AN ENTROPY STABLE SCHEME FOR THE NON-LINEAR BOLTZMANN EQUATION

SHASHANK JAISWAL

Abstract. The kinetic equations govern the behavior of gaseous flows, compressibility, turbulence, reactions with internal energy exchange, that form the essence of numerous physical processes. Despite their wide applicability, their six-dimensional (seven including time) nature presents a huge computational challenge. With the advent of the modern high performance computing systems and few recent advances in numerical methods, it is now possible to *numerically* study the behavior of these systems. However, to understand the instability of rich molecular processes and find a structure in the apparent chaos, a scheme that is efficient in multi-dimensions, exhibits high parallel-efficiency, and foremost produces an entropy solution as per the theory of solutions of hyperbolic systems may prove useful. In this work, first, we construct an entropy stable flux for non-linear inhomogeneous (full) Boltzmann equation. Second, to ensure geometrical flexibility, we couple the scheme with a class of high order discontinuous Galerkin discretization (Jaiswal 2019b [41]) which satisfy summation-by-part (SBP) discretely. Third, utilizing SBP, we prove that the resulting *semi-discrete* scheme is locally and globally conservative in flat spaces; efficient and simple; and therefore suitable for dealing with highly complex *non-smooth* flow problems. Fourth, we show that the fully-discrete kinetic scheme, utilizing an implicit-explicit time-discretization, in the limit of vanishing Knudsen number, becomes an entropy-stable explicit scheme applied to the Euler system. Fifth, we carry out a series of verification tests to illustrate the stability, accuracy, and conservation properties of the proposed method. These tests involve over 85 million degrees of freedom and over 50 billion operations per time-step.

Key words. Entropy stability, Boltzmann equation, High order schemes, Discontinuous Galerkin.
AMS subject classifications. 76P05, 82B40, 82C40, 82D05, 35Q20, 65M60, 65M70

1. Introduction

The premise that the kinetic models, arising at a detailed molecular level description of the evolution of matter, are linked with the continuum equations in a limit, for example mean free path approaching zero, has been emphasized since the beginning of the kinetic theory of gases. It has always been hoped that this limit will reveal and illuminate some of the complex phenomena present at the macroscopic level [53]. Kinetic theory may hold the key to understanding turbulence [24, 25]. Indeed, incompressible Navier-Stokes [13] can be formally derived from the asymptotics of Boltzmann equation[1]. The widely-used approaches for numerical solution of the kinetic equations are based on the Monte Carlo (MC) methods [3, 4] because the process of statistical sampling is often computationally efficient for evaluating multi-dimensional integrals contained therein. For example, the direct simulation MC (DSMC) method, based on the kinetic theory of dilute gases, introduces a set of particles in the computational domain, transports them according to their velocities, and then models the binary interactions between them using MC. In the limit of infinitely-many number of particles and interactions, DSMC method converges to the Boltzmann equation [66]. These methods are useful for simulating the molecular processes, compressible, turbulent, non-equilibrium chemical reactions with internal energy exchange, that form the core of our understanding of the physical reality at microscale. However, the DSMC methods are based on the assumption that the binary collision processes are local processes and therefore only particles within a computational cell may collide with each other. This implies that the size of the cell, in which the particle resides, should be less than the mean-free path (MFP) of the particle, and the particle, during a single time-step, should move within the cell so that it may actually collide with other particles. In near-continuum cases, as MFP approaches zero, naturally these methods become expensive due to increase in number of cells and decrease in time-step (ensures that the particle remains in the cell during interaction). For reference, the turbulence simulations in [24] were carried out on half-a-million core system, inaccessible to the most of academia or industry. In this work, we aim to solve the same problem using full Boltzmann equation on few graphics processing units, accessible to the most of academia or industry.

Date: May 6, 2022.
J. Comput. Phys.; Submitted to associate editor on 9 Nov 2021; Revised 22 April 2022; Accepted 6 May 2022.
[1]This is different than Chapman-Enskog expansions which have never received a satisfactory mathematical justification [65].

The other approaches for numerical simulations of kinetic equations are deterministic, which seek to discretize the Boltzmann equation and compute numerically approximate solutions. We refer to [16] for a recent review. To understand the instability of rich molecular processes and find a structure in the apparent chaos, we need a scheme that doesn't break down numerically, is efficient in multi-dimensions on parallel computers, and foremost produces a unique weak solution as per current understanding of the theory of solutions of hyperbolic systems [48]. In this regard, we mention that DSMC simulations are unconditionally stable, so we may not be able to verify if the simulation has actually converged to the correct solution, other than the hope that it probably has. Rerunning the simulation may not be always an option because of the enormous amount of effort required, computationally, economically, or otherwise. For example, the half-a-million core simulations [24] are at least an order of magnitude costly than equivalent laboratory experiments (estimate based on the pricing of Amazon-EC2). The aim of modelling and simulation has always been the other way around. This issue also appears in direct numerical simulations of reactive turbulence [9]. Nevertheless, we emphasize that the potential of kinetic simulations are enormous, for example, the computational cost of continuum reactive turbulence calculations is affected by the number of species being modelled; whereas the same is not true for DSMC. Therefore, kinetic simulations open up avenues for the assessment of the accuracy of reactive turbulence theories.

The solution of hyperbolic system which maximizes entropy is referred as entropy solution. The uniqueness has been recently disputed, however [23]. The schemes that produce an entropy solution are called entropy-stable schemes. For linear hyperbolic systems, schemes are designed to be L^2-stable at a discrete level [33]; whereas for nonlinear convection-dominated systems, the entropy stability plays the corresponding role. Osher [57] established a general class of such schemes, called E-schemes, that preserve entropy inequalities. Entropy is crucial in the theory, numerics, and physics of non-linear problems [49] such as Navier-Stokes [8], Boltzmann [17]. To proceed further, let us introduce the notations for the Boltzmann equation.

1.1. **Kinetic equations.** In the present work, we are concerned with the following class of kinetic equations:

$$\partial_t f + \boldsymbol{v}\cdot\nabla_{\boldsymbol{x}} f = \frac{1}{\mathrm{Kn}}\mathcal{Q}(f,f), \quad t\geq 0,\ \boldsymbol{x}\in\Omega_{\boldsymbol{x}}\subset\mathbb{R}^{d_x},\ \boldsymbol{v}\in\mathbb{R}^3,\ \mathrm{Kn}\in(0,\infty) \tag{1}$$

where $f = f(t,\boldsymbol{x},\boldsymbol{v})$ is the one-particle distribution function at time t, position $\boldsymbol{x}$, and particle velocity $\boldsymbol{v}$. $f\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}$ gives the number of particles to be found in an infinitesimal volume $\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}$ centered at the point $(\boldsymbol{x},\,\boldsymbol{v})$ of the phase space. $\mathcal{Q}(f,f)$ is the collision operator describing the interactions among particles, and acts only in the velocity space. For instance, the non-dimensional full Boltzmann collision operator (c.f. [40] for details on non-dimensionalization) is given as

$$\mathcal{Q}(f,f)(\boldsymbol{v}) = \int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\mathcal{B}(|\boldsymbol{v}-\boldsymbol{v}_*|,\cos\chi)[f(\boldsymbol{v}')f(\boldsymbol{v}'_*)-f(\boldsymbol{v})f(\boldsymbol{v}_*)]\,\mathrm{d}\boldsymbol{\omega}\,\mathrm{d}\boldsymbol{v}_*, \tag{2}$$

where $(\boldsymbol{v},\boldsymbol{v}_*)$ and $(\boldsymbol{v}',\boldsymbol{v}'_*)$ denote the pre- and post- collision velocity pairs, which are related through momentum and energy conservation as

$$\boldsymbol{v}' = \frac{\boldsymbol{v}+\boldsymbol{v}_*}{2}+\frac{|\boldsymbol{v}-\boldsymbol{v}_*|}{2}\boldsymbol{\omega},\quad \boldsymbol{v}'_* = \frac{\boldsymbol{v}+\boldsymbol{v}_*}{2}-\frac{|\boldsymbol{v}-\boldsymbol{v}_*|}{2}\boldsymbol{\omega},\quad \cos\chi = \frac{\boldsymbol{\omega}\cdot(\boldsymbol{v}-\boldsymbol{v}_*)}{|\boldsymbol{v}-\boldsymbol{v}_*|}, \tag{3}$$

with the vector $\boldsymbol{\omega}$ varying over the unit sphere $\mathbb{S}^2$. The quantity $\mathcal{B}$ (≥ 0) is the collision kernel depending only on $|\boldsymbol{v}-\boldsymbol{v}_*|$ and the scattering angle χ (angle between $\boldsymbol{v}-\boldsymbol{v}_*$ and $\boldsymbol{v}'-\boldsymbol{v}'_*$). The gain and the loss terms are defined as

$$\begin{aligned}\overline{\mathcal{Q}} &= \int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\mathcal{B}(|\boldsymbol{v}-\boldsymbol{v}_*|,\cos\chi)\ f(\boldsymbol{v}')f(\boldsymbol{v}'_*)\,\mathrm{d}\boldsymbol{\omega}\,\mathrm{d}\boldsymbol{v}_*,\\ \underline{\mathcal{Q}} &= f\cdot\nu(f),\quad \nu(f) = \int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\mathcal{B}(|\boldsymbol{v}-\boldsymbol{v}_*|,\cos\chi)\,f(\boldsymbol{v}_*)\,\mathrm{d}\boldsymbol{\omega}\,\mathrm{d}\boldsymbol{v}_*.\end{aligned} \tag{4}$$

Given the distribution function, f, one can recover the macroscopic observable via moments,

$$\begin{aligned} n &= \int_{\mathbb{R}^3} f\,\mathrm{d}\boldsymbol{v},\quad \boldsymbol{u} = \frac{1}{n}\int_{\mathbb{R}^3} f\,\boldsymbol{v}\,\mathrm{d}\boldsymbol{v},\quad T = \frac{2}{3n}\int_{\mathbb{R}^3} f\,|\boldsymbol{v}-\boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{v},\\ \mathbb{P} &= 2\int_{\mathbb{R}^3} f\,(\boldsymbol{v}-\boldsymbol{u})\otimes(\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}\boldsymbol{v},\quad \boldsymbol{q} = \int_{\mathbb{R}^3} f\,(\boldsymbol{v}-\boldsymbol{u})\,|\boldsymbol{v}-\boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{v}.\end{aligned} \tag{5}$$

where n, $\boldsymbol{u}$, T, $\mathbb{P}$, $\boldsymbol{q}$ are, respectively, flow number density, bulk-velocity, temperature, pressure-tensor, and heat-flux. Here, $\otimes$ refers to the vector outer-product which, in index-notation, reads $(\mathbf{a}\otimes\mathbf{b})_{ij} = a_i b_j$ for some vectors

a and **b**. Throughout this work, we use u and v, respectively, to denote the first component of macroscopic and microscopic velocity.

The collision operator, $\mathcal{Q}$, is expected to satisfy the following basic conditions:

– Mass/momentum/energy conservation:

$$\int_{\mathbb{R}^3} \mathcal{Q}\,\mathrm{d}\boldsymbol{v} = \int_{\mathbb{R}^3} \boldsymbol{v}\,\mathcal{Q}\,\mathrm{d}\boldsymbol{v} = \int_{\mathbb{R}^3} |\boldsymbol{v}|^2\,\mathcal{Q}\,\mathrm{d}\boldsymbol{v} = 0. \tag{6}$$

– H-theorem:

$$H = \int_{\mathbb{R}^3} \partial_t\big(f\ln(f)\big)\,\mathrm{d}\boldsymbol{v} + \int_{\mathbb{R}^3} \nabla_{\boldsymbol{x}}\cdot\big(\boldsymbol{v} f\ln(f)\big)\,\mathrm{d}\boldsymbol{v} \le 0, \quad \text{or} \quad \int_{\mathbb{R}^3} \mathcal{Q}(f)\,\ln(f)\,\mathrm{d}\boldsymbol{v} \le 0, \tag{7}$$

– Equilibrium state is Maxwellian:

$$\mathcal{Q} = 0 \Leftrightarrow \int_{\mathbb{R}^3} \ln(f)\,\mathcal{Q}\,\mathrm{d}\boldsymbol{v} = 0 \Leftrightarrow f = \mathcal{M} = \frac{n}{(\pi T)^{3/2}}\exp\Big(-\frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{T}\Big). \tag{8}$$

– Weak form:

$$\int_{\mathbb{R}^3} \mathcal{Q}\,\Psi(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v} = \frac{1}{4}\frac{1}{\mathrm{Kn}}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2} \mathcal{B}[f(\boldsymbol{v}')f(\boldsymbol{v}'_*) - f(\boldsymbol{v})f(\boldsymbol{v}_*)]\Big(\Psi(\boldsymbol{v}) + \Psi(\boldsymbol{v}_*) - \Psi(\boldsymbol{v}') - \Psi(\boldsymbol{v}'_*)\Big)\,\mathrm{d}\boldsymbol{\omega}\,\mathrm{d}\boldsymbol{v}_*\,\mathrm{d}\boldsymbol{v}. \tag{9}$$

For non-linear systems, the total amount of entropy, $\int H\mathrm{d}x$, does not increase in time, which is a generalization of the weighted L^2-energy bound for the linear systems. H-theorem plays an important role in the well-posedness of Boltzmann equation [17], including convergence [14]. Therefore, it is natural to seek numerical schemes which satisfy equation (7) discretely. The primary objective of the present work is to construct a (high order) entropy stable scheme for Boltzmann equation which can be used for studying critical problems such as turbulence. We mention that it is straightforward to construct *first* order monotone entropy stable scheme along the lines of [31, 32, 57]. We are interested in high resolution schemes. To the best of our knowledge, this paper is the first attempt to construct entropy stable schemes for the Boltzmann equation, their implementation, and application.

In the section, (2), that follows, we describe a method of constructing high order entropy stable schemes for Boltzmann equation. We propose an entropy-stable flux in the process. Section (3) describes a class of high order collocated discontinuous Galerkin (DG) schemes which satisfy integration by parts at a *discrete* level. Due to this property, these class of DG schemes can be modified to fulfill an *arbitrary* entropy condition, while maintaining conservation and high order accuracy. Such schemes are referred as entropy stable DG (EDG) schemes. EDG is utilized to construct a one-dimensional semi-discrete scheme for Boltzmann equation in section (3.3). The section also provides proof for conservation and entropy. The fully-discrete scheme is detailed in section (4). Numerical tests are provided in section (5) and (6).

2. Entropy stable fluxes for Boltzmann equation

Consider a system of one dimensional hyperbolic conservation laws

$$\boldsymbol{f}_t + \boldsymbol{g}(\boldsymbol{f})_x = 0, \quad \boldsymbol{f}(x,0) = \boldsymbol{f}_0(x). \tag{10}$$

Here $\boldsymbol{f} = (f_1, ..., f_j)^{\mathrm{T}}$ is the state vector of unknowns and $\boldsymbol{g}(\boldsymbol{f}) = (g_1, ..., g_j)^{\mathrm{T}}$, the flux, is a vector valued function of j components, equipped with a $j \times j$ Jacobian matrix,

$$\partial\boldsymbol{g}/\partial\boldsymbol{f} = \begin{bmatrix} \dfrac{\partial \boldsymbol{g}(\boldsymbol{f})}{\partial f_1} & \cdots & \dfrac{\partial \boldsymbol{g}(\boldsymbol{f})}{\partial f_j} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial g_1}{\partial f_1} & \cdots & \dfrac{\partial g_1}{\partial f_j} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_j}{\partial f_1} & \cdots & \dfrac{\partial g_j}{\partial f_j} \end{bmatrix},$$

with j real eigenvalues, $\{g_1(\boldsymbol{f}) \le g_2(\boldsymbol{f}) \le \ldots g_j(\boldsymbol{f})\}$, and a complete set of j linearly independent right-eigenvectors. We are interested in the entropy solutions [62] of equation (10), i.e., bounded solutions in the sense of distributions [53] which satisfy Lax's [48] entropy inequalities,

$$A(\boldsymbol{f})_t + B(\boldsymbol{f})_x \le 0, \tag{11}$$

for any *convex* functional $A(\boldsymbol{f})$ satisfying $A''(\boldsymbol{f}) \le 0$ and the corresponding flux

$$B(\boldsymbol{f}) = \int^{\boldsymbol{f}} A'(\boldsymbol{f})\,\boldsymbol{g}'(\boldsymbol{f})\,\mathrm{d}\boldsymbol{f} \implies B'(\boldsymbol{f}) = A'(\boldsymbol{f})\,\boldsymbol{g}'(\boldsymbol{f}). \tag{12}$$

Here A' and A'', respectively, denote the first and second derivative of A with respect to f. $A(\boldsymbol{f})$ is generally referred as entropy functional, and $B(\boldsymbol{f})$ entropy flux. For systems with convex entropy, one can define entropy variables

$$\boldsymbol{a}^{\mathrm{T}} = A'(\boldsymbol{f}) = \partial A/\partial \boldsymbol{f}. \tag{13}$$

Other way around, equation (11) can be obtained by left-multiplying $\boldsymbol{a}^{\mathrm{T}}$ to equation (10) and using equation (12).

The convexity guarantees that the mapping between state and entropy variables, $\boldsymbol{f} \mapsto \boldsymbol{a}$, is one-to-one; and therefore we may recast equation (10) as

$$\boldsymbol{f}(\boldsymbol{a})_t + \boldsymbol{b}(\boldsymbol{f}(\boldsymbol{a}))_x = 0 \implies \boldsymbol{f}'(\boldsymbol{a})\,\boldsymbol{a}_t + \boldsymbol{b}'(\boldsymbol{a})\,\boldsymbol{a}_x = 0. \tag{14}$$

One can assert that $\boldsymbol{f}'(\boldsymbol{a})$ and $\boldsymbol{b}'(\boldsymbol{a})$ are both symmetric, so there exist functions $\varphi(\boldsymbol{a})$ and $\psi(\boldsymbol{a})$, called potential function and potential fluxes (c.f. [54, 30]), such that

$$\boldsymbol{\varphi}'(\boldsymbol{a}) = \boldsymbol{f}(\boldsymbol{a})^{\mathrm{T}}, \quad \boldsymbol{\psi}'(\boldsymbol{a}) = \boldsymbol{b}(\boldsymbol{a})^{\mathrm{T}}, \tag{15}$$

and therefore

$$\boldsymbol{\varphi}(\boldsymbol{a}) = \boldsymbol{f}(\boldsymbol{a})^{\mathrm{T}}\boldsymbol{a} - A(\boldsymbol{f}(\boldsymbol{a})), \quad \boldsymbol{\psi}(\boldsymbol{a}) = \boldsymbol{b}(\boldsymbol{a})^{\mathrm{T}}\boldsymbol{a} - B(\boldsymbol{f}(\boldsymbol{a})), \tag{16}$$

using equations (12), (13), and symmetry.

Next, we multiply (14) by $\boldsymbol{a}^{\mathrm{T}}$, use (16), and integrate over a reference physical space $[\alpha, \beta]$ to find

$$\int_\alpha^\beta \mathrm{d}x\, A(\boldsymbol{f})_t + \big(B(\beta) - B(\alpha)\big) = \int_\alpha^\beta \mathrm{d}x\, (\boldsymbol{a}_x)^{\mathrm{T}}\boldsymbol{g} - \boldsymbol{\psi}(\boldsymbol{a})\Big|_\alpha^\beta \tag{17}$$

Comparing equations (11) and (17), we arrive at two simple definitions: a) when the right side of equation (17) vanishes, the scheme is entropy conservative; b) when the right side is less than zero, the scheme is entropy stable. equation (17) prompts us to introduce two point functions.

