G. A. Bird, *Molecular Gas Dynamics and the Direct Simulation*, Clarendon, Oxford, 1994.
- [4] A. B. Huang, D. P. Giddens, The discrete ordinate method for the linearized boundary value problems in kinetic theory of gases, in: C. L. Brundin (Ed.), *Rarefied Gas Dynamics*, Volume 1, 1967, p. 481.
- [5] V. V. Aristov, *Direct Methods for Solving the Boltzmann Equation and Study of Nonequilibrium Flows*, Springer Netherlands, Dordrecht, 2001.
- [6] W. Wagner, Approximation of the Boltzmann equation by discrete velocity models, *Journal of Statistical Physics* 78 (5) (1995) 1555–1570.
- [7] C. Buet, A discrete-velocity scheme for the Boltzmann operator of rarefied gas dynamics, *Transport Theory and Statistical Physics* 25 (1) (1996) 33–60.
- [8] L. L. Baker, N. G. Hadjiconstantinou, Variance-reduced Monte Carlo solutions of the Boltzmann equation for low-speed gas flows: A discontinuous Galerkin formulation, *International Journal for Numerical Methods in Fluids* 58 (4) (2008) 381–402.
- [9] J. Yang, J. Huang, Rarefied flow computations using nonlinear model Boltzmann equations, *Journal of Computational Physics* 120 (2) (1995) 323 – 339.
- [10] V. Kolobov, R. Arslanbekov, V. Aristov, A. Frolova, S. Zabelok, Unified solver for rarefied and continuum flows with adaptive mesh and algorithm refinement, *Journal of Computational Physics* 223 (2) (2007) 589 – 608.
- [11] L. Mieussens, Discrete-velocity models and numerical schemes for the Boltzmann-BGK equation in plane and axisymmetric geometries, *Journal of Computational Physics* 162 (2) (2000) 429 – 466.
- [12] V. A. Titarev, Efficient deterministic modelling of three-dimensional rarefied gas flows, *Communications in Computational Physics* 12 (1) (2012) 162192.
- [13] M. Gobbert, S. Webster, T. Cale, A Galerkin method for the simulation of the transient 2-D/2-D and 3-D/3-D linear Boltzmann equation, *Journal of Scientific Computing* 30 (2) (2007) 237–273.
- [14] G. Kitzler, J. Schöberl, A high order space-momentum discontinuous Galerkin method for the Boltzmann equation, *Computers & Mathematics with Applications* 70 (7) (2015) 1539 – 1554.
- [15] N. Crouseilles, T. Respaud, E. Sonnendrcker, A forward semi-Lagrangian method for the numerical solution of the Vlasov equation, *Computer Physics Communications* 180 (10) (2009) 1730 – 1745.
- [16] N. Crouseilles, M. Mehrenberger, E. Sonnendrcker, Conservative semi-Lagrangian schemes for Vlasov equations, *Journal of Computational Physics* 229 (6) (2010) 1927 – 1953.
- [17] Y. Gl, W. Hitchon, A high order cell-centered semi-Lagrangian scheme for multi-dimensional kinetic simulations of neutral gas flows, *Journal of Computational Physics* 231 (8) (2012) 3289 – 3316.
- [18] G. Dimarco, R. Loubere, Towards an ultra efficient kinetic scheme. Part I: Basics on the BGK equation, *Journal of Computational Physics* 255 (2013) 680 – 698.
- [19] G. Dimarco, R. Loubere, Towards an ultra efficient kinetic scheme. Part II: The high order case, *Journal of Computational Physics* 255 (2013) 699 – 719.
- [20] P. L. Bhatnagar, E. P. Gross, M. Krook, A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems, *Physical Review* 94 (1954) 511–525.
- [21] L. H. Holway, New statistical models for kinetic theory: Methods of construction, *Physics of Fluids* 9 (9) (1966) 1658–1673.