Lemma 2.1 (Harten [30], Osher [57], Tadmor [61], Shu [43])**.** *Suppose $g_s(a(\alpha), a(\beta))$ is a scalar monotone function, non-decreasing in first argument and non-increasing in second argument. Then g_s is entropy stable.*

Proof. Let $a_\alpha = a(\alpha)$, $a_\beta = a(\beta)$, and so on. By the mean value theorem, there exists $a_\alpha < a_\gamma < a_\beta$ such that

$$\psi(a_\beta) - \psi(a_\alpha) = (a_\beta - a_\alpha)b(a_\gamma) = (a_\beta - a_\alpha)g(f(a_\gamma))$$

Since $f(a)$ is an increasing function, $f(a_\alpha) < f(a_\gamma) < f(a_\beta)$. By the monotonicity of g_s, we recover

$$(f_\beta - f_\alpha)\big(g_s(f_\alpha, f_\beta) - g(f_\gamma)\big) \le 0 \implies \big(\psi(a_\beta) - \psi(a_\alpha)\big) - (a_\beta - a_\alpha)g_s(f_\alpha, f_\beta) \le 0, \tag{18}$$

where $f_\square = f(a_\square)$. □

The idea can be generalized to systems [62, 8, 10] utilizing convexity. We will directly state the needed theorems.

Corollary 2.2 (Tadmor [61], Carpenter [8], Shu [10])**.** *A consistent, symmetric two-point numerical flux $\boldsymbol{g}_s(\boldsymbol{f}_\alpha, \boldsymbol{f}_\beta)$ is*

- *entropy conservative if*
$$(\boldsymbol{a}_\beta - \boldsymbol{a}_\alpha)^{\mathrm{T}}\,\boldsymbol{g}_s(\boldsymbol{f}_\alpha, \boldsymbol{f}_\beta) - \big(\boldsymbol{\psi}_\beta - \boldsymbol{\psi}_\alpha\big) = 0$$
- *entropy stable if*
$$(\boldsymbol{a}_\beta - \boldsymbol{a}_\alpha)^{\mathrm{T}}\,\boldsymbol{g}_s(\boldsymbol{f}_\alpha, \boldsymbol{f}_\beta) - \big(\boldsymbol{\psi}_\beta - \boldsymbol{\psi}_\alpha\big) \le 0$$
- *symmetric and consistent in the sense that $\boldsymbol{g}_s(\boldsymbol{f}(x), \boldsymbol{f}(y)) = \boldsymbol{g}(\boldsymbol{f}(x))$, and therefore*
$$\frac{\partial \boldsymbol{g}}{\partial x}(x) = \frac{\partial \boldsymbol{g}_s}{\partial x}(x, x) + \frac{\partial \boldsymbol{g}_s}{\partial y}(x, x) = 2\frac{\partial \boldsymbol{g}_s}{\partial y}(x, x).$$

where $\boldsymbol{a}_{\alpha/\beta}$ and $\boldsymbol{\psi}_{\alpha/\beta}$ are entropy variables and potential fluxes at the left (α) and right (β) states.

For scalar systems (c.f. (18)),

$$g_s = \begin{cases} \dfrac{\psi_\beta - \psi_\alpha}{a_\beta - a_\alpha} & f_\alpha \neq f_\beta \\ g(f_\alpha) & f_\alpha = f_\beta \end{cases},$$

is unique; whereas for multi-component systems, the equation is undetermined. It can be solved using eigenvalue decomposition (Theorem 6.1 of [62]), however is expensive in practice. Historically, such ideas have also appeared early for high order phase field equations [20, 18].

To put aforementioned ideas into Boltzmann equation context, we discretize (1) in velocity space.

Definition 2.3 (Discretization in velocity space)**.** *To discretize $\boldsymbol{v}$, we employ finite difference. Each velocity component v^i ($i \in \{1, \dots, 3\}$) is discretized uniformly with N points in the interval $[-L, L]$. The grid points, v_i, are chosen as $-L + (m - 1/2)\Delta v$, with $m = 1, \dots, N$ and $\Delta v = 2L/N$. We use $\boldsymbol{v}_j \in \mathbb{R}^3$ with $j = 1, \dots, N^3$ to denote the velocity grid point, and $\vartheta_j = \Delta v^3$ to denote the associated weight. For clarity of notation, let $N_v = N^3$. We assume that the velocity space is sufficiently resolved such that (6) holds discretely*[2]*.*

Using this, we define the state vector $\boldsymbol{f}$ with $f_j = f(t, \boldsymbol{x}, \boldsymbol{v}_j)$; and the flux vector $\boldsymbol{g}$ with $g_j = v_j\, f(t, \boldsymbol{x}, \boldsymbol{v}_j)$. Comparing (7) and (11), we get

$$A(f_j) = f_j \ln(f_j), \quad B(f_j) = v_j f_j \ln(f_j). \tag{19}$$

with

$$a_j = 1 + \ln(f_j), \quad \psi(f_j) = v_j f_j, \quad g_s(f_{j,\beta},\, f_{j,\alpha}) = \frac{v_j \big(f_{j,\beta} - f_{j,\alpha}\big)}{\ln(f_{j,\beta}) - \ln(f_{j,\alpha})}. \tag{20}$$

The statement of H-theorem (7) also contains integration over the velocity space. This will be enforced once we introduce time-discretization.

Remark 2.4. *(19, 20) utilize the "physical" entropy. However, our procedure is general. It is possible to find a family of such entropy pairs. We use (19, 20) for a concise demonstration of the idea in the present work.*

3. A semi-discrete high order entropy stable scheme in one-dimension

Equations (20) are non-linear. To enforce them at a discrete level, we need a specific class of numerical schemes. In this section, we describe the EDG spatial discretization schemes that will help us construct entropy stable scheme for Boltzmann equation. We prefer DG because most of the problems of relevance are *multi*-dimensional by nature, for example turbulence is a three-dimensional phenomenon. Compared to high-order finite volume/difference methods, DG provides easy formulation on arbitrary meshes, high-order flux reconstruction, straightforward implementation of boundary conditions, as well as *linear* scaling on parallel processors due to the compactness of the scheme. Historically speaking, *classical* DG schemes for Boltzmann equation have been proposed in [40], extended to multi-species in [41], and implemented in [42, Table 3] with a parallel efficiency close to 99%. Recent works in this direction includes construction of DG schemes for Lattice-Boltzmann system [26]; *h/p* adaptive modal/nodal DG [39] and isogeometric [38] schemes for full Boltzmann equation and relaxation-type kinetic models; DG schemes for Grad's moment system [64]; dynamical low-rank methods for kinetic equation [19] and non-linear Boltzmann equation [35]; and a mode cut-off strategy for solving high-speed problems [36].

First, we, recall that DG methods assume that solution within an element of the discretized domain, may be expressed as a polynomial. Therefore, all the operations in the physical space, including, gradient, divergence, curl, are element local operators. For example, derivative of a function is simply the matrix-vector product of the coefficients of the polynomial and the derivative of the basis of the polynomial. Because, there is no universal basis for polynomials, there exists a wide class of DG methods [11, 45, 34, 8], each with different properties. A certain class of DG operators can be constructed to mimic summation by part property [27] at a *discrete* level. Due to this property, these class of DG schemes can be modified to fulfill an *arbitrary* entropy condition, while maintaining conservation and high order accuracy.

3.1. Summation by parts.

Consider a reference element, $I = [-1, 1]$, associated with Gauss-Legendre-Lobatto (GLL) quadrature points

$$-1 = \xi_0 < \xi_1 \cdots \xi_p = 1 \tag{21}$$

[2]We use a Fourier-spectral method (see [40] for details) for evaluating the collision operator. These methods rely on convolutive-decomposition of the collision operator. The resulting convolutions are evaluated using fast Fourier-transform (FFT). FFT is naturally performed (and faster) on uniform grids. Note that, one can, in-principle, use non-uniform velocity grid and/or possibly other approaches for evaluating the collision operator.

and the quadrature weights $\{w_r\}_{r=0}^{p}$. Define the Lagrange interpolating polynomials,

$$\ell_r(\xi) := \prod_{\substack{q=0 \\ q\neq r}}^{p} \frac{\xi - \xi_q}{\xi_r - \xi_q}, \tag{22}$$

such that $\ell_r(\xi_q) = \delta_{rq}$ where δ is the standard Kronecker delta function. Then, the discrete inner product of two functions f and g are

$$(f, g) = \int_I f\, g \,\mathrm{d}\xi \approx \sum_{q=0}^{p} f(\xi_q)\, g(\xi_q)\, w_q. \tag{23}$$

It can be shown that for any $l_1, l_2 \in \{\ell_0, \ldots, \ell_p\}$, integration-by-parts yields

$$\int_I l_1 \frac{\mathrm{d}l_2}{\mathrm{d}\xi}\mathrm{d}\xi + \int_I l_2 \frac{\mathrm{d}l_1}{\mathrm{d}\xi}\mathrm{d}\xi = l_1 l_2\Big|_{-1}^{1} \tag{24}$$

The nodal mass matrix, difference matrix, and stiffness matrix, respectively, are defined as

$$M_{rs} = \int_I \ell_r(\xi)\, \ell_s(\xi)\,\mathrm{d}\xi \approx \delta_{rq}\delta_{sq}\, w_q, \quad D_{rq} = \ell_q'(\xi_r), \quad S_{rs} = \int_I \ell_r(\xi)\, \ell_s'(\xi)\,\mathrm{d}\xi. \tag{25}$$

For convenience, we define a diagonal boundary matrix

$$B = \mathrm{diag}\begin{pmatrix} -1 & 0 & \ldots & 0 & 1 \end{pmatrix}^{\mathrm{T}}. \tag{26}$$

equation (24) holds for any polynomial assuming a sufficiently accurate quadrature rule, or under analytical integration. The advantage of the Lagrange polynomial constructed on GLL quadrature is that equation (24) holds discretely.

Lemma 3.1 (Gassner [27])**.** *The aforementioned matrices satisfy the following properties*

- *Discrete summation by parts*

$$S = MD, \quad MD + D^{\mathrm{T}}M = S + S^{\mathrm{T}} = B. \tag{27}$$

- *For each* $0 \le r \le p$ *we have*

$$\sum_{q=0}^{p} D_{rq} = \sum_{q=0}^{p} S_{rq} = 0, \quad \sum_{q=0}^{p} S_{qr} = \tau_r = \begin{cases} -1 & r = 0 \\ 1 & r = p \\ 0 & 1 \le r \le p-1. \end{cases} \tag{28}$$

A general construction procedure can be found in [21].

3.2. **A semi-discrete DG scheme for Boltzmann equation.** Let us first return to equation (10) and add the collision operator $\mathcal{Q}(\boldsymbol{f})$ to the right side. Our starting point is the classical DG scheme [11]. Given a domain decomposition

$$x_{1/2} < x_{3/2} < \cdots < x_{N_e+1/2}, \quad I_e = [x_{e-1/2},\, x_{e+1/2}], \quad \Delta x_e = x_{e+1/2} - x_{e-1/2}$$

and the discrete DG space of orthonormal Legendre polynomial of degree p

$$\boldsymbol{V}_h^p = \{\boldsymbol{\phi}_h : \boldsymbol{\phi}_h \,|\, I_e \in [P_p(I_e)]^j,\ 1 \le e \le N_e\}$$

we seek $\boldsymbol{f}_h \in \boldsymbol{V}_h^p$ such that for each $\boldsymbol{\phi}_h \in \boldsymbol{V}_h^p$ and $1 \le e \le N_e$,

$$\int_{I_e} \frac{\partial \boldsymbol{f}_h}{\partial t}\boldsymbol{\phi}_h \mathrm{d}x - \int_{I_e} \boldsymbol{g}(\boldsymbol{f}_h)^{\mathrm{T}} \frac{\mathrm{d}\boldsymbol{\phi}_h}{\mathrm{d}x}\mathrm{d}x = -\hat{\boldsymbol{g}}_{e+1/2}^{\mathrm{T}}\boldsymbol{\phi}_h(x_{e+1/2}^-) + \hat{\boldsymbol{g}}_{e-1/2}^{\mathrm{T}}\boldsymbol{\phi}_h(x_{e-1/2}^+) + \frac{1}{\mathrm{Kn}}\int_{I_e} \mathcal{Q}(\boldsymbol{f}_h, \boldsymbol{f}_h)\boldsymbol{\phi}_h \,\mathrm{d}x \tag{29}$$

where $\hat{\boldsymbol{g}}_{e+1/2}$ is a single-valued numerical flux at the element interface, depending on the values of numerical solution from both sides

$$\hat{\boldsymbol{g}}_{e+1/2} = \hat{\boldsymbol{g}}\big(\boldsymbol{f}_h(x_{e+1/2}^-),\ \boldsymbol{f}_h(x_{e+1/2}^+)\big) \tag{30}$$

equation (29) is usually called the weak form. We obtain the strong form after a simple integration by parts

$$\int_{I_e} \frac{\partial \boldsymbol{f}_h}{\partial t}\boldsymbol{\phi}_h \mathrm{d}x - \int_{I_e} \boldsymbol{g}(\boldsymbol{f}_h)^{\mathrm{T}} \frac{\mathrm{d}\boldsymbol{\phi}_h}{\mathrm{d}x}\mathrm{d}x = -\left(\boldsymbol{g}(\boldsymbol{f}_h(x^-_{e+1/2})) - \hat{\boldsymbol{g}}_{e+1/2}\right)^{\mathrm{T}} \boldsymbol{\phi}_h(x^-_{e+1/2}) + \left(\boldsymbol{g}(\boldsymbol{f}_h(x^+_{e-1/2})) - \hat{\boldsymbol{g}}_{e-1/2}\right)^{\mathrm{T}} \boldsymbol{\phi}_h(x^+_{e-1/2}) + \frac{1}{\mathrm{Kn}}\int_{I_e} \mathcal{Q}(\boldsymbol{f}_h, \boldsymbol{f}_h)\boldsymbol{\phi}_h \,\mathrm{d}x. \tag{31}$$

By the change of variables between I_e and the reference element $I = [-1, 1]$

$$x_e(\xi) = x_{e-1/2}\frac{(1-\xi)}{2} + x_{e+1/2}\frac{(1+\xi)}{2} \tag{32}$$

the weak form, equation (29), on I is

$$\frac{\Delta x_e}{2}\int_I \frac{\partial \boldsymbol{f}_h}{\partial t}\boldsymbol{\phi}_h \mathrm{d}\xi - \int_I \boldsymbol{g}(\boldsymbol{f}_h)^{\mathrm{T}} \frac{\mathrm{d}\boldsymbol{\phi}_h}{\mathrm{d}\xi}\mathrm{d}\xi = -\hat{\boldsymbol{g}}^{\mathrm{T}}_{e+1/2}\boldsymbol{\phi}_h(x_e(1)) + \hat{\boldsymbol{g}}^{\mathrm{T}}_{e-1/2}\boldsymbol{\phi}_h(x_e(-1)) + \frac{1}{\mathrm{Kn}}\int_I \mathcal{Q}(\boldsymbol{f}_h, \boldsymbol{f}_h)\boldsymbol{\phi}_h \,\mathrm{d}\xi. \tag{33}$$

In vector notations, equation (33) can be recast as

$$\frac{\Delta x_e}{2}\boldsymbol{M}\frac{\mathrm{d}\vec{\boldsymbol{f}}^e}{\mathrm{d}t} - \boldsymbol{S}^{\mathrm{T}}\vec{\boldsymbol{g}}^e = -\boldsymbol{B}\vec{\boldsymbol{g}}^e_\dagger + \frac{\Delta x_e}{2}\boldsymbol{M}\mathcal{Q}\big((\vec{\boldsymbol{f}}^e)^{\mathrm{T}}, (\vec{\boldsymbol{f}}^e)^{\mathrm{T}}\big)^{\mathrm{T}}, \tag{34}$$

where

$$\vec{\boldsymbol{f}}^e = \begin{bmatrix} \boldsymbol{f}_h(x_e(\xi_0)) \\ \vdots \\ \boldsymbol{f}_h(x_e(\xi_k)) \end{bmatrix}, \quad \vec{\boldsymbol{g}}^e = \begin{bmatrix} \boldsymbol{g}_h(x_e(\xi_0)) \\ \vdots \\ \boldsymbol{g}_h(x_e(\xi_k)) \end{bmatrix}, \quad \vec{\boldsymbol{g}}^e_\dagger = \begin{bmatrix} \hat{\boldsymbol{g}}_{e-1/2} \\ 0 \\ \vdots \\ 0 \\ \hat{\boldsymbol{g}}_{e+1/2} \end{bmatrix},$$

$$\boldsymbol{M} = M \otimes I_{N_v}, \quad \boldsymbol{D} = D \otimes I_{N_v}, \quad \boldsymbol{S} = S \otimes I_{N_v}, \quad \boldsymbol{B} = B \otimes I_{N_v}. \tag{35}$$

Here $\vec{\boldsymbol{f}}^e$, $\vec{\boldsymbol{g}}^e$, $\vec{\boldsymbol{g}}^e_\dagger$, each, are matrices of dimension $N_p \times N_v$; I_{N_v} is an identity matrix of size $N_v \times N_v$; and $\boldsymbol{M}$, $\boldsymbol{D}$, $\boldsymbol{S}$, $\boldsymbol{B}$, each, are matrices of dimension $N_p \times N_p$. Note that $\mathcal{Q}$ is an operator which is local in the physical space, so it takes a row of $\vec{\boldsymbol{f}}^e$ as input, and outputs a row vector; which is the reason for using transpose on right side of equation (34). We may simplify equation (34) as

$$\frac{\mathrm{d}\vec{\boldsymbol{f}}^e}{\mathrm{d}t} - \frac{2}{\Delta x_e}(\boldsymbol{M}^{-1}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{M})\vec{\boldsymbol{g}}^e = -\frac{2}{\Delta x_e}\boldsymbol{M}^{-1}\boldsymbol{B}\vec{\boldsymbol{g}}^e_\dagger + \frac{1}{\mathrm{Kn}}\mathcal{Q}\big((\vec{\boldsymbol{f}}^e)^{\mathrm{T}}, (\vec{\boldsymbol{f}}^e)^{\mathrm{T}}\big)^{\mathrm{T}}. \tag{36}$$

Similarly the strong form equation (31) can be recast in vector notations as

$$\frac{\mathrm{d}\vec{\boldsymbol{f}}^e}{\mathrm{d}t} + \frac{2}{\Delta x_e}\boldsymbol{D}\vec{\boldsymbol{g}}^e = \frac{2}{\Delta x_e}\boldsymbol{M}^{-1}\boldsymbol{B}(\vec{\boldsymbol{g}}^e - \vec{\boldsymbol{g}}^e_\dagger) + \frac{1}{\mathrm{Kn}}\mathcal{Q}\big((\vec{\boldsymbol{f}}^e)^{\mathrm{T}}, (\vec{\boldsymbol{f}}^e)^{\mathrm{T}}\big)^{\mathrm{T}}. \tag{37}$$

An important takeaway is that non-linear operators, local to the physical space, are evaluated utilizing the Kronecker delta property of the basis. This idea results in the fastest-known discontinuous Galerkin solvers for non-linear Boltzmann equation [42].

3.3. **A semi-discrete entropy stable scheme for Boltzmann equation on single element.** We begin with a description of an entropy stable scheme for Boltzmann equation on a single element with periodic or compactly supported boundary condition. First, we recast equation (1) in terms of entropy flux following section (2) as

$$\boldsymbol{f}_t + 2(\boldsymbol{g}_s)_x = \mathcal{Q}(\boldsymbol{f}, \boldsymbol{f}) \tag{38}$$

Utilizing the idea put forth in the previous section, this can be discretized as (dropping e for simplicity)

$$\frac{\mathrm{d}\boldsymbol{f}_r}{\mathrm{d}t} + \sum_{q=0}^{p}\frac{4}{\Delta x}D_{rq}\,\boldsymbol{g}_s(\boldsymbol{f}_r, \boldsymbol{f}_q) = \frac{2}{\Delta x}\frac{\tau_r}{w_r}(\boldsymbol{g}_r - \boldsymbol{g}_{\dagger,r}) + \frac{1}{\mathrm{Kn}}\mathcal{Q}\big(\boldsymbol{f}_r, \boldsymbol{f}_r\big), \quad r = 0, \ldots, p. \tag{39}$$

This yields the following set of *discrete* conservation laws.

Theorem 3.2. *The scheme* (39) *conserves mass, momentum, and energy i.e.,*

$$\int_{\Omega_{\boldsymbol{x}}}\int_{\mathbb{R}^3} \boldsymbol{f}\begin{pmatrix} 1 \\ \boldsymbol{v} \\ |\boldsymbol{v}|^2 \end{pmatrix}\mathrm{d}\boldsymbol{v}\,\mathrm{d}x \text{ are independent of time.}$$

which in discrete form reads

$$\sum_j \sum_{r=0}^{p} f_{j,r} \begin{pmatrix} 1 \\ \boldsymbol{v}_j \\ |\boldsymbol{v}_j|^2 \end{pmatrix} \vartheta_j \, w_r \quad \textit{are independent of time.} \tag{40}$$

Proof. Let $\Lambda_j = \begin{pmatrix} 1 & \boldsymbol{v}_j & |\boldsymbol{v}_j|^2 \end{pmatrix}^{\mathrm{T}}$, then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Big(\sum_{r,j} f_{j,r} \, \Lambda_j \, \vartheta_j \, w_r \Big) = & \\
& \frac{2}{\Delta x} \sum_{r,j} \tau_r \big(g_{j,r} - g_{\dagger,j,r} \big) \, \Lambda_j \, \vartheta_j - \frac{4}{\Delta x} \sum_{r,j,q} D_{rq} g_s(f_{j,r}, f_{j,q}) \, \Lambda_j \, \vartheta_j \, w_r \\
& + \frac{1}{\mathrm{Kn}} \sum_{r,j} \mathcal{Q}(f_{\cdot,r}, \, f_{\cdot,r}) \, \Lambda_j \, \vartheta_j \, w_r && \text{(collision invariants (6))} \\
= & \frac{2}{\Delta x} \sum_{r,j} \tau_r \big(g_{j,r} - g_{\dagger,j,r} \big) \, \Lambda_j \, \vartheta_j - \frac{4}{\Delta x} \sum_{r,j,q} (M D_{rq}) \; g_s(f_{j,r}, f_{j,q}) \, \Lambda_j \, \vartheta_j \\
= & \frac{2}{\Delta x} \sum_{r,j} \tau_r \big(g_{j,r} - g_{\dagger,j,r} \big) \, \Lambda_j \, \vartheta_j - \frac{2}{\Delta x} \sum_{r,j,q} (M D_{rq} + D_{qr} M) \; g_s(f_{j,r}, f_{j,q}) \, \Lambda_j \, \vartheta_j \\
& && \text{(symmetry)} \\
= & \frac{2}{\Delta x} \sum_{r,j} \tau_r \big(g_{j,r} - g_{\dagger,j,r} \big) \, \Lambda_j \, \vartheta_j - \frac{2}{\Delta x} \sum_{r,j} \tau_r \, g(f_{j,r}) \, \Lambda_j \, \vartheta_j \\
& && \text{(summation by parts (27,28))} \\
= & \frac{2}{\Delta x} \sum_{j} \big(g_{\dagger,j,k} - g_{\dagger,j,0} \big) \, \Lambda_j \, \vartheta_j && \text{(local conservation)} \\
= & \, 0 \\
& && \text{(compactly supported boundary, single element, global conservation)}
\end{aligned}$$

□

Theorem 3.3. *Let $f(\boldsymbol{v}') = f'$, $f(\boldsymbol{v}'_*) = f'_*$, $f(\boldsymbol{v}_*) = f_*$. The scheme, equation* (39)*, satisfies the entropy identity*

$$\int_{\Omega_{\boldsymbol{x}}} \int_{\mathbb{R}^3} \partial_t \big(f \ln(f) \big) \, \mathrm{d}\boldsymbol{v} = -\frac{1}{4} \frac{1}{\mathrm{Kn}} \int_{\Omega_{\boldsymbol{x}}} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{S}^2} \mathcal{B}[f' f'_* - f f_*] \, \ln\Big(\frac{f' f'_*}{f f_*} \Big) \, \mathrm{d}\boldsymbol{\omega} \, \mathrm{d}\boldsymbol{v}_* \, \mathrm{d}\boldsymbol{v},$$

which in discrete form reads

$$\sum_{j,r} \frac{\mathrm{d}}{\mathrm{d}t} \Big(f_{j,r} \, \ln(f_{j,r}) \Big) \, \vartheta_j \, w_r + \frac{1}{4} \sum_{r,j} \vartheta_j \, w_r \int_{\mathbb{R}^3} \int_{\mathbb{S}^2} \mathcal{B}[f'_{\cdot,r} f'_{*,\cdot,r} - f_{\cdot,r} f_{*,\cdot,r}] \, \ln\Big(\frac{f'_{\cdot,r} f'_{*,\cdot,r}}{f_{\cdot,r} f_{*,\cdot,r}} \Big) \, \mathrm{d}\boldsymbol{\omega} \, \mathrm{d}\boldsymbol{v}_* = 0. \tag{41}$$

Proof. Multiply equation (39) by $a_{j,r} = (1 + \ln(f_{j,r}))$ we recover

$$\sum_{j,r} a_{j,r} \frac{\mathrm{d} f_{j,r}}{\mathrm{d}t} \, \vartheta_j \, w_r + \sum_{r,j,q} \frac{4}{\Delta x} a_{j,r} \, D_{rq} \, g_s(f_{j,r}, f_{j,q}) \, \vartheta_j \, w_r = \frac{2}{\Delta x} \sum_{j,r} \tau_r \, a_{j,r} (g_{j,r} - g_{\dagger,j,r}) \, \vartheta_j + \frac{1}{\mathrm{Kn}} \sum_{j,r} a_{j,r} \, \mathcal{Q} \, \vartheta_j \, w_r$$

Term 1:

$$\sum_{j,r} a_{j,r} \frac{\mathrm{d} f_{j,r}}{\mathrm{d}t} \, \vartheta_j \, w_r = \sum_{j,r} (1 + \ln(f_{j,r})) \frac{\mathrm{d} f_{j,r}}{\mathrm{d}t} \, \vartheta_j \, w_r = \sum_{j,r} \frac{\mathrm{d}}{\mathrm{d}t} \Big(f_{j,r} \, \ln(f_{j,r}) \Big) \, \vartheta_j \, w_r$$

Term 2:

$$\begin{aligned}\sum_{r,j,q}\frac{4}{\Delta x}a_{j,r}\,D_{rq}\,g_s(f_{j,r},f_{j,q})\,\vartheta_j\,w_r &= \frac{2}{\Delta x}\sum_{r,j,q}a_{j,r}\,(D_{rq}-D_{qr}+(M^{-1}B)_{rq})\,g_s(f_{j,r},f_{j,q})\,\vartheta_j\,w_r\\ &\qquad\qquad\text{(summation by parts (27))}\\ &= \frac{2}{\Delta x}\sum_{r,j}a_{j,r}\,\tau_r\,g_{j,r}\,\vartheta_j+\frac{2}{\Delta x}\sum_{r,j,q}D_{rq}\,(a_{j,r}-a_{j,q})g_s(f_{j,r},f_{j,q})\,\vartheta_j\,w_r\\ &= \frac{2}{\Delta x}\sum_{r,j}a_{j,r}\,\tau_r\,g_{j,r}\,\vartheta_j+\frac{2}{\Delta x}\sum_{r,j,q}D_{rq}\,(\psi_{j,r}-\psi_{j,q})\,\vartheta_j\,w_r\\ &= \frac{2}{\Delta x}\sum_{r,j}\tau_r\,(a_{j,r}\,g_{j,r}-\psi_{j,r})\,\vartheta_j\end{aligned}$$

Term 4:

$$\sum_{r,j}a_{j,r}\mathcal{Q}(f_{\cdot,r},f_{\cdot,r})\,\vartheta_j\,w_r=-\frac{1}{4}\sum_{r,j}\vartheta_j\,w_r\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\mathcal{B}[f'_{\cdot,r}f'_{*,\cdot,r}-f_{\cdot,r}f_{*,\cdot,r}]\,\ln\Big(\frac{f'_{\cdot,r}f'_{*,\cdot,r}}{f_{\cdot,r}f_{*,\cdot,r}}\Big)\,\mathrm{d}\boldsymbol{\omega}\,\mathrm{d}\boldsymbol{v}_*$$

(weak form (9))

Subtracting Term 3 from Term 2, we recover

$$\begin{aligned}\frac{2}{\Delta x}\sum_{r,j}\tau_r\,(a_{j,r}\,g_{j,r}-\psi_{j,r})\,\vartheta_j-\frac{2}{\Delta x}\sum_{j,r}\tau_r\,a_{j,r}(g_{j,r}-g_{\dagger,j,r})\,\vartheta_j &= \frac{2}{\Delta x}\sum_{r,j}\tau_r\,(a_{j,r}\,g_{\dagger,j,r}-\psi_{j,r})\,\vartheta_j\\ &= \frac{2}{\Delta x}\sum_{j}\Big((\psi_{j,p}-a_{j,p}\,g_{\dagger,j,k})-(\psi_{j,0}-a_{j,0}\,g_{\dagger,j,0})\Big)\,\vartheta_j\\ &\qquad\qquad\text{(local conservation)}\\ &= 0\\ &\qquad\qquad\text{(global conservation if } g_{\dagger,j}\text{ is entropy conservative)}\end{aligned}$$

Utilizing these terms, we recover equation (41). □

3.4. **A semi-discrete entropy stable discrete scheme for Boltzmann equation on multiple elements.** The last step of theorem (3.3) requires $g_{\dagger,j}$ to be entropy conservative. One may let $g_{\dagger,j}$ be the entropy conservative flux, g_s, yielding an entropy conservative scheme. However, entropy should be dissipated at shock waves and entropy conservative schemes will produce strong oscillations near shocks (c.f. Fig. (4) in [63]). Dispersive oscillations on the *mesh scale* are observed due to the absence of any dissipation mechanism [50, 52] since entropy vanishes at discontinuities. To this end, we may utilize the upwind numerical fluxes which are well-known to be entropy *stable*. Upwinding works very well for kinetic systems and has been central to construction of many kinetic schemes [59, 44]. Since we are dealing with constant advection system, the upwind flux can be simplified as

$$\hat{g}_{e-1/2,j}=\begin{cases}g(f_j(x^+_{e-1/2})) & \boldsymbol{v}_j\cdot\boldsymbol{n}<0\\ g(f_j(x^-_{e-1/2})) & \boldsymbol{v}_j\cdot\boldsymbol{n}\geq 0\end{cases},\quad e=1,\ldots,N_e, \tag{42}$$

where $\boldsymbol{n}$ is the interface normal. Alternatively, one may add an upwind term to the entropy-conservative flux [37]

$$\hat{g}_{e-1/2,j}=g_s(f_j(x^+_{e-1/2}),\,f_j(x^-_{e-1/2}))-\frac{1}{2}|\boldsymbol{v}_j|\big(f_j(x^+_{e-1/2})-f_j(x^-_{e-1/2})\big),\quad e=1,\ldots,N_e, \tag{43}$$

so that entropy is dissipated at shock waves.

The entropy stable scheme on multiple elements reads as

$$\frac{\mathrm{d}f^e_{j,r}}{\mathrm{d}t}+\sum_{q=0}^{p}\frac{4}{\Delta x_e}D_{rq}\,g_s(f^e_{j,r},f^e_{j,q})=\frac{2}{\Delta x_e}\frac{\tau_r}{w_r}(g^e_{j,r}-\hat{g}^e_{\dagger,j,r})+\frac{1}{\mathrm{Kn}}\mathcal{Q}(f^e_{j,r},\,f^e_{j,r}). \tag{44}$$

It can be recast in a conservative form along the lines of [62, 8, 10]

$$\frac{\mathrm{d}f^e_{j,r}}{\mathrm{d}t} + \frac{2}{\Delta x_e}\frac{1}{w_r}\big(g^e_{j,r+1/2} - g^e_{j,r-1/2}\big) = \mathcal{Q}(f^e_{\cdot,r},\, f^e_{\cdot,r})$$

$$g^e_{j,r+1/2} = \begin{cases} \hat{g}_{e-1/2,j} & r = -1 \\ \hat{g}_{e+1/2,j} & r = p \\ 2\sum_{p=0}^{r}\sum_{q=r+1}^{p} S_{pq}\, g_s(f^e_{j,p},\, f^e_{j,q}) & 0 \le r \le p-1 \end{cases}. \tag{45}$$

The entropy *stability* on multiple elements reads as

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(f^e_{j,r}\,\ln(f^e_{j,r})\big) + \sum_{q=0}^{p}\frac{4}{\Delta x_e} D_{rq}\, a^e_{j,r}\, g_s(f^e_{j,r}, f^e_{j,q}) - \frac{2}{\Delta x_e}\frac{\tau_r}{w_r} a^e_{j,r}\,(g^e_{j,r} - \hat{g}^e_{\dagger,j,r}) \le 0, \tag{46}$$

and can be also recast in a conservative form as

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(f^e_{j,r}\,\ln(f^e_{j,r})\big) + \frac{2}{\Delta x_e}\frac{1}{w_r}\big(G^e_{j,r+1/2} - G^e_{j,r-1/2}\big) \le 0$$

$$G^e_{j,r+1/2} = \begin{cases} \frac{1}{2}\big((a^{e-1}_{j,p} + a^e_{j,0})\hat{g}_{e-1/2,j} - (\psi^{e-1}_{j,p} + \psi^e_{j,0})\big) & r = -1 \\ \frac{1}{2}\big((a^e_{j,p} + a^{e+1}_{j,0})\hat{g}_{e+1/2,j} - (\psi^e_{j,p} + \psi^{e+1}_{j,0})\big) & r = p \\ \sum_{p=0}^{r}\sum_{q=r+1}^{p} S_{pq}\,\big((a^e_{j,p} + a^e_{j,q})\, g_s(f^e_{j,p},\, f^e_{j,q}) - (\psi^e_{j,p} + \psi^e_{j,q})\big) & 0 \le r \le p-1 \end{cases}. \tag{47}$$

To enforce the velocity averaging in equation (47), we need to fully-discretize the scheme (to be described later).

3.5. **A positivity-preserving limiter for Boltzmann equation.** The probability distribution function (PDF), f, is naturally non-negative. A discrete scheme must ensure that its values are within the physical bounds. A simple positivity preserving limiter can be constructed along the lines of [67, 68] as

- Compute the element average as

$$\bar{f}^e_j = \sum_r f^e_{j,r}\, w_r/2$$

- Define a lower bound $\varepsilon = \min_j(10^{-13},\, \bar{f}_j)$
- In each element, modify the probability distribution as

$$\breve{f}^e_{j,r} = \bar{f}^e_j + \theta_1(f^e_{j,r} - \bar{f}^e_j), \qquad \theta_1 = \min\left\{\left|\frac{\bar{f}^e_j - \varepsilon}{\bar{f}^e_j - (f^e_j)_{\min}}\right|, 1\right\}, \quad (f^e_j)_{\min} = \min_r f^e_{j,r}.$$

It has been shown in [67] that this procedure preserves the formal accuracy of the scheme.

3.6. **A slope limiter for Boltzmann equation.** Positivity-preserving limiter enforces the physical bound of the PDF. We note that the entropy stable schemes are stable in the presence of shocks. However, the solution may still contain oscillations. A slope limiter can be used for damping spurious oscillations. The procedure can be defined as follows [10]:

- Compute the element average as

$$\bar{f}^e_j = \sum_r f^e_{j,r}\, w_r/2$$

- In each element, modify the probability distribution as

$$\breve{f}^e_{j,0} = \bar{f}^e_j + m(f^e_{j,0} - \bar{f}^e_j,\; \bar{f}^e_j - \bar{f}^{e+1}_j,\; \bar{f}^{e-1}_j - \bar{f}^e_j), \quad \breve{f}^e_{j,p} = \bar{f}^e_j + m(f^e_{j,p} - \bar{f}^e_j,\; \bar{f}^{e+1}_j - \bar{f}^e_j,\; \bar{f}^e_j - \bar{f}^{e-1}_j),$$

$$\breve{f}^e_{j,r} = \bar{f}^e_j + \theta_2(f^e_{j,r} - \bar{f}^e_j), \quad \theta_2 = \frac{(\breve{f}^e_{j,0} - \bar{f}^e_j) + (\breve{f}^e_{j,p} - \bar{f}^e_j)}{(f^e_{j,0} - \bar{f}^e_j) + (f^e_{j,p} - \bar{f}^e_j)}, \quad r = \{1, \ldots, p-1\},$$

$$m(a_1, \ldots, a_n) = \begin{cases} s\min_{1\le i\le n}|a_e|, & |s| = 1, \\ 0 & \text{otherwise} \end{cases}, \quad s = \frac{1}{n}\sum_{i=1}^{n}\mathrm{sign}(a_e) \tag{48}$$

- Reconstruct the solution as $f^e_j(\xi_q) = f^e_{j,r}$.

Slope limiters are known to destroy the formal accuracy of the scheme. For entropy stable schemes, the limiters are generally not required. Nonetheless, slope limiters will improve the robustness of the scheme. In non-ideal real-world non-smooth industrial level simulations, a numerical scheme must not break down at any cost. The limiter, equation (48), *provably* satisfies the total variation diminishing (TVD) property for constant advection systems [11].

3.7. **A high-order slope limiter for Boltzmann equation.** (48) is TVD, however it may destroy the formal accuracy of the scheme. To maintain the formal accuracy of the scheme, at the cost of TVD property, we may use a high order moment limiter [5, 33].

First, we recall that a central point in the construction of discontinuous Galerkin schemes is the equivalence relation [45, 34]

$$f_j^e(\xi) = \sum_r^p \hat{f}_{j,r}^e\, \phi_r(\xi) = \sum_q^p f_j^e(\xi_q)\, \ell_r(\xi), \quad \phi \in \boldsymbol{V}_h^p,$$

where $\hat{f}_{j,r}^e$ are called as modal values (polynomial coefficients), and $f_j^e(\xi_q)$ are called nodal values (they have a physical meaning that they represent the approximate solution at the point ξ_q since $\ell_{rq} = \delta_{rq}$). Mathematically speaking, $f_j^e(\xi)$ is the ground truth, and that should be independent of the projection space.

The high-order slope limiting procedure can be described as follows

- Compute the modal values $\hat{f}_{j,r}^e$. To do this evaluate the basis at the Gauss-Legendre-Lobatto quadrature points. Let this basis be defined as a matrix, $\mathcal{V} = \{\phi_0, \dots, \phi_p\}$, of dimension $\mathbb{R}^{p+1} \times \mathbb{R}^{p+1}$ such that the polynomial modes are placed as columns of the matrix. Then $\hat{f}_{j,r}^e = \mathcal{V}^{-1} f_j^e(\xi_q)$
- In each element, modify the probability distribution as

$$\breve{f}_{j,r}^e = m(\hat{f}_{j,r}^e,\ \theta_1(\hat{f}_{j,r-1}^{e+1} - \hat{f}_{j,r-1}^e),\ \theta_1(\hat{f}_{j,r-1}^e - \hat{f}_{j,r-1}^{e-1})), \quad \kappa = \sqrt{(2r+1)(2r+3)}, \quad r = \{p, \dots, 1\}, \tag{49}$$

  where $\theta_1 \sim O(\kappa^{-1})$ is a free parameter. Note that the limiting procedure is done from top-to-bottom. If the top mode is not limited i.e., the m function returns the first argument, we do not limit the lower mode. This is the reason that the solution retains as high order accuracy as possible. The constant κ comes from the fact that orthonormal Legendre polynomials, our basis functions, satisfy a recurrence relation

$$\sqrt{2p+1}\, P_p = \frac{1}{\sqrt{2p+3}} P'_{p+1} - \frac{1}{\sqrt{2p-1}} P'_{p-1}.$$

- Reconstruct the solution as $f_j^e(\xi_q) = \mathcal{V} \breve{f}_{j,r}^e$.