- [22] E. Shakhov, Generalization of the Krook kinetic relaxation equation, *Fluid Dynamics* 3 (5) (1968) 95–96.
- [23] D. Goldstein, B. Sturtevant, J. E. Broadwell, Investigations of the motion of discrete-velocity gases, *Progress in Astronautics and Aeronautics* 117 (1989) 100–117.
- [24] A. Bobylev, A. Palczewski, J. Schneider, On approximation of the Boltzmann equation by discrete velocity models, *Comptes Rendus de l'Academie des Sciences* 320 (1995) 639–644.
- [25] A. B. Morris, P. L. Varghese, D. B. Goldstein, Improvement of a discrete velocity Boltzmann equation solver with arbitrary postcollision velocities, *AIP Conference Proceedings* 1084 (1) (2008) 458–463.

- [26] Y. Sone, T. Ohwada, K. Aoki, Temperature jump and Knudsen layer in a rarefied gas over a plane wall: Numerical analysis of the linearized Boltzmann equation for hardsphere molecules, *Physics of Fluids A: Fluid Dynamics* 1 (2) (1989) 363–370.
- [27] T. Ohwada, Structure of normal shock waves: Direct numerical analysis of the Boltzmann equation for hard-sphere molecules, *Physics of Fluids A: Fluid Dynamics* 5 (1) (1993) 217–234.
- [28] T. Ohwada, Heat flow and temperature and density distributions in a rarefied gas between parallel plates with different temperatures. Finitedifference analysis of the nonlinear Boltzmann equation for hard-sphere molecules, *Physics of Fluids* 8 (8) (1996) 2153–2160.
- [29] S. Kosuge, K. Aoki, S. Takata, Shock-wave structure for a binary gas mixture: finite-difference analysis of the Boltzmann equation for hard-sphere molecules, *European Journal of Mechanics - B/Fluids* 20 (1) (2001) 87 – 126.
- [30] F. Tcheremissine, Conservative evaluation of Boltzmann collision integral in discrete ordinates approximation, *Computers & Mathematics with Applications* 35 (1) (1998) 215 – 221.
- [31] F. G. Tcheremissine, Solution to the Boltzmann kinetic equation for high-speed flows, *Computational Mathematics and Mathematical Physics* 46 (2) (2006) 315–329.
- [32] A. Alekseenko, T. Nguyen, A. Wood, A deterministic-stochastic method for computing the Boltzmann collision integral in  $\mathcal{O}(mn)$  operations, *Kinetic & Related Models* 11 (5) (2018) 1211–1234.
- [33] A. Bobylev, The theory of the nonlinear spatially uniform Boltzmann equation for Maxwell molecules, *Mathematical Physics Reviews* 7 (1988) 111–233.
- [34] L. Pareschi, B. Perthame, A Fourier spectral method for homogeneous Boltzmann equations, *Transport Theory and Statistical Physics* 25 (3-5) (1996) 369–382.
- [35] A. Bobylev, R. Rjasanow, Difference scheme for the Boltzmann equation based on fast Fourier transformation, *European Journal of Mechanics B/Fluids* 16 (2) (1997) 293–306.
- [36] C. Watchararuangwita, Y. N. Grigoriev, S. V. Meleshko, A deterministic spectral method for solving the Boltzmann equation for one-dimensional flows, *Science Asia* 35 (1) (2009) 70–79.
- [37] A. Bobylev, S. Rjasanow, Fast deterministic method of solving the Boltzmann equation for hard spheres, *European Journal of Mechanics - B/Fluids* 18 (5) (1999) 869 – 887.
- [38] I. Ibragimov, S. Rjasanow, Numerical solution of the Boltzmann equation on the uniform grid, *Computing* 69 (2) (2002) 163–186.
- [39] L. Pareschi, G. Russo, Numerical solution of the Boltzmann equation I: Spectrally accurate approximation of the collision operator, *SIAM Journal on Numerical Analysis* 37 (4) (2000) 1217–1245.