An intuition behind such high order limiters can be found in [46]. Simply speaking, equation (49) aims to control spurious growth in $\hat{f}_{j,r}$ — which is approximately the r^{th} derivative of the solution — by comparing them to the forward and backward differences of the $(r-1)^{th}$ derivative, which are alternative approximations to the r^{th} derivative. From a WENO viewpoint, equation (49) may be reinterpreted as smoothness indicators in hiding: it tries to estimate how smooth the solution and its derivatives are.

4. A fully-discrete positivity-preserving implicit-explicit high order entropy stable scheme

In this section, we describe a first-order implicit-explicit time-discretization for (38). Semi-discrete analysis is a crucial assumption [10]. A fully discrete entropy stability analysis is available for first-order time-integration schemes [51]. The entropy stability of high-order schemes equipped with high-order time-integration is still an open problem, particularly for non-linear systems.

4.1. **The implicit-explicit time discretization.** Denoting the distribution function at time-index l by f^l, and the transport operator by $\mathcal{T}(f)$, an implicit-explicit scheme [22, 15] for equation (1) or (38) reads

$$\begin{aligned} f^{(1)} &= f^l, \\ f^{(2)} &= f^l - \Delta t\, \mathcal{T}(f^{(1)}) + \frac{\Delta t}{\mathrm{Kn}}\Big[\mathcal{Q}(f^{(1)}, f^{(1)}) - \mathcal{P}(f^{(1)})\Big] + \frac{\Delta t}{\mathrm{Kn}}\mathcal{P}(f^{(2)}), \\ f^{l+1} &= f^{(2)}, \end{aligned} \tag{50}$$

where $\Delta t = t_{l+1} - t_l$ is time step, and $\mathcal{P}$ is the penalization operator.

In practical simulations, regions of flow are characterised by multiple scales (low and high Kn). As $\mathrm{Kn} \to 0$, the collision operator becomes stiff. Therefore, this scheme decomposes the Boltzmann equation into a stiff and non-stiff part, wherein the stiff part is treated implicitly, and the non-stiff part is treated explicitly so that the evolution is constrained only by the Courant–Friedrichs–Lewy condition. For the non-linear Boltzmann collision operators, this is achieved by penalizing the collision operator by a penalty function $\mathcal{P}(f)$ i.e.,

$$\underbrace{-\mathcal{T}(f)}_{\text{non stiff part}} + \underbrace{\frac{\mathcal{Q}(f,f)}{\mathrm{Kn}}}_{\text{stiff part}} = \underbrace{-\mathcal{T}(f) + \frac{\mathcal{Q}(f,f) - \mathcal{P}(f)}{\mathrm{Kn}}}_{\text{non-stiff part}} + \underbrace{\frac{\mathcal{P}(f)}{\mathrm{Kn}}}_{\text{stiff part}}. \tag{51}$$

In equation (50), the penalization operator acts implicitly on $f^{(2)}$. To develop an efficient iteration-free scheme, the penalization operator should be sufficiently simple, typically, linear, as far as the present literature is concerned. There is no fixed rule for selecting a penalization operator. The generally used criterion is to use an operator that preserves the asymptotic transition from the microscopic kinetic regime to the macroscopic fluid regime. In the present work, we utilize the BGK [2] operator (c.f. [39, Appendix] for other choices) i.e.,

$$\mathcal{P}(f) = \nu(\mathcal{M}[f] - f), \tag{52}$$

where $\mathcal{M}$ is the Maxwellian.

As per [15], ν is selected so that it is a valid estimate for the largest value of the negative term in the Boltzmann operator i.e.,

$$\nu \geq \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} B(|\boldsymbol{v} - \boldsymbol{v}_*|, \cos\chi)\, f(v_*)\, \mathrm{d}\boldsymbol{\omega}\, \mathrm{d}\boldsymbol{v}_*. \tag{53}$$

Remark 4.1. *We emphasize that there is no* precise *rule for selecting the optimal value of this parameter. We can simply utilize* $\nu = \max_{\Omega_{\boldsymbol{x}} \otimes \Omega_{\boldsymbol{v}}} \nu$. *This results in a stable scheme (a simple analogy follows from local vs global Lax–Friedrichs flux. The panalty is simply a method of adding dissipation). But, for the generality of presentation, we let* ν *to be a function of* $\Omega_{\boldsymbol{x}}$*: it is solely determined by the moments of probability distribution.*

equation (50) may nonetheless appear implicit because of presence of $\mathcal{P}(f^{(2)})$ in the second equation. To this end, one can utilize the properties of the BGK and Boltzmann collision operator. We first define

$$\boldsymbol{U}(f) = \int_{\mathbb{R}^3} \begin{pmatrix} 1 \\ \boldsymbol{v} \\ |\boldsymbol{v}|^2 \end{pmatrix} f \mathrm{d}\boldsymbol{v} = \begin{pmatrix} n \\ n\boldsymbol{u} \\ \frac{3}{2}nT + n|\boldsymbol{u}|^2 \end{pmatrix},$$

and recognize that

$$\boldsymbol{U}(\mathcal{M} - f) = \boldsymbol{U}(\mathcal{P}(f)) = \boldsymbol{U}(\mathcal{Q}(f,f)) = 0. \tag{54}$$

The implicit assumption behind such a construction is that (54) is satisfied at a discrete level as well.

Now, we compute the moments of (50), which yields

$$\begin{aligned} &\boldsymbol{U}(f^{(1)}) = \boldsymbol{U}(f^l), \\ &\boldsymbol{U}(f^{(2)}) = \boldsymbol{U}(f^l) - \Delta t \Big(\int_{\mathbb{R}^3} \begin{pmatrix} 1 \\ \boldsymbol{v} \\ |\boldsymbol{v}|^2 \end{pmatrix} \mathcal{T}(f^{(1)}) \mathrm{d}\boldsymbol{v} \Big), \\ &\boldsymbol{U}(f^{l+1}) = \boldsymbol{U}(f^{(2)}). \end{aligned} \tag{55}$$

Here all the three equations can be computed explicitly. Once, $\boldsymbol{U}(f^{(2)})$ is known i.e., $(n^{(2)}, n^{(2)}\boldsymbol{u}^{(2)}, \frac{3}{2}n^{(2)}T^{(2)} + n^{(2)}|\boldsymbol{u}^{(2)}|^2)$, one can use them to construct $\mathcal{M}(n^{(2)}, \boldsymbol{u}^{(2)}, T^{(2)}) \equiv \mathcal{M}(f^{(2)})$. Observe that when $\mathrm{Kn} \to 0$, $f \to \mathcal{M}(f)$, then equation (55) degenerates to a forward-Euler scheme for solving the compressible Euler equation [22]. The procedure of computing moments of equation (50) to recover equation (55) is innocuous, however it will ultimately enforce the velocity averaging required, for example, in equation (47) to achieve entropy stability.

The final scheme then reads:

$$\begin{aligned} &f^{(1)} = f^l, \\ &f^{(2)}\Big(1 + (\Delta t/\mathrm{Kn})\,\nu^{(2)}\Big) = f^l - \Delta t\, \mathcal{T}(f^{(1)}) + \frac{\Delta t}{\mathrm{Kn}}\Big[\mathcal{Q}(f^{(1)}, f^{(1)}) - \nu^{(1)}(\mathcal{M}(f^{(1)}) - f^{(1)})\Big] + \frac{\Delta t}{\mathrm{Kn}}\nu^{(2)}\mathcal{M}(f^{(2)}), \\ &f^{l+1} = f^{(2)}. \end{aligned} \tag{56}$$

This scheme is positivity-preserving (PP) provided spatial discretization and collision operator are positivity-preserving. Our spatial discretization is already positivity-preserving. For positivity-preserving collision operator, the direct simulation Monte Carlo (DSMC) method [3], the discrete velocity method [56], or the entropic Fourier method [7] can be used.

4.2. **Fully discrete scheme.** In a more compact form, the fully-discrete version of (56) reads

(57)
$$f_{j,r}^{l+1,e}\Big(1+\frac{\Delta t}{\mathrm{Kn}}\,\nu_r^{l+1,e}\Big)=f_{j,r}^{l,e}-\Delta t\,\mathcal{T}(f_{j,r}^{l,e})+\frac{\Delta t}{\mathrm{Kn}}\Big[\mathcal{Q}(f_{\cdot,r}^{l,e},f_{\cdot,r}^{l,e})-\nu_r^{l,e}(\mathcal{M}(f_{j,r}^{l,e})-f_{j,r}^{l,e})\Big]+\frac{\Delta t}{\mathrm{Kn}}\nu_r^{l+1,e}\mathcal{M}(f_{j,r}^{l+1,e}),$$

where, for example,

(58)
$$\mathcal{T}(f_{j,r}^{l,e})=\sum_{q=0}^{p}\frac{4}{\Delta x_e}D_{rq}\,g_s(f_{j,r}^{l,e},f_{j,q}^{l,e})-\frac{2}{\Delta x_e}\frac{\tau_r}{w_r}(g_{j,r}^{l,e}-\hat{g}_{\dagger,j,r}^{l,e}),$$

for one-dimensional physical space, and

(59)
$$\begin{pmatrix}n\\ n\boldsymbol{u}\\ \frac{3}{2}nT+n|\boldsymbol{u}|^2\end{pmatrix}_r^{l+1,e}=\begin{pmatrix}n\\ n\boldsymbol{u}\\ \frac{3}{2}nT+n|\boldsymbol{u}|^2\end{pmatrix}_r^{l,e}-\Delta t\sum_j\Lambda_j\mathcal{T}(f_{j,r}^{l,e})\,\vartheta_j.$$

To construct $\mathcal{M}(f_{j,r}^{l+1,e})$, we let

(60)
$$f_{j,r}^{*,e}=f_{j,r}^{l,e}-\Delta t\,\mathcal{T}(f_{j,r}^{l,e}),\quad \boldsymbol{U}(f_{j,r}^{l+1,e})=\sum_j\Lambda_j f_{j,r}^{*,e}\,\vartheta_j,\quad \mathcal{M}_{j,r}^{l+1}=\mathcal{M}[\boldsymbol{U}(f_{j,r}^{l+1,e})].$$

Similarly, $\nu_r^{l+1,e}$ can be constructed from $\boldsymbol{U}(f_{j,r}^{l+1,e})$ or $f_{j,r}^{*,e}$. Simply speaking, $\nu_r^{l+1,e}$ and $\boldsymbol{U}$ equip the system with an *apriori* estimate of mass, energy and entropy.

Remark 4.2. *$g_s(f_{j,r}^{l,e},f_{j,q}^{l,e})$ is numerically ill-posed when $f_{j,r}^{l,e}\to f_{j,q}^{l,e}$. See appendix of [37] for a numerical recipe.*

equation (59) is a system of conservation laws with entropy , $\int_{\mathbb{R}^3} f\ln(f)\mathrm{d}\boldsymbol{v}$, and entropy flux

(61)
$$\int_{\mathbb{R}^3}\Lambda\otimes\frac{\boldsymbol{v}(f_\alpha-f_\beta)}{\ln(f_\alpha)-\ln(f_\beta)}\mathrm{d}\boldsymbol{v},$$

where $\Lambda=\begin{pmatrix}1 & \boldsymbol{v} & |\boldsymbol{v}|\end{pmatrix}^{\mathrm{T}}$. Neither equation (59) nor (61) are closed. However, in the limit of vanishing Knudsen number, $f\to\mathcal{M}$, (59) becomes a consistent discretization to the limiting Euler system

(62)
$$\begin{cases}\partial_t n+\nabla_x\cdot(n\boldsymbol{u})=0,\\ \partial_t(n\boldsymbol{u})+\nabla_x\cdot(n\boldsymbol{u}\otimes\boldsymbol{u}+nT\,\mathrm{Id}/2)=0,\\ \partial_t E+\nabla_x\cdot\Big(\frac{5}{2}nT\boldsymbol{u}+n|\boldsymbol{u}|^2\,\boldsymbol{u}\Big)=0,\end{cases}$$

where $E=\frac{3}{2}nT+n|\boldsymbol{u}|^2$, and Id is an identity matrix; with entropy

(63)
$$n\ln\Big(\frac{n}{(\pi T)^{3/2}}\Big)-\frac{3}{2}n,$$

and the entropy flux

(64)
$$\int_{\mathbb{R}^3}\Lambda\frac{\theta_\alpha\exp(-\lambda_\alpha|\boldsymbol{v}-\boldsymbol{u}_\alpha|^2)\,\boldsymbol{v}}{C_1-\lambda_\alpha|\boldsymbol{v}-\boldsymbol{u}_\alpha|^2+\lambda_\beta|\boldsymbol{v}-\boldsymbol{u}_\beta|^2}\mathrm{d}\boldsymbol{v}-\int_{\mathbb{R}^3}\Lambda\frac{\theta_\beta\exp(-\lambda_\beta|\boldsymbol{v}-\boldsymbol{u}_\beta|^2)\,\boldsymbol{v}}{C_1-\lambda_\alpha|\boldsymbol{v}-\boldsymbol{u}_\alpha|^2+\lambda_\beta|\boldsymbol{v}-\boldsymbol{u}_\beta|^2}\mathrm{d}\boldsymbol{v},$$

where

$$C_1=\ln(\theta_\alpha)-\ln(\theta_\beta),\quad \theta=\frac{n}{(\pi T)^{3/2}},\quad \lambda=\frac{1}{T},\quad \mathcal{M}=\theta\exp(-\lambda|\boldsymbol{v}-\boldsymbol{u}|^2).$$

It is difficult to find a closed form analytical expression for equation (64). However, observe that equation (63) is consistent with the theory put forward in [30], and can be rewritten in terms of Euler thermodynamic entropy as

$$\frac{-n(s+(\ln(\pi)-1))}{\gamma-1},\quad s=\ln(p)-\gamma\ln(n),\quad \gamma=\frac{5}{3}.$$

Note that if $h(s) = (s + (\ln(\pi) - 1)))$, then $h' - \gamma h'' > 0$, $h' > 0$ [62]. Following the same argument put forward in section (3.3), it is easy to show that this limiting system is entropy-stable.

Remark 4.3. *The aforementioned discussion can be easily extended to structured quadrilateral domains using a tensor product construction. We will skip the details for brevity. We will directly present numerical tests in the following section.*

5. Numerical tests in one-dimensional physical space

In this section, we present verification for the scheme given by equation (57). We emphasize that in one-dimensional physical space, Boltzmann equation is a five-dimensional equation system (1 space, 3 velocity, and 1 time). These systems are computationally demanding. For example, many of the problems reported in this section involve degrees of freedom on the order of hundred millions. Due to the quadratic nature of the collision operator, these test cases involve floating point operations on the order of hundred billions at every time step. To appreciate the computational cost, recall that Boltzmann collision operator has complexity of $O(cN^3 \log(N)$, where c is algorithmic constant. This collision operator is called for each degree of freedom. Assume that our physical space consists of $N_e = 128$ elements, each with $k = 2$ order approximation. Furthermore assume that our velocity space consists of $N = 48$ points and let $c = 32$. Therefore, one requires floating point operations on order of $128 \times 3 \times 32 \times 48^3 \approx 1.39 \times 10^9$. If we evolve the system for 1000 steps, the total computation is on order of 1.39 trillion. Note that this kind of algorithmic intensity is similar to what one would generally expect from simulation of flow over a full-scale Boeing aircraft using compressible Euler or Navier-Stokes.

5.1. **1-D Spatial accuracy: Mixing regime.** We begin with the spatial accuracy test in mixing regime i.e., Knudsen number is a function of space variable. This results in a wide range of mixing scales which is typical of practical aerodynamic problems. Tests akin to these have been widely used in kinetic theory [22] since analytical solutions for Boltzmann equation do not exist. The purpose of this case is to confirm high order spatial accuracy for variable Kn. Consider equation (1) with initial condition as

$$f(0, x, \boldsymbol{v}) = \frac{1}{2}\Big(\mathcal{M}(n, u, T) + \mathcal{M}(n, -u, T)\Big), \tag{65}$$

with

$$n = 1 + 0.2\sin(\pi x), \quad u = \Big(\frac{1}{2}, 0, 0\Big), \quad T = \frac{1}{1 + 0.2\sin(\pi x)},$$

$$\mathrm{Kn} = \frac{1}{2}\Big(\tanh(6 - 5\,x) + \tanh(-4 + 5\,x)\Big) + \varepsilon_0, \tag{66}$$

where ε_0 is some fixed positive constant. The spatial domain is taken as $x \in [0, 2]$ with periodic boundary condition. The velocity domain is truncated into $[-|v|_{\max}, |v|_{\max}]^3$ with $|v|_{\max} = 5$ and discretized by a finite difference scheme using $N_v = 24^3$ grid points. We choose $h = \Delta x = 2/N_e$ and set $\Delta t \ll c\ \Delta x/|v|_{\max}$, where c is CFL number of the p-order DG scheme. Since the exact solution is not available, the numerical solution on a finer mesh $\Delta x/2$ is used as a reference solution to compute the error for the solution on the mesh size Δx i.e.,

$$E_{\Delta t,\,\Delta x} := \|f_{\Delta t,\,\Delta x} - f_{\Delta t/2,\,\Delta x/2}\|_{L^2_{x,\boldsymbol{v}}} = \frac{1}{(2|v|_{\max})^3}\,\frac{1}{2}\left(\int_{\Omega_{\boldsymbol{v}}}\int_{\Omega_{\boldsymbol{x}}} |f_{\Delta t,\,\Delta x} - f_{\Delta t/2,\,\Delta x/2}|^2\,\mathrm{d}x\,\mathrm{d}\boldsymbol{v}\right)^{1/2}. \tag{67}$$

The solution is computed for $\varepsilon_0 = \{10^{-4}\}$ with $N_e = \{2, 4, 8, 16, 32, 64\}$. In Tab. (1), we present the error, equation (67), at $t = 0.1$ using the Boltzmann equation with entropy stable flux. The scheme recovers the standard convergence rate of nodal schemes constructed on Gauss-Lobatto quadrature for non-linear problems [34, 33]. Next, to verify that the limiters are consistent (produce correct reconstruction with entropy-stable flux), we apply limiter on all elements. In Tab. (1), we show the corresponding numerical evidence.

5.2. **1-D conservation test: shock-front propagation.** Next, we consider two initial-value Riemann problems involving shock-front propagation. A normal shock is placed at $x = 0.5$ in a unit domain, $x \in [0, 1]$; and then marched in time. The full set of simulation parameters have been reiterated in Tab. (2).

In the first test case, Case SF-01, the shock is kept stationary. We run the test for 20,000 steps. The results have been shown in Fig. (1). The reference results are obtained using OpenFOAM [28] (with adiabatic index, $\gamma = 5/3$) which implements the Kurganov-Tadmor [47] scheme. The results show that error in mean velocity is $0.2/u_0 \times 100 \approx 0.05\%$. This asserts that the scheme is conservative. The error can be decreased further by refining the velocity grid.

$\downarrow p$	h	$h/2$	$h/4$	$h/8$	$h/16$	$\approx$ Rate
$\varepsilon_0 = 10^{-4}$						
	2	4	8	16	32	
1	1.95×10^{-5}	4.10×10^{-6}	1.15×10^{-6}	3.01×10^{-7}	7.77×10^{-8}	1.96
2	1.51×10^{-6}	6.39×10^{-7}	7.38×10^{-8}	9.39×10^{-9}	1.17×10^{-9}	3.01
$\varepsilon_0 = 10^{-4}$ (with high order limiter (49) on all elements)						
	2	4	8	16	32	
1	1.72×10^{-5}	7.50×10^{-6}	2.09×10^{-6}	4.70×10^{-7}	1.07×10^{-7}	2.14
2	6.09×10^{-6}	5.60×10^{-6}	8.72×10^{-7}	5.06×10^{-8}	5.97×10^{-9}	3.08
$\varepsilon_0 = 10^{-4}$ (with TVD limiter (48) on all elements)						
	2	4	8	16	32	
1	1.72×10^{-5}	7.48×10^{-6}	2.10×10^{-6}	4.52×10^{-7}	1.01×10^{-7}	2.16
2	5.88×10^{-6}	5.48×10^{-6}	1.80×10^{-6}	4.36×10^{-7}	1.16×10^{-7}	1.91

TABLE 1. Accuracy test for 1D-3V Boltzmann equation with entropy stable flux.

In the second test case, Case SF-02, the shock is assigned an initial velocity along x-axis. We again run the test for 20,000 steps. The results have been shown in Fig. (1). First, because our DG approximation endows every element with a solution expressed as a polynomial, we expect to see the classical Gibb's phenomenon caused by approximating discontinuous functions using a (piecewise) continuous polynomial. In classical DG, one uses a slope limiter which smooths out these oscillations. In continuous Galerkin, one adds shock-capturing terms [6] to the equation system to achieve the same. Both the ideas are related [12]. In entropy stable DG, we do not add such stabilization. Second, although the oscillations are present, we observe that the shock-front is propagated at a correct speed, and the scheme is stable. A crucial point to note here is: We use a first order time-marching scheme, and we still recover a sharp shock front. As a point of reference, the streamlined methods [6] add remarkably high amount of diffusion for these kinds of problems (see [29]).

5.3. **1-D initial-boundary value problem: Couette flow.** In the present test case, we test the entropy fluxes for initial boundary value problem with fully diffuse wall boundary condition. Consider a standard 1-D Couette flow setup: A rarefied Argon gas is trapped between two walls in a millimeter domain. The left wall moves with velocity $(0, -250, 0)$ m/s and the right wall with $(0, 250, 0)$ m/s. Both the walls are at temperature of 273 K. We use the standard VHS model properties [3] for modelling Argon. We consider three (number densities, Kn): a) $(3.537 \times 10^{20}, 3.7)$; b) $(3.537 \times 10^{21}, 3.7 \times 10^{-1})$; c) $(3.537 \times 10^{22}, 3.7 \times 10^{-2})$. All cases use a fixed time step of 1.5×10^{-8} s for a 4 element 3rd order DG scheme. The mesh is *not* resolved to O(Kn) [22]. The full set of simulation parameters have been reiterated in Tab. (3). The results have been presented in Fig. (2). Our scheme recovers the correct temperature and velocity which matches DSMC for a range of Knudsen numbers.

Note that the velocity recovered for this case using BGK model is generally consistent with the DSMC results. But the temperature predicted by BGK is inconsistent: it is remarkably high due to its Prandtl number defect. Our scheme recovers the correct temperature which matches DSMC thereby asserting that the scheme solves the *actual* Boltzmann equation asymptotically. BGK is merely used as a penalty (observe that BGK operator has been added and subtracted in equation (51)).

5.4. **1-D Sod-shock tube problem.** In the current test case, we consider the classical Sod shock tube problem. The full set of simulation parameters have been reiterated in Tab. (4). In summary, we utilize four different (upstream, downstream) densities [kg/m^3]: a) $(1 \times 10^{-6}, 0.125 \times 10^{-6})$; b) $(1 \times 10^{-5}, 0.125 \times 10^{-5})$; c) $(1 \times 10^{-4}, 0.125 \times 10^{-4})$; d) $(1 \times 10^{-3}, 0.125 \times 10^{-3})$. These upstream densities correspond to Knudsen number [40] of a) 4.96×10^{-1}; b) 4.96×10^{-2}; c) 4.96×10^{-3}; d) 4.96×10^{-4}. We use a 256 element 3rd order DG scheme with a time step of $\approx 4 \times 10^{-6}\, s$ for all cases. The velocity domain $[-6.14, 6.14]^3$ is discretized using 32^3 points.

The results have been shown in Fig. (3): the symbols denote the results using entropy stable flux *without* limiter; the lines denote the results using traditional fluxes with TVD limiter given by equation (48). We emphasize that one should expect overshoots near contact discontinuities as seen in last row of Fig. (3). This is a known fact, for example, in one of the earliest papers on entropy-stable schemes (Fig. 4, [63]), author noted that such results demonstrate the purely dispersive character of the entropy stable schemes. Dispersive oscillations on

Parameter	SF-01	SF-02
Working Gas	Air	
Physical space (m)	[0, 1]	
Characteristic length: H_0 (m)	1	
Char. velocity: u_0 (m/s)	400.05	
Char. time: t_0 (s)	0.0025	
Char. temperature: T_0 (K)	278.746	
Char. no. density: n_0 (m^{-3})	2.599×10^{25}	
Velocity space	$[-5.44u_0, 5.44u_0]^3$	
Points in velocity mesh: N_v	32^3	
Number of elements: N_e	128	
Polynomial degree: p	2	
Time step: $\Delta t/t_0$:	1×10^{-4}	
Final time: T/t_0:	2	
Upstream region: [0, 0.5]:		
Density: n/n_0	1.4	
Velocity: u/u_0	0	0.1
Temperature: T/T_0	0.714	
Downstream region: (0.5, 1]:		
Density: n/n_0	1	
Velocity: u/u_0	0	0.1
Temperature: T/T_0	1	

TABLE 2. Numerical parameters for shock-front propagation test cases. The molecular parameters for "Air" are as indicated in Appendix-A of [3]. We solve the Boltzmann equation in non-dimensional form (c.f. [40] for non-dimensionalization convention). The characteristic variables indicated above are needed there. The variables that remain unchanged, are not repeated.

the *mesh scale* are observed due to the absence of any dissipation mechanism [50, 52] (because entropy vanishes at contact discontinuities, whereas in slope-limiters, there is dissipation applied at contact discontinuities. Note that DG schemes have remarkable dissipation and dispersion properties [1] and therefore they have very low numerical viscosity by virtue). The key point is: *numerical solutions do not blow up.* Since oscillations are on order of mesh scale, observe that in Fig. 4 of [63], they die down upon refining the mesh. In fact, in a related paper on entropy-stable schemes [37], authors constructed a flux, equation (43), that adds necessary dissipation at shocks. The rationale is that for many calculations, the shock structure is of no interest and a far coarser resolution is adequate. Nevertheless, if one wants a visually smooth solution, a strategy is to introduce a TVD limiter at the *end of the post-processing* step.

In Fig. (4) we present a comparison of DSMC and the proposed method for the sod-shock tube problem at low Knudsen number. We observe an excellent agreement between the DSMC and the proposed scheme. Note that Sod-shock tube is a transient problem. In general, it's difficult to obtain time-accurate results using DSMC. In our simulations, we have used 1000 cells, 4000 particles per cell, and a time-step of 10^{-8} s (since the final time is 7×10^{-3} s, this results in 700,000 time-steps). On an 8-core machine, these simulations took 153.39 hours each. There is still some statistical noise in the DSMC results which can be certainly reduced further by increasing the number of particles per cell, however, the same remains elusive from a computational viewpoint.

5.5. **1-D Lax-shock tube problem.** Next, we consider the Lax shock tube problem. The full set of simulation parameters have been reiterated in Tab. (5). The results have been shown in Fig. (5): the symbols denote the results using entropy stable fluxes with TVD limiter; the lines denote the results using classical Boltzmann flux with TVD limiter (48). We observe a remarkably close agreement, both in near continuum regime (Euler limit), as well as the usual rarefied regime.

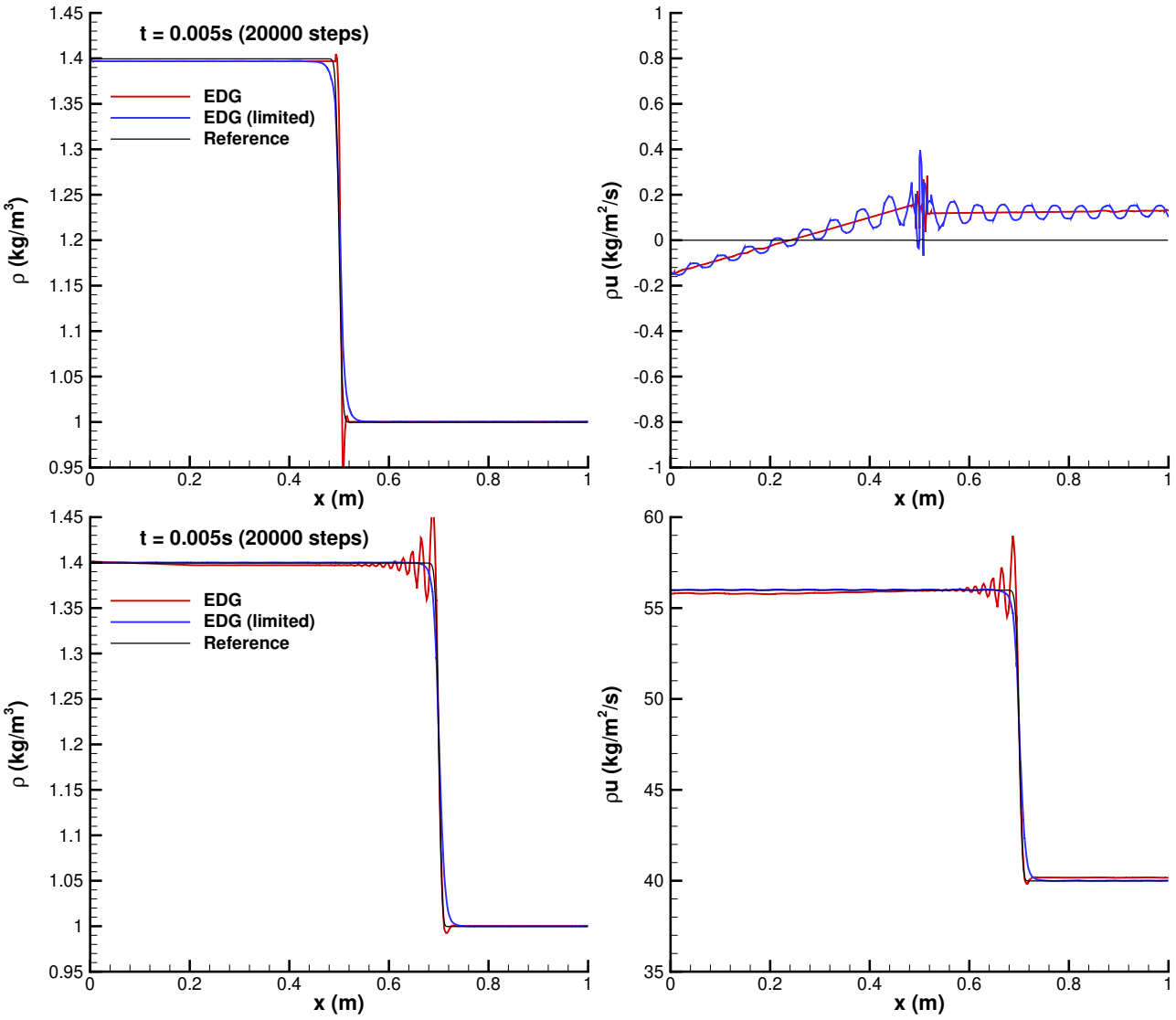


FIGURE 1. Shock front propagation: Variation of mass-density and mass-flux. Top row: Case SF-01, Bottom row: Case SF-02. Results after 20,000 time-steps. EDG refers to Boltzmann equation with entropy stable flux without limiter; EDG (limited) refers to entropy stable flux with TVD limiter. Reference refers to results using OpenFOAM [28] Euler solver which implements the Kurganov-Tadmor [47] scheme. Observe that the error in mean velocity and momentum is $0.2/u_0 \times 100 \approx 0.05\%$, which asserts that the scheme is conservative.

6. Numerical tests in two-dimensional physical space

6.1. **2-D heat-conduction in a square cavity.** We consider heat-conduction in a unit milli-meter square cavity $[0, 1\,\text{mm}]^2$. The top wall is heated to 373 K, whereas other walls are kept at a constant temperature of 273 K. All the walls are at rest. At $t > 0$, this heat propagates through the domain. The exact simulation parameters have been provided in Tab. (6).

In Fig. (6), we illustrate the residual, $\|f^{n+1} - f^n\|$, history as a function of non-dimensional time. Here f^n is the distribution function at n^{th} step. We observe that both the schemes produce qualitatively same behavior for "smooth" problems with temperature gradients.

In Fig. (7), we illustrate results from DG and EDG schemes at steady state. Ignoring statistical noise, the results from both DG and EDG schemes agree well with the DSMC simulations. To understand the differences between the DG and EDG schemes, in Fig. (8) we show the solution at different time instants. The results from EDG schemes agrees very well with conventional DG scheme for Boltzmann equation.

6.2. **2-D lid-driven cavity flow.** We consider lid-driven cavity flow problem. The flow is again setup in a unit milli-meter square cavity $[0, 1\,\text{mm}]^2$. The top wall is moving with a velocity of $(0, 50, 0)$ m/s, whereas other walls

Parameter	C-01	C-02	C-03
Working Gas	Ar		
Physical space (mm)	$[0, 1]$		
Characteristic length: H_0 (mm)	1		
Char. velocity: u_0 (m/s)	337.2		
Char. temperature: T_0 (K)	273		
Char. no. density: n_0 (m^{-3})	3.537×10^{20}	3.537×10^{21}	3.537×10^{22}
Velocity space	$[-5u_0, 5u_0]^3$		
Points in velocity mesh: N_v	24^3		
Number of elements: N_e	4		
Polynomial degree: p	2		
Time step: Δt (s):	1.5×10^{-8}		
Initial conditions			
Number density: n (m^{-3})	3.537×10^{20}	3.537×10^{21}	3.537×10^{22}
Knudsen number: Kn	3.7	3.7×10^{-1}	3.7×10^{-2}
Velocity: u (m/s)	$(0, 0, 0)$		
Temperature: T (K)	273		
Left wall conditions			
Velocity: $\boldsymbol{u}$ (m/s)	$(0, -250, 0)$		
Temperature: T (K)	273		
Right wall conditions			
Velocity: $\boldsymbol{u}$ (m/s)	$(0, 250, 0)$		
Temperature: T (K)	273		

TABLE 3. Numerical parameters for Couette flow test case. The molecular parameters for "Ar" are as indicated in Appendix-A of [3].

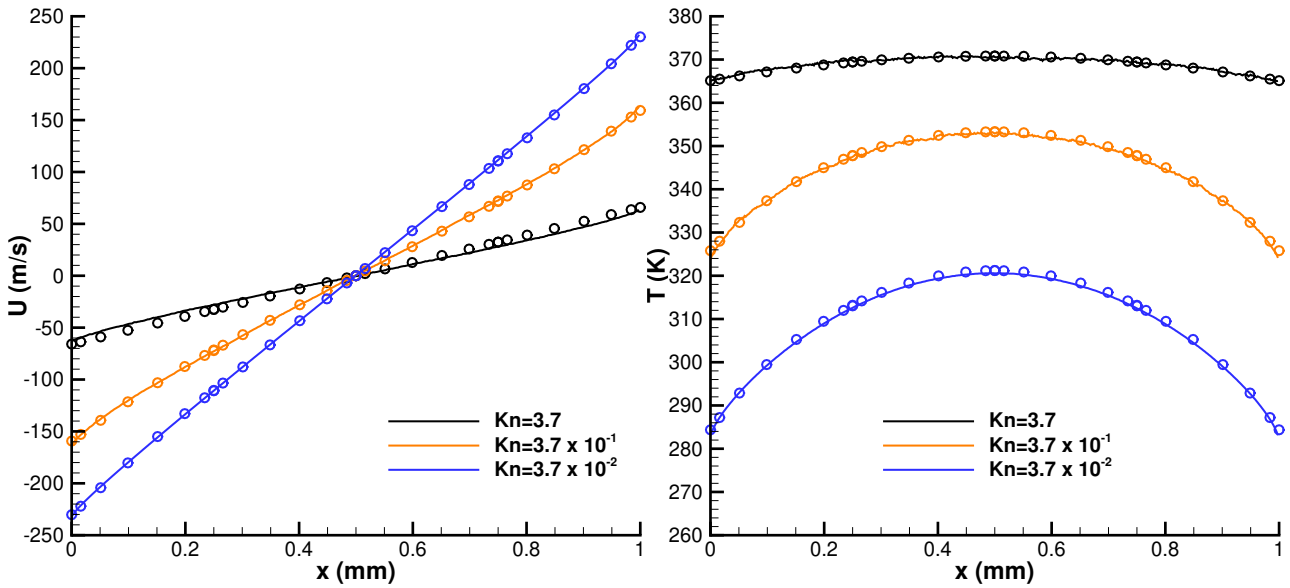


FIGURE 2. 1-D Couette flow: Velocity and temperature using Boltzmann equation with entropy-stable flux without any limiter. The reference DSMC (lines) results were obtained using SPARTA [60].

Parameter	SOD-01	SOD-02	SOD-03	SOD-04
Working Gas	Air			
Physical space (m)	$[-5, 5]$			
Characteristic length: H_0 (m)	1			
Char. velocity: u_0 (m/s)	400.05			
Char. temperature: T_0 (K)	278.746			
Char. no. density: n_0 (m^{-3})	2.599×10^{18}	2.599×10^{19}	2.599×10^{20}	2.599×10^{21}
Velocity space	$[-6.14u_0, 6.14u_0]^3$			
Points in velocity mesh: N_v	32^3			
Number of elements: N_e	256			
Polynomial degree: p	2			
Time step: Δt (s):	4×10^{-6}			
Upstream region: $[-5, 0]$:				
Kn:	4.96×10^{-1}	4.96×10^{-2}	4.96×10^{-3}	4.96×10^{-4}
Density: n/n_0	8			
Velocity: u/u_0	0			
Temperature: T/T_0	1.25			
Downstream region: $(0, 5]$:				
Density: n/n_0	1			
Velocity: u/u_0	0			
Temperature: T/T_0	1			

TABLE 4. Numerical parameters for the Sod shock tube problem. The molecular parameters for "Air" are as indicated in Appendix-A of [3].

Parameter	LAX-01	LAX-02	LAX-03	LAX-04
Working Gas	Air			
Physical space (m)	$[-5, 5]$			
Characteristic length: H_0 (m)	1			
Char. velocity: u_0 (m/s)	400.05			
Char. temperature: T_0 (K)	278.746			
Char. no. density: n_0 (m^{-3})	2.08×10^{19}	2.08×10^{20}	2.08×10^{22}	2.08×10^{25}
Velocity space	$[-9.89u_0, 9.89u_0]^3$			
Points in velocity mesh: N_v	48^3			
Number of elements: N_e	256			
Polynomial degree: p	2			
Time step: Δt (s):	2.5×10^{-6}			
Upstream region: $[-5, 0]$:				
Kn:	6.2×10^{-2}	6.2×10^{-3}	6.2×10^{-5}	6.2×10^{-8}
Density: n/n_0	0.445			
Velocity: u/u_0	0.6989			
Temperature: T/T_0	5.68			
Downstream region: $(0, 5]$:				
Density: n/n_0	0.5			
Velocity: u/u_0	0			
Temperature: T/T_0	1.142			

TABLE 5. Numerical parameters for the Lax shock tube problem.