- [40] F. Filbet, G. Russo, High order numerical methods for the space non-homogeneous Boltzmann equation, *Journal of Computational Physics* 186 (2) (2003) 457 – 480.
- [41] F. Filbet, C. Mouhot, L. Pareschi, Solving the Boltzmann equation in  $n \log^2 n$ , *SIAM Journal on Scientific Computing* 28 (3) (2006) 1029–1053.
- [42] C. Mouhot, L. Pareschi, Fast algorithms for computing the Boltzmann collision operator, *Mathematics of Computation* 75 (256) (2006) 1833–1852.
- [43] L. Wu, C. White, T. J. Scanlon, J. M. Reese, Y. Zhang, Deterministic numerical solutions of the Boltzmann equation using the fast spectral method, *Journal of Computational Physics* 250 (2013) 27 – 52.
- [44] L. Wu, J. M. Reese, Y. Zhang, Solving the Boltzmann equation deterministically by the fast spectral method: application to gas microflows, *Journal of Fluid Mechanics* 746 (2014) 5384.
- [45] L. Wu, H. Liu, Y. Zhang, J. M. Reese, Influence of intermolecular potentials on rarefied gas flows: Fast spectral solutions of the Boltzmann equation, *Physics of Fluids* 27 (8) (2015) 082002.
- [46] G. A. Radtke, N. G. Hadjiconstantinou, W. Wagner, Low-noise monte carlo simulation of the variable hard sphere gas, *Physics of Fluids* 23 (3) (2011) 030606.
- [47] W. H. Reed, T. R. Hill, Triangular Mesh Methods for the Neutron Transport Equation, *Tech. Rep.* 836 (1973).
- [48] B. Cockburn, C.-W. Shu, The Runge-Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems, *Journal of Computational Physics* 141 (2) (1998) 199 – 224.
- [49] B. Cockburn, C.-W. Shu, Runge-Kutta discontinuous Galerkin methods for convection-dominated prob-

- lems, *Journal of Scientific Computing* 16 (3) (2001) 173–261.
- [50] W. Su, A. A. Alexeenko, G. Cai, A parallel Runge–Kutta discontinuous Galerkin solver for rarefied gas flows based on 2D Boltzmann kinetic equations, *Computers & Fluids* 109 (2015) 123 – 136.
- [51] S. Jaiswal, A. A. Alexeenko, J. Hu, A discontinuous Galerkin fast spectral method for the full Boltzmann equation with general collision kernels, *Journal of Computational Physics*.
- [52] E. J. Kubatko, C. Dawson, J. J. Westerink, Time step restrictions for Runge–Kutta discontinuous Galerkin methods on triangular grids, *Journal of Computational Physics* 227 (23) (2008) 9697 – 9710.
- [53] G. Dimarco, R. Loubre, J. Narski, T. Rey, An efficient numerical method for solving the Boltzmann equation in multidimensions, *Journal of Computational Physics* 353 (2018) 46 – 81.
- [54] C. Liu, K. Xu, Q. Sun, Q. Cai, A unified gas-kinetic scheme for continuum and rarefied flows IV: Full Boltzmann and model equations, *Journal of Computational Physics* 314 (2016) 305 – 340.
- [55] Y. Zhu, C. Zhong, K. Xu, Implicit unified gas-kinetic scheme for steady state solutions in all flow regimes, *Journal of Computational Physics* 315 (2016) 16 – 38.
- [56] Z. Guo, K. Xu, R. Wang, Discrete unified gas kinetic scheme for all Knudsen number flows: Low-speed isothermal case, *Phys. Rev. E* 88 (2013) 033305.
- [57] Z. Guo, R. Wang, K. Xu, Discrete unified gas kinetic scheme for all Knudsen number flows. II. Thermal compressible case, *Physical Review E* 91 (3) (2015) 033313.