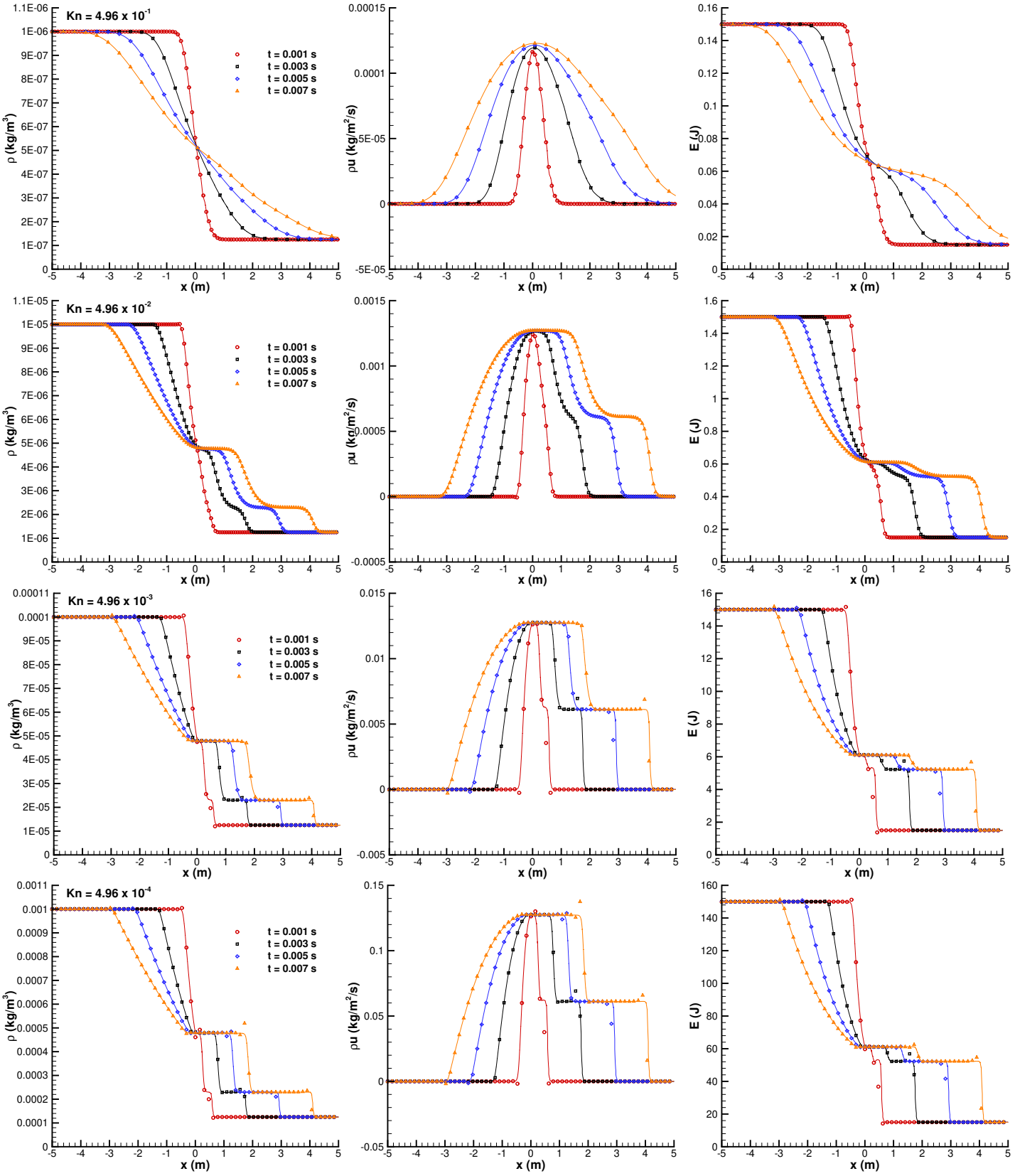


FIGURE 3. Sod shock tube problem: Variation of conserved flow properties, viz. mass-density, mass-flux, and energy. Symbols denote the results from Boltzmann equation with entropy stable flux without any limiter; and lines denote the results from Boltzmann equation with classical flux and TVD limiter.

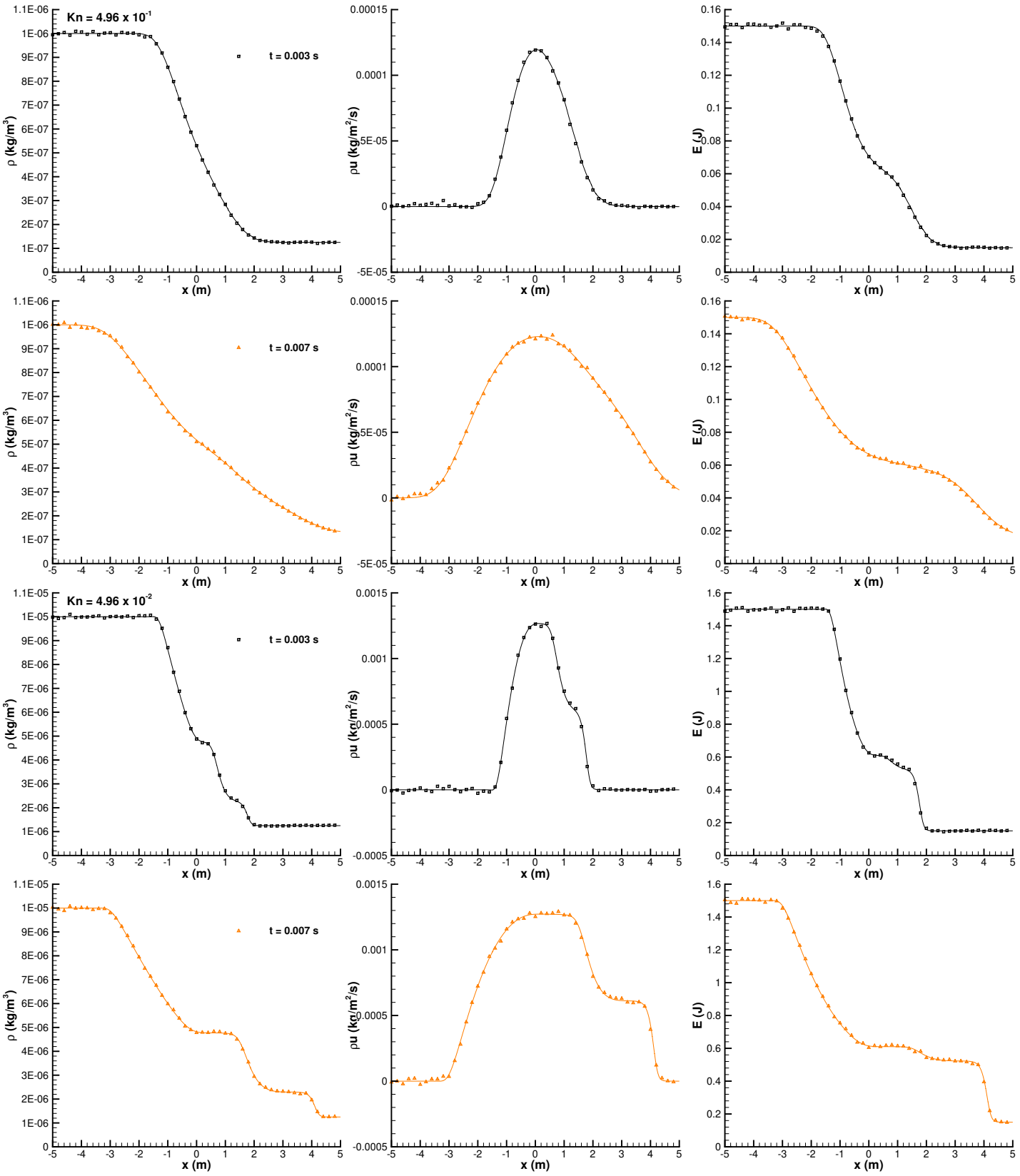


FIGURE 4. Sod shock tube problem: Comparison of conserved flow properties from proposed method and direct simulation Monte Carlo (DSMC) method. Symbols denote the results from DSMC; and lines denote the results from Boltzmann equation with entropy stable flux without any limiter. Row 1: Kn=4.96×10^{-1}, $t = 3 \times 10^{-3}$ s; Row 2: Kn=4.96×10^{-1}, $t = 7 \times 10^{-3}$ s; Row 3: Kn=4.96×10^{-2}, $t = 3 \times 10^{-3}$ s; Row 4: Kn=4.96×10^{-2}, $t = 7 \times 10^{-3}$ s.

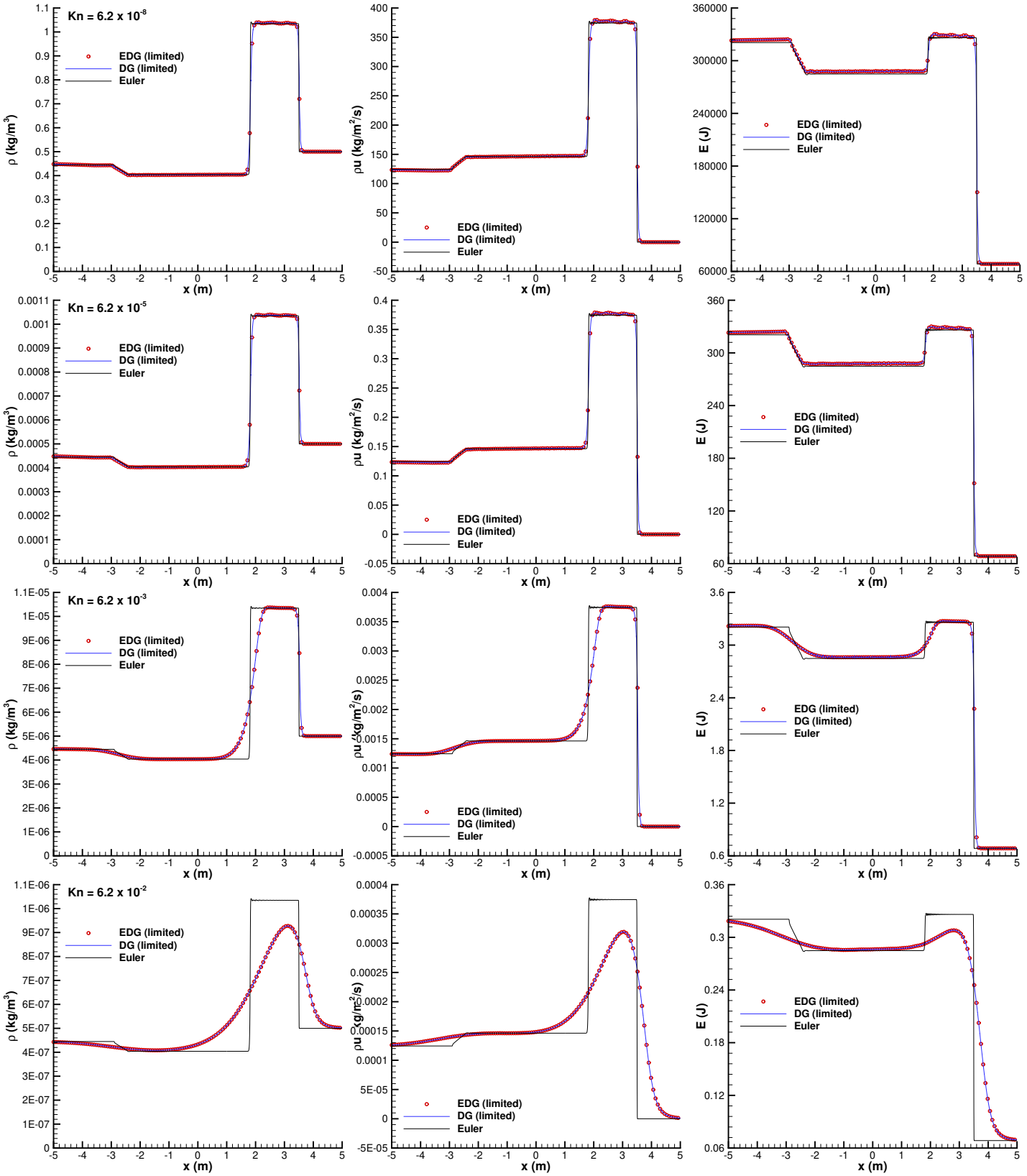


FIGURE 5. Lax shock tube problem: Variation of conserved flow properties, viz. mass-density, mass-flux, and energy. Symbols denote the results from Boltzmann equation with entropy stable flux with TVD limiter; blue lines denote the results from Boltzmann equation with classical flux and TVD limiter; and black line refers to results using OpenFOAM [28] Euler solver which implements the Kurganov-Tadmor [47] scheme.

Parameter	Case HC-01
Working gas	Ar
Physical space (mm)	$[0, 1]^2$
Characteristic length: H_0 (mm)	1
Characteristic velocity: u_0 (m/s)	337.2
Characteristic temperature: T_0 (K)	273
Characteristic no. density: n_0 (m^{-3})	3.469×10^{22}
Characteristic time: t_0 (s)	2.96×10^{-6}
Velocity space	$[-5u_0, 5u_0]^3$
No. of points in velocity mesh: N_v^3	24^3
Initialization conditions	
Number density: n (m^{-3})	3.469×10^{22}
Velocity: $\boldsymbol{u}$ (m/s)	$(0, 0, 0)$
Temperature: T (K)	273
Top wall conditions	
Velocity: $\boldsymbol{u}$ (m/s)	$(0, 0, 0)$
Temperature: T (K)	373
Other wall conditions	
Velocity: $\boldsymbol{u}$ (m/s)	$(0, 0, 0)$
Temperature: T (K)	273

TABLE 6. Test case from section (6.1): Numerical parameters for 2-D heat conduction case. The molecular parameters for "Ar" are as indicated in Appendix-A of [3].

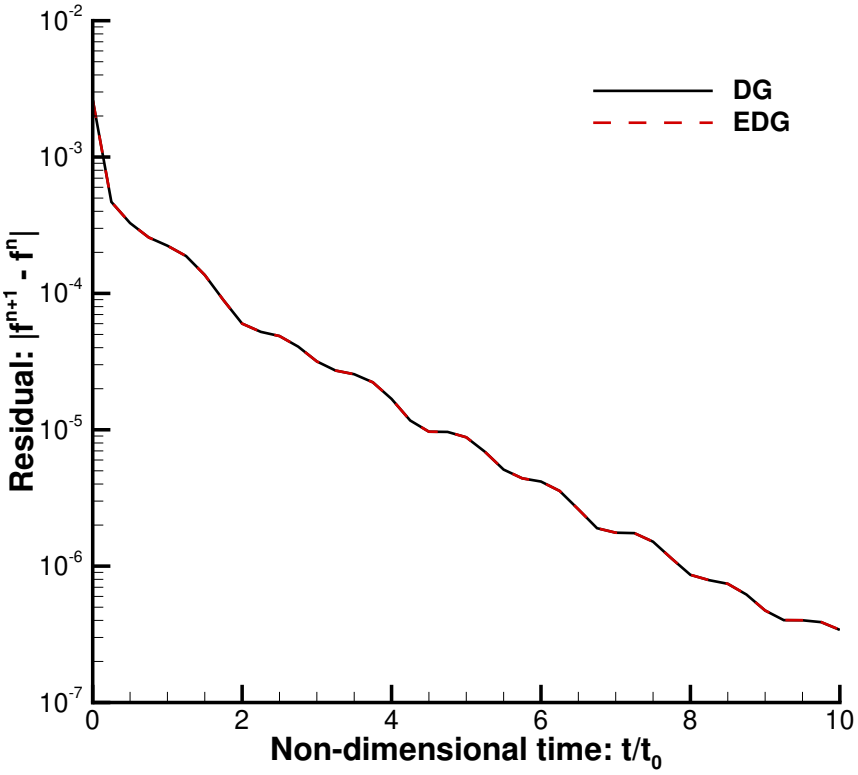
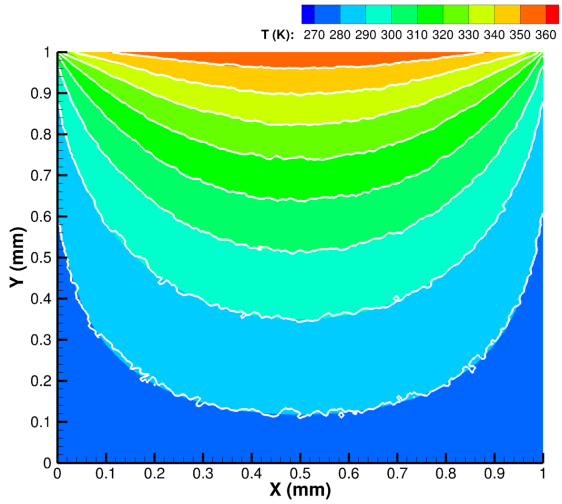
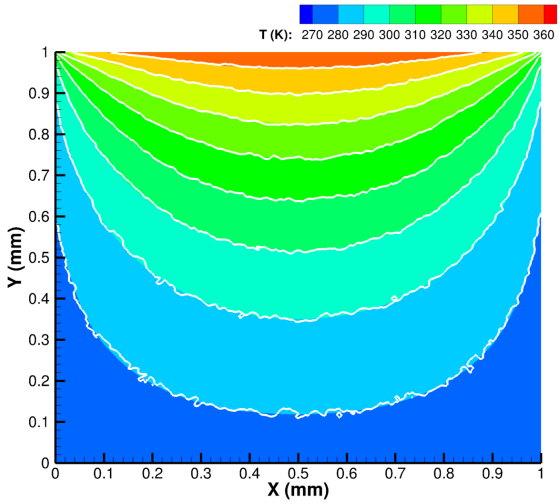


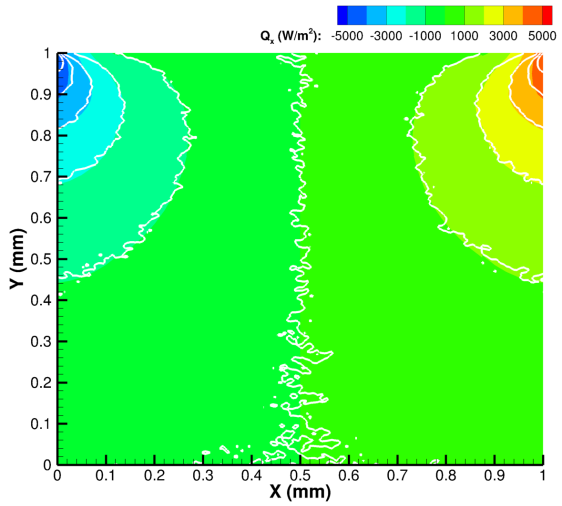
FIGURE 6. Test case from section (6.1): Residual history of the convention DG scheme, and entropy-stable DG scheme for Boltzmann equation.



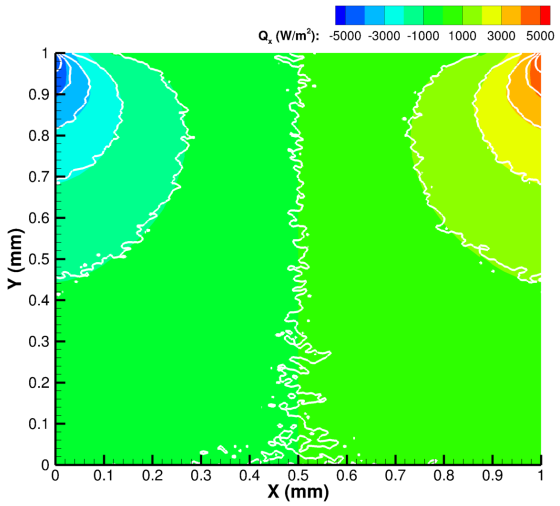
(A) Temperature: T (K)



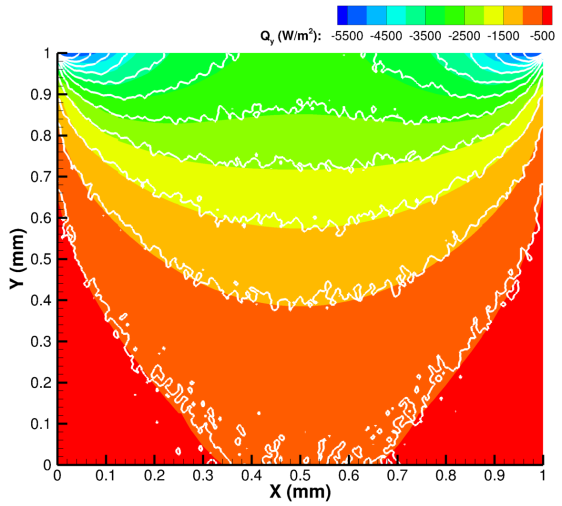
(B) Temperature: T (K)



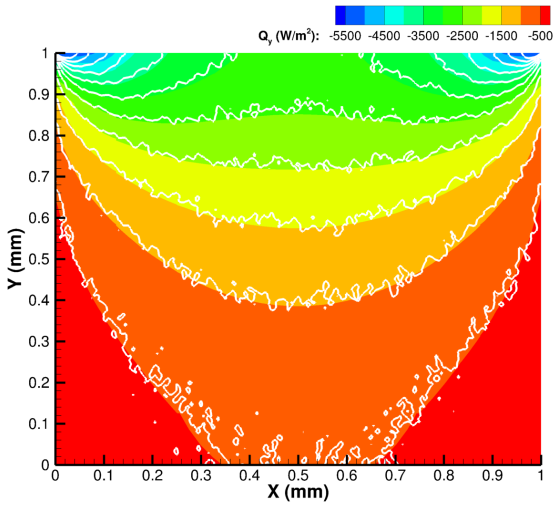
(C) x-component of heat-flux: q_x (W/m^2)



(D) x-component of heat-flux: q_x (W/m^2)

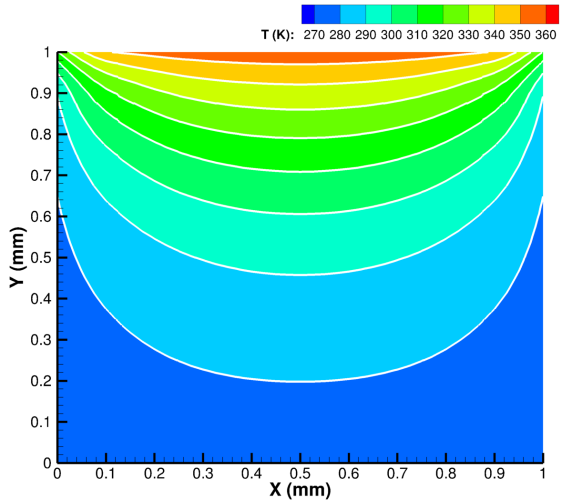


(E) y-component of heat-flux: q_y (W/m^2)

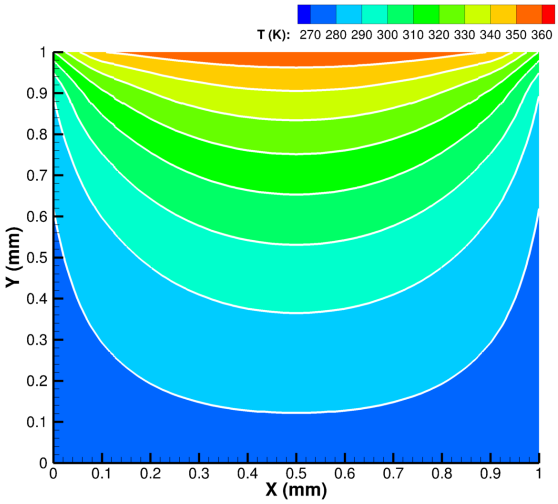


(F) y-component of heat-flux: q_y (W/m^2)

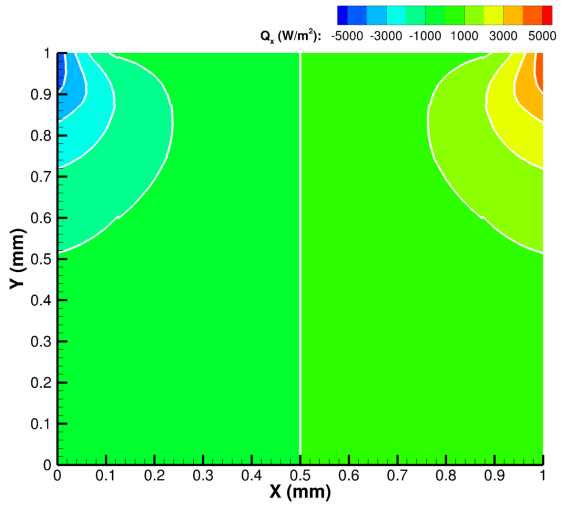
FIGURE 7. Test case from section (6.1): Results for 2-D heat conduction case at steady-state. In each of the plot, white lines indicate the results from DSMC simulations; whereas the background contour is from DG methods. The left column illustrates the results from conventional DG formulation for Boltzmann equation, whereas the right column shows the results from entropy-stable DG formulation for Boltzmann equation.



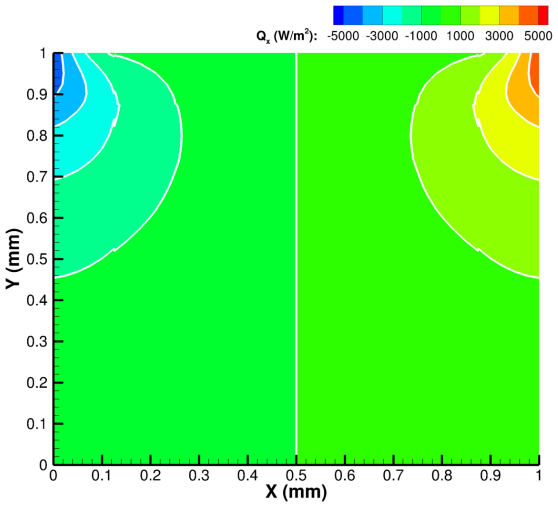
(A) Temperature: T (K)



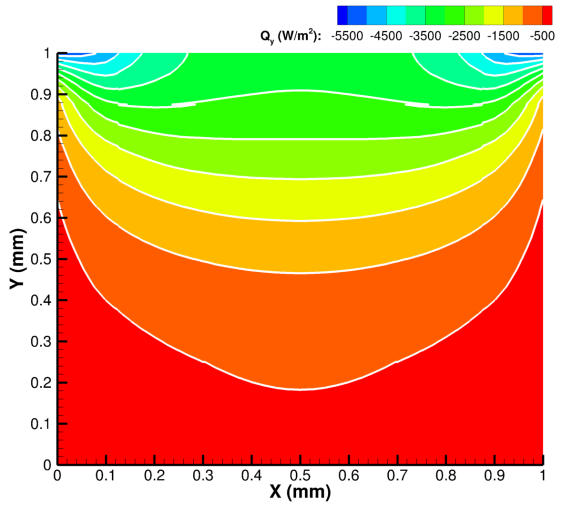
(B) Temperature: T (K)



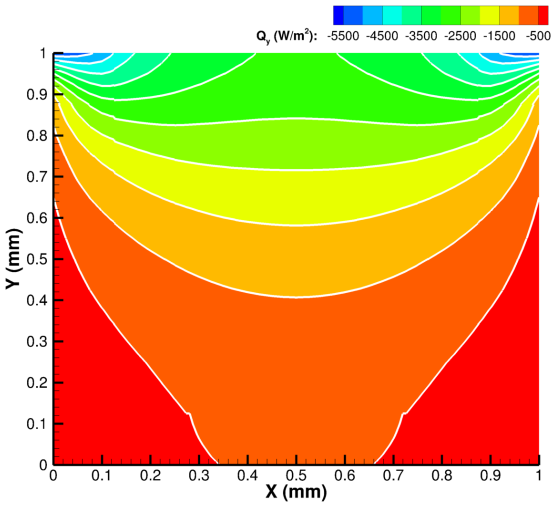
(C) x-component of heat-flux: q_x (W/m^2)



(D) x-component of heat-flux: q_x (W/m^2)



(E) y-component of heat-flux: q_y (W/m^2)



(F) y-component of heat-flux: q_y (W/m^2)

FIGURE 8. Test case from section (6.1): Results for 2-D heat conduction case at different times. In each of the plot, white lines indicate the results from conventional DG formulation for Boltzmann equation; whereas the background contour is from entropy-stable DG formulation for Boltzmann equation. The left column shows results at $t = 2$, whereas the right column shows results at $t = 4$.

are kept at rest. All the walls are assigned a temperature of 273 K. At $t > 0$, the flow develops due to shearing of fluid layers. The exact simulation parameters have been provided in Tab. (7).

In Fig. (9), we illustrate the residual, $\|f^{n+1} - f^n\|$, history as a function of non-dimensional time. Here f^n is the distribution function at n^{th} step. We observe that the residual from the entropy-stable DG scheme for the Boltzmann equation is lower for "smooth" problems with velocity-gradient.

In Fig. (9), we observe that the error-norm is on order of 10^{-7}, and it doesn't drop further. To appreciate why this is the case, first recall that we use a Fourier-spectral method [58, 40] for evaluating the collision operator. A natural question is: How do these methods perform? In [58, Figs. (7, 8)], [55, Tab. (1)], [40, Fig. (1)] authors have solved the spatially-homogeneous Boltzmann equation and have presented the error norms. In all these works, we see that the error-norm is on order of 10^{-7}. As these methods are spectrally accurate, it is possible to reduce the error-norm further (c.f. [55, Tab. (1)]). However, doing so incurs a tremendous amount of computational cost. Recall that we are solving a 6-dimensional problem (2-space, 3-velocity, 1 time).

In Fig. (10), we illustrate results from DG and EDG schemes at steady state. Ignoring statistical noise, the results from both DG and EDG schemes agree well with the DSMC simulations. To understand the differences between the DG and EDG schemes, in Fig. (11) we show the solution at different time instants. The results from EDG schemes agrees very well with conventional DG scheme for Boltzmann equation.

Parameter	Case LD-01
Working gas	Ar
Physical space (mm)	$[0,\ 1]^2$
Characteristic length: H_0 (mm)	1
Characteristic velocity: u_0 (m/s)	337.2
Characteristic temperature: T_0 (K)	273
Characteristic no. density: n_0 (m^{-3})	1.678×10^{21}
Characteristic time: t_0 (s)	2.96×10^{-6}
Velocity space	$[-6u_0,\ 6u_0]^3$
No. of points in velocity mesh: N_v^3	48^3
Initialization conditions	
Number density: n (m^{-3})	1.678×10^{21}
Velocity: $\boldsymbol{u}$ (m/s)	$(0, 0, 0)$
Temperature: T (K)	273
Top wall conditions	
Velocity: $\boldsymbol{u}$ (m/s)	$(0, 50, 0)$
Temperature: T (K)	273
Other wall conditions	
Velocity: $\boldsymbol{u}$ (m/s)	$(0, 0, 0)$
Temperature: T (K)	273

TABLE 7. Test case from section (6.2): Numerical parameters for 2-D lid-driven cavity case. The molecular parameters for "Ar" are as indicated in Appendix-A of [3].

7. Conclusions

We have presented a high-order entropy-stable deterministic numerical method for the solution of the full Boltzmann equation. The method combines a summation-by-parts collocated discontinuous Galerkin scheme in physical space together with the proposed entropy stable flux. The resulting semi-discrete scheme is shown to be locally and globally conservative, and it satisfies the celebrated Boltzmann's entropy inequality at a *discrete* level. This will ensure that the scheme converges to the unique entropy solution. A fully-discrete scheme is then constructed by utilizing a first order implicit-explicit discretization. The resulting schemes are iteration free, and suitable for solving problems in a wide range of Knudsen numbers, from continuum to free-molecular. The DG-type formulation employed in the present work has an element-local compact nature which enables effective parallelization on massively parallel architectures, contrary to its predecessors such as weighted essentially non-oscillatory schemes.

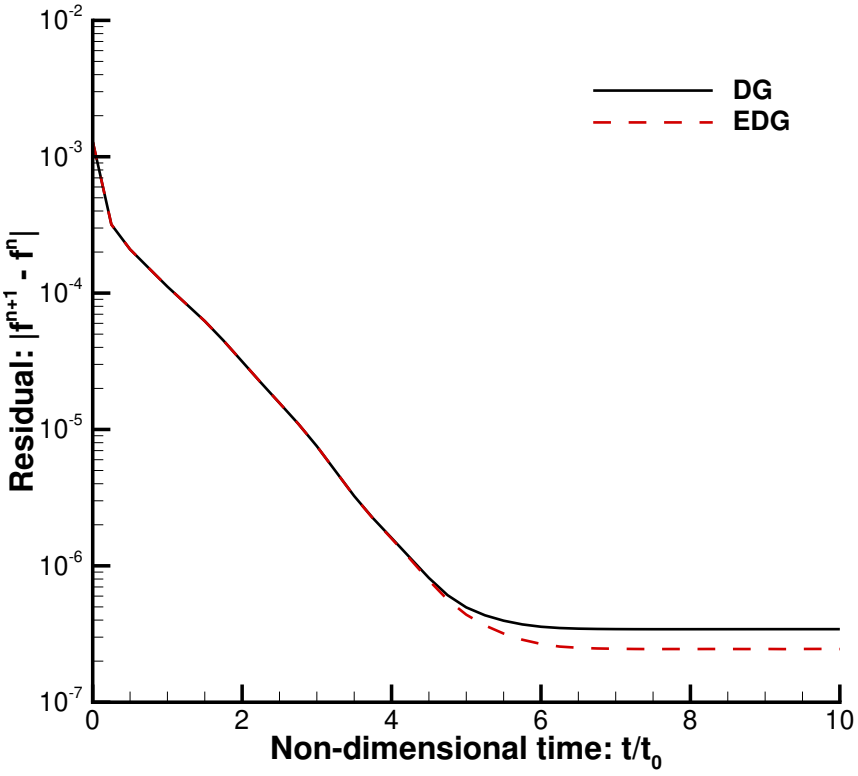


FIGURE 9. Test case from section (6.2): Residual history of the convention DG scheme, and entropy-stable DG scheme for Boltzmann equation.

To verify the scheme, we carry out simulations on well-known Riemann and mixing regime problems. Many of these cases have been run with different Knudsen numbers to highlight the general nature of the scheme. The scheme is high order accurate and stable in presence of shocks both in continuum and rarefied regime. The present findings may open up direction for provably non-linearly stable schemes for the entire family of Boltzmann equations, for instance phonon Boltzmann for phonon transport, Fokker-Planck for ionic transport in charged plasmas, quantum Boltzmann for transport processes in Fermi/Bose gases, generalized (active particles) Boltzmann for epidemics and virus mutations, etc.

REFERENCES

[1] AINSWORTH, M. Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods. *Journal of Computational Physics 198*, 1 (2004), 106–130.

[2] BHATNAGAR, P. L., GROSS, E. P., AND KROOK, M. A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems. *Physical review 94*, 3 (1954), 511.

[3] BIRD, G. A. *Molecular gas dynamics and the direct simulation of gas flows*. Clarendon, 1994.

[4] BIRDSALL, C. K., AND LANGDON, A. B. *Plasma physics via computer simulation*. CRC press, 2004.

[5] BISWAS, R., DEVINE, K. D., AND FLAHERTY, J. E. Parallel, adaptive finite element methods for conservation laws. *Applied Numerical Mathematics 14*, 1-3 (1994), 255–283.

[6] BROOKS, A. N., AND HUGHES, T. J. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Computer methods in applied mechanics and engineering 32*, 1-3 (1982), 199–259.

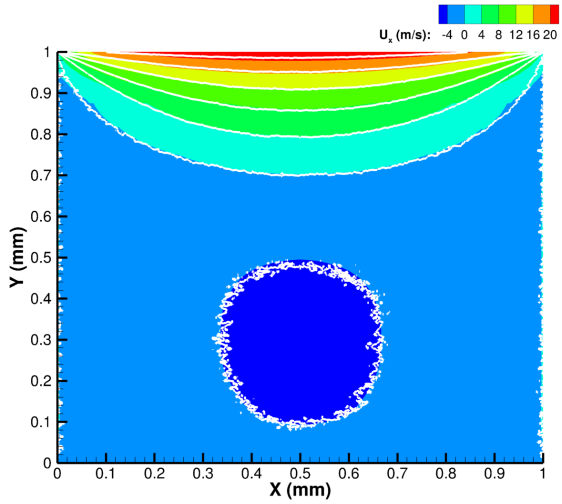
[7] CAI, Z., FAN, Y., AND YING, L. An entropic Fourier method for the Boltzmann equation. *SIAM Journal on Scientific Computing 40*, 5 (2018), A2858–A2882.

[8] CARPENTER, M. H., FISHER, T. C., NIELSEN, E. J., AND FRANKEL, S. H. Entropy stable spectral collocation schemes for the Navier–Stokes equations: Discontinuous interfaces. *SIAM Journal on Scientific Computing 36*, 5 (2014), B835–B867.

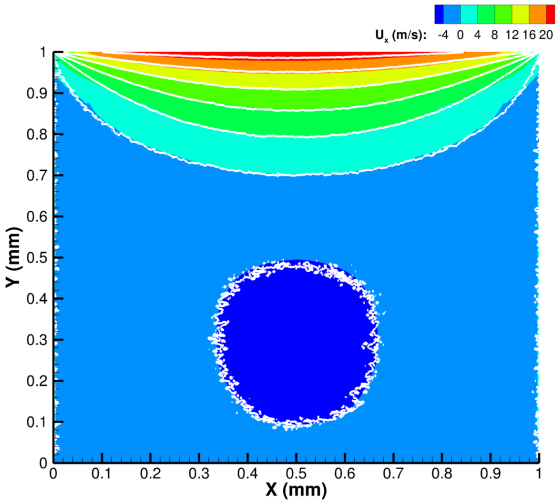
[9] CHEN, J. H., CHOUDHARY, A., DE SUPINSKI, B., DEVRIES, M., HAWKES, E. R., KLASKY, S., LIAO, W.-K., MA, K.-L., MELLOR-CRUMMEY, J., PODHORSZKI, N., ET AL. Terascale direct numerical simulations of turbulent combustion using S3D. *Computational Science & Discovery 2*, 1 (2009), 015001.

[10] CHEN, T., AND SHU, C.-W. Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws. *Journal of Computational Physics 345* (2017), 427–461.

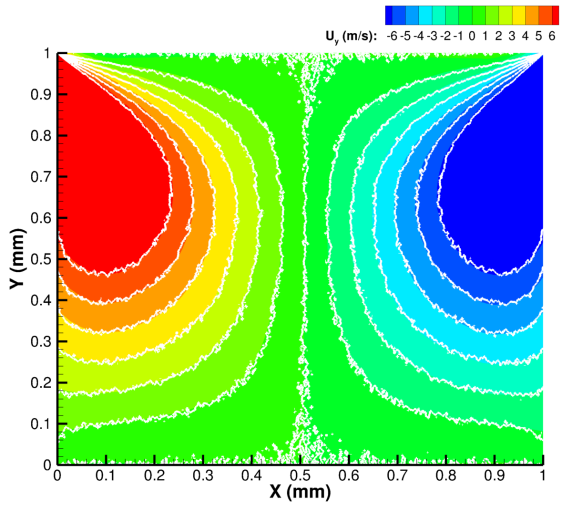
[11] COCKBURN, B. Discontinuous Galerkin methods for convection-dominated problems. In *High-order methods for computational physics*. Springer, 1999, pp. 69–224.