- [58] W. Su, S. Lindsay, H. Liu, L. Wu, Comparative study of the discrete velocity and lattice Boltzmann methods for rarefied gas flows through irregular channels, *Physics Review E* 96 (2017) 023309.
- [59] A. Crivellini, F. Bassi, An implicit matrix-free discontinuous Galerkin solver for viscous and turbulent aerodynamic simulations, *Computers & Fluids* 50 (1) (2011) 81 – 93.
- [60] S. Murphy, Methods for solving discontinuous-galerkin finite element equations with application to neutron transport, Ph.D. thesis, Institute National Polytechnique de Toulouse (2015).
- [61] H. Alsmeyer, Density profiles in argon and nitrogen shock waves measured by the absorption of an electron beam, *Journal of Fluid Mechanics* 74 (3) (1976) 497513.
- [62] Z. Tang, B. He, G. Cai, Investigation on a coupled Navierstokes direct simulation Monte Carlo method for the simulation of plume flowfield of a conical nozzle, *International Journal for Numerical Methods in Fluids* 76 (2) 95–108.
- [63] K. Xu, J.-C. Huang, A unified gas-kinetic scheme for continuum and rarefied flows, *Journal of Computational Physics* 229 (20) (2010) 7747 – 7764.
- [64] P. Valentini, T. E. Schwartzentruber, Large-scale molecular dynamics simulations of normal shock waves in dilute argon, *Physics of Fluids* 21 (6) (2009) 066101.
- [65] S. Yen, Temperature overshoot in shock waves, *The Physics of Fluids* 9 (7) (1966) 1417–1418.
- [66] S. Chen, C. Zhang, L. Zhu, Z. Guo, A unified implicit scheme for kinetic model equations. Part I. Memory reduction technique, *Science Bulletin* 62 (2) (2017) 119 – 129.
- [67] B. John, X.-J. Gu, D. R. Emerson, Investigation of heat and mass transfer in a lid-driven cavity under nonequilibrium flow conditions, *Numerical Heat Transfer, Part B: Fundamentals* 58 (5) (2010) 287–303.
- [68] J. C. Chai, H. S. Lee, S. V. Patankar, Ray effect and false scattering in the discrete ordinates method, *Numerical Heat Transfer, Part B: Fundamentals* 24 (4) (1993) 373–389.
- [69] P. Coelho, The role of ray effects and false scattering on the accuracy of the standard and modified discrete ordinates methods, *Journal of Quantitative Spectroscopy and Radiative Transfer* 73 (2) (2002) 231 – 238, third International Symposium on Radiative Transfer.
- [70] A. Passian, R. J. Warmack, T. L. Ferrell, T. Thundat, Thermal transpiration at the microscale: A crookes cantilever, *Phys. Rev. Lett.* 90 (2003) 124503.
- [71] J. Nabeth, S. Chigullapalli, A. A. Alexeenko, Quantifying the Knudsen force on heated microbeams: A compact model and direct comparison with measurements, *Phys. Rev. E* 83 (2011) 066306.
- [72] A. Strongrich, A. Alexeenko, Microstructure actuation and gas sensing by the Knudsen thermal force, *Applied Physics Letters* 107 (19) (2015) 193508.
- [73] Y. Sone, M. Yoshimoto, Demonstration of a rarefied gas flow induced near the edge of a uniformly heated plate, *Physics of Fluids* 9 (11) (1997) 3530–3534.
- [74] T. Zhu, W. Ye, Origin of Knudsen forces on heated microbeams, *Phys. Rev. E* 82 (2010) 036308.

- [75] L. Zhu, Z. Guo, Application of discrete unified gas kinetic scheme to thermally induced nonequilibrium flows, *Computers & Fluids*.
- [76] F. Sharipov, Modeling of transport phenomena in gases based on quantum scattering, *Physica A: Statistical Mechanics and its Applications* 508 (2018) 797 – 805.