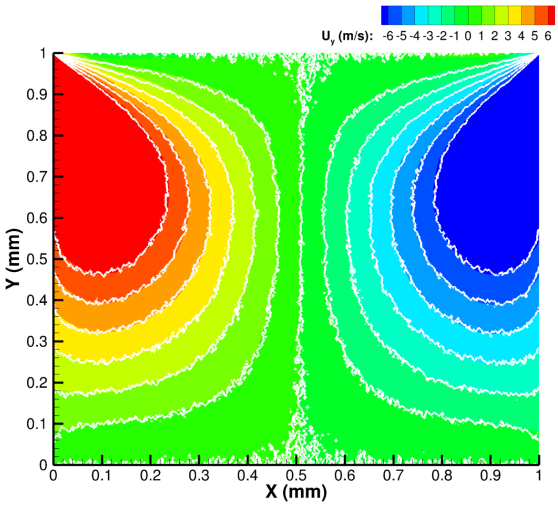
(A) x-component of velocity: U_x (m/s)



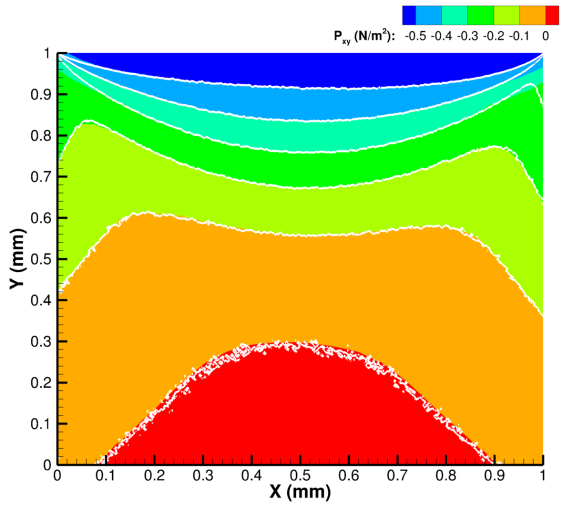
(B) x-component of velocity: U_x (m/s)



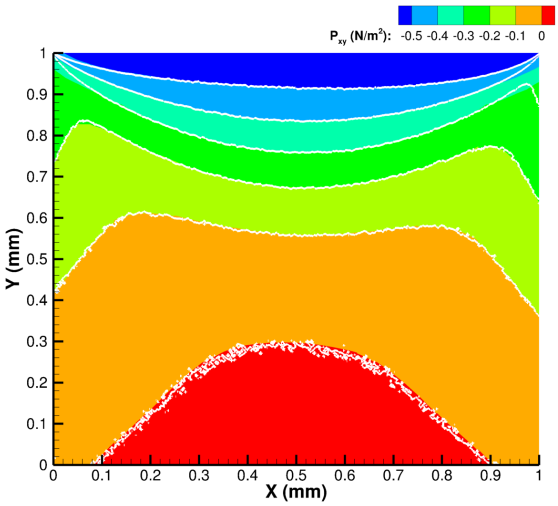
(C) y-component of velocity: U_y (m/s)



(D) y-component of velocity: U_y (m/s)

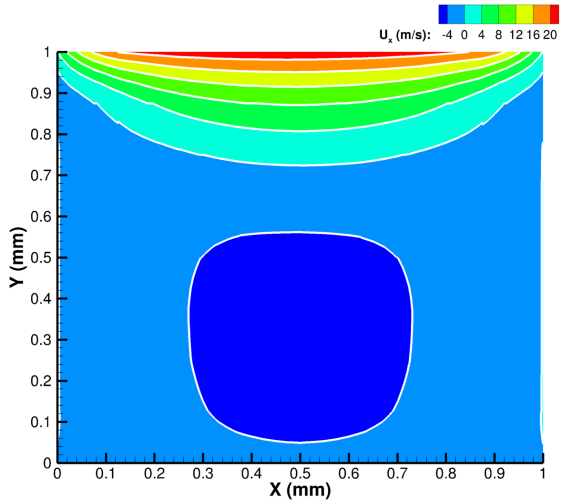


(E) xy-component of shear-stress: $\mathbb{P}_{xy}$ (m/s)

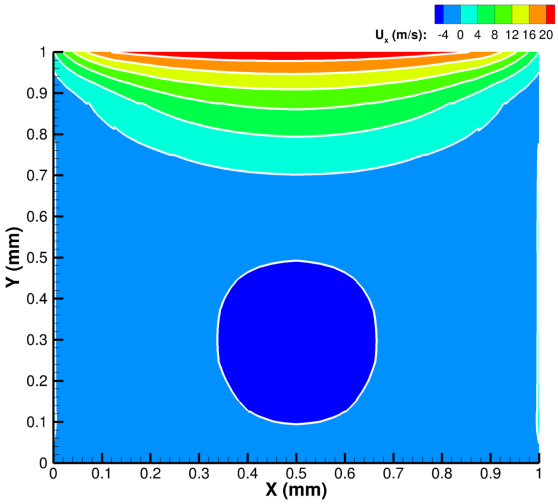


(F) xy-component of shear-stress: $\mathbb{P}_{xy}$ (m/s)

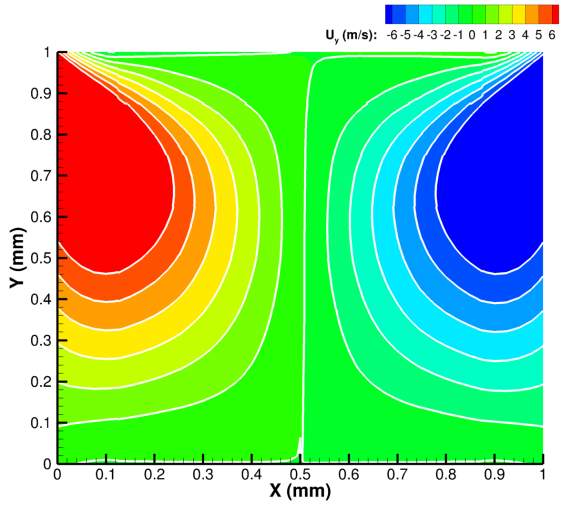
FIGURE 10. Test case from section (6.2): Results for 2-D lid-driven cavity flow at steady-state. In each of the plot, white lines indicate the results from DSMC simulations; whereas the background contour is from DG methods. The left column illustrates the results from conventional DG formulation for Boltzmann equation, whereas the right column shows the results from entropy-stable DG formulation for Boltzmann equation.



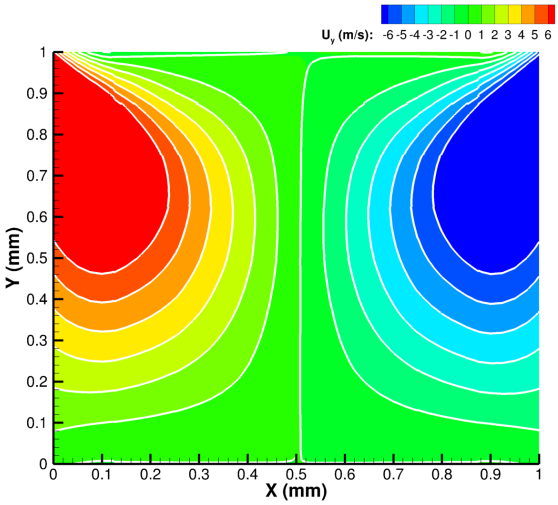
(A) x-component of velocity: U_x (m/s)



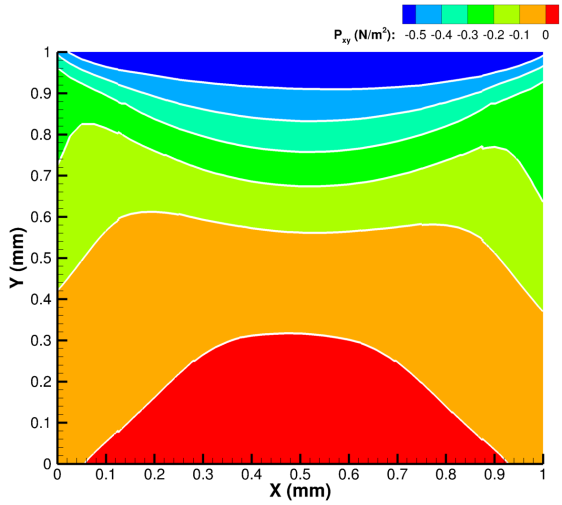
(B) x-component of velocity: U_x (m/s)



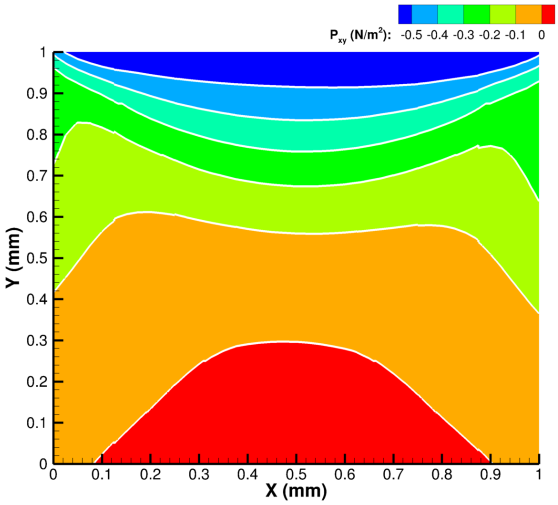
(C) y-component of velocity: U_y (m/s)



(D) y-component of velocity: U_y (m/s)



(E) xy-component of shear-stress: $\mathbb{P}_{xy}$ (m/s)



(F) xy-component of shear-stress: $\mathbb{P}_{xy}$ (m/s)

FIGURE 11. Test case from section (6.2): Results for 2-D lid-driven cavity flow at different time instants. In each of the plot, white lines indicate the results from conventional DG formulation for Boltzmann equation, whereas the background contour is from entropy-stable DG formulation for Boltzmann equation. The left column shows results at $t = 2$, whereas the right column shows results at $t = 4$.

[12] Cockburn, B. Devising discontinuous Galerkin methods for non-linear hyperbolic conservation laws. *Journal of Computational and Applied Mathematics 128*, 1-2 (2001), 187–204.

[13] de Masi, A., Esposito, R., and Lebowitz, J.-L. Incompressible Navier-Stokes and Euler limits of the Boltzmann equation. *Communications on pure and applied mathematics 42*, 8 (1989), 1189–1214.

[14] Desvillettes, L., and Villani, C. On the trend to global equilibrium for spatially inhomogeneous kinetic systems: the Boltzmann equation. *Inventiones mathematicae 159*, 2 (2005), 245–316.

[15] Dimarco, G., and Pareschi, L. Asymptotic preserving implicit-explicit Runge–Kutta methods for nonlinear kinetic equations. *SIAM Journal on Numerical Analysis 51*, 2 (2013), 1064–1087.

[16] Dimarco, G., and Pareschi, L. Numerical methods for kinetic equations. *Acta Numerica 23* (2014), 369–520.

[17] DiPerna, R. J., and Lions, P.-L. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Annals of Mathematics* (1989), 321–366.

[18] Du, Q., and Nicolaides, R. A. Numerical analysis of a continuum model of phase transition. *SIAM Journal on Numerical Analysis 28*, 5 (1991), 1310–1322.

[19] Einkemmer, L. A low-rank algorithm for weakly compressible flow. *SIAM Journal on Scientific Computing 41*, 5 (2019), A2795–A2814.

[20] Elliott, C. M., and French, D. A. Numerical studies of the Cahn-Hilliard equation for phase separation. *IMA Journal of Applied Mathematics 38*, 2 (1987), 97–128.

[21] Fernández, D. C. D. R., Boom, P. D., and Zingg, D. W. A generalized framework for nodal first derivative summation-by-parts operators. *Journal of Computational Physics 266* (2014), 214–239.

[22] Filbet, F., and Jin, S. A class of asymptotic-preserving schemes for kinetic equations and related problems with stiff sources. *Journal of Computational Physics 229*, 20 (2010), 7625–7648.

[23] Fjordholm, U. S., Mishra, S., and Tadmor, E. On the computation of measure-valued solutions. *Acta numerica 25* (2016), 567.

[24] Gallis, M., Bitter, N., Koehler, T., Torczynski, J., Plimpton, S., and Papadakis, G. Molecular-level simulations of turbulence and its decay. *Physical Review Letters 118*, 6 (2017), 064501.

[25] Gallis, M., Torczynski, J., Krygier, M., Bitter, N., and Plimpton, S. Turbulence at the edge of continuum. *Physical Review Fluids 6*, 1 (2021), 013401.

[26] Ganeshan, K., and Williams, D. M. A finite element discrete boltzmann method for high knudsen number flows. *arXiv preprint arXiv:2012.13487* (2020).

[27] Gassner, G. J. A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods. *SIAM Journal on Scientific Computing 35*, 3 (2013), A1233–A1253.

[28] Greenshields, C. J., Weller, H. G., Gasparini, L., and Reese, J. M. Implementation of semi-discrete, non-staggered central schemes in a colocated, polyhedral, finite volume framework, for high-speed viscous flows. *International journal for numerical methods in fluids 63*, 1 (2010), 1–21.

[29] Guermond, J.-L., Pasquetti, R., and Popov, B. Entropy viscosity method for nonlinear conservation laws. *Journal of Computational Physics 230*, 11 (2011), 4248–4267.

[30] Harten, A. "On the symmetric form of systems of conservation laws with entropy". *Journal of Computational Physics 49*, 1 (1983), 151 – 164.

[31] Harten, A., Hyman, J. M., Lax, P. D., and Keyfitz, B. On finite-difference approximations and entropy conditions for shocks. *Communications on pure and applied mathematics 29*, 3 (1976), 297–322.

[32] Harten, A., Lax, P. D., and Leer, B. v. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM review 25*, 1 (1983), 35–61.

[33] Hesthaven, J. S. *Numerical methods for conservation laws: From analysis to algorithms*. SIAM, 2017.

[34] Hesthaven, J. S., and Warburton, T. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*. Springer Science & Business Media, 2007.

[35] Hu, J., and Wang, Y. An adaptive dynamical low rank method for the nonlinear boltzmann equation. *arXiv preprint arXiv:2112.02695* (2021).

[36] Hu, Z., and Cai, Z. Burnett spectral method for high-speed rarefied gas flows. *SIAM Journal on Scientific Computing 42*, 5 (2020), B1193–B1226.

[37] Ismail, F., and Roe, P. L. Affordable, entropy-consistent Euler flux functions II: Entropy production at shocks. *Journal of Computational Physics 228*, 15 (2009), 5410–5436.

[38] Jaiswal, S. Isogeometric schemes in rarefied gas dynamics context. *Computer Methods in Applied Mechanics and Engineering 383* (2021), 113926.

[39] Jaiswal, S. Non-linear Boltzmann equation on hybrid-unstructured non-conforming multi-domains. *Journal of Computational Physics 450* (2022), 110687.

[40] Jaiswal, S., Alexeenko, A. A., and Hu, J. A discontinuous Galerkin fast spectral method for the full Boltzmann equation with general collision kernels. *Journal of Computational Physics 378* (2019), 178–208.

[41] Jaiswal, S., Alexeenko, A. A., and Hu, J. A discontinuous Galerkin fast spectral method for the multi-species Boltzmann equation. *Computer Methods in Applied Mechanics and Engineering 352* (2019), 56–84.

[42] Jaiswal, S., Hu, J., Brillon, J. K., and Alexeenko, A. A. A discontinuous Galerkin fast spectral method for multi-species full Boltzmann on streaming multi-processors. In *Proceedings of the Platform for Advanced Scientific Computing Conference* (2019), pp. 1–9.

[43] Jiang, G. S., and Shu, C.-W. On a cell entropy inequality for discontinuous Galerkin methods. *Mathematics of Computation 62*, 206 (1994), 531–538.

[44] Jin, S., and Xin, Z. The relaxation schemes for systems of conservation laws in arbitrary space dimensions. *Communications on pure and applied mathematics 48*, 3 (1995), 235–276.
[45] Karniadakis, G., and Sherwin, S. *Spectral/hp element methods for computational fluid dynamics*. Oxford University Press, 2013.
[46] Krivodonova, L. Limiters for high-order discontinuous Galerkin methods. *Journal of Computational Physics 226*, 1 (2007), 879–896.
[47] Kurganov, A., and Tadmor, E. New high-resolution central schemes for nonlinear conservation laws and convection–diffusion equations. *Journal of Computational Physics 160*, 1 (2000), 241–282.
[48] Lax, P. D. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Communications on pure and applied mathematics 7*, 1 (1954), 159–193.
[49] Lax, P. D. The formation and decay of shock waves. *The American Mathematical Monthly 79*, 3 (1972), 227–241.
[50] Lax, P. D. On dispersive difference schemes. *Physica D: Nonlinear Phenomena 18*, 1-3 (1986), 250–254.
[51] Lefloch, P. G., Mercier, J.-M., and Rohde, C. Fully discrete, entropy conservative schemes of arbitrary order. *SIAM Journal on Numerical Analysis 40*, 5 (2002), 1968–1992.
[52] Levermore, C. D., and Liu, J.-G. Oscillations arising in numerical experiments. In *Singular Limits of Dispersive Waves*. Springer, 1994, pp. 329–346.
[53] Lions, P.-L., Perthame, B., and Tadmor, E. A kinetic formulation of multidimensional scalar conservation laws and related equations. *Journal of the American Mathematical Society 7*, 1 (1994), 169–191.
[54] Mock, M. S. Systems of conservation laws of mixed type. *Journal of Differential equations 37*, 1 (1980), 70–88.
[55] Mouhot, C., and Pareschi, L. Fast algorithms for computing the boltzmann collision operator. *Mathematics of computation 75*, 256 (2006), 1833–1852.
[56] Mouhot, C., Pareschi, L., and Rey, T. Convolutive decomposition and fast summation methods for discrete-velocity approximations of the Boltzmann equation. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique 47*, 5 (2013), 1515–1531.
[57] Osher, S. Riemann solvers, the entropy condition, and difference. *SIAM Journal on Numerical Analysis 21*, 2 (1984), 217–235.
[58] Pareschi, L., and Russo, G. Numerical solution of the boltzmann equation i: Spectrally accurate approximation of the collision operator. *SIAM journal on numerical analysis 37*, 4 (2000), 1217–1245.
[59] Pareschi, L., and Russo, G. Implicit–explicit Runge–Kutta schemes and applications to hyperbolic systems with relaxation. *Journal of Scientific computing 25*, 1 (2005), 129–155.
[60] Plimpton, S., Moore, S., Borner, A., Stagg, A., Koehler, T., Torczynski, J., and Gallis, M. Direct simulation Monte Carlo on petaflop supercomputers and beyond. *Physics of Fluids 31*, 8 (2019), 086101.
[61] Tadmor, E. The numerical viscosity of entropy stable schemes for systems of conservation laws. I. *Mathematics of Computation 49*, 179 (1987), 91–103.
[62] Tadmor, E. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica 12*, 1 (2003), 451–512.
[63] Tadmor, E., and Zhong, W. Entropy stable approximations of Navier–Stokes equations with no artificial numerical viscosity. *Journal of Hyperbolic Differential Equations 3*, 03 (2006), 529–559.
[64] Van Heyningen, R. L. *Discontinuous Galerkin solutions of the Boltzmann equation: spectral collocation and moment methods*. PhD thesis, Massachusetts Institute of Technology, 2021.
[65] Villani, C. A review of mathematical topics in collisional kinetic theory. *Handbook of mathematical fluid dynamics 1*, 71-305 (2002), 3–8.
[66] Wagner, W. A convergence proof for Bird's direct simulation Monte Carlo method for the Boltzmann equation. *Journal of Statistical Physics 66*, 3-4 (1992), 1011–1044.
[67] Zhang, X., and Shu, C.-W. On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *Journal of Computational Physics 229*, 23 (2010), 8918–8934.
[68] Zhang, X., and Shu, C.-W. Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms. *Journal of Computational Physics 230*, 4 (2011), 1238–1248.

(S. Jaiswal) School of Aeronautics & Astronautics; and Department of Physics, Purdue University, West Lafayette, Indiana 47907, USA

Email address: jaiswal0@purdue.edu