SIAM J. Sci. Comput.
Vol. 44, No. 3, pp. A1497–A1524

A FAST PETROV–GALERKIN SPECTRAL METHOD FOR THE MULTIDIMENSIONAL BOLTZMANN EQUATION USING MAPPED CHEBYSHEV FUNCTIONS[*]

JINGWEI HU[†], XIAODONG HUANG[‡], JIE SHEN[‡], AND HAIZHAO YANG[‡]

Abstract. Numerical approximation of the Boltzmann equation presents a challenging problem due to its high-dimensional, nonlinear, and nonlocal collision operator. Among the deterministic methods, the Fourier–Galerkin spectral method stands out for its relative high accuracy and possibility of being accelerated by the fast Fourier transform. However, this method requires a domain truncation which is unphysical since the collision operator is defined in $\mathbb{R}^d$. In this paper, we introduce a Petrov–Galerkin spectral method for the Boltzmann equation in the unbounded domain. The basis functions (both test and trial functions) are carefully chosen mapped Chebyshev functions to obtain desired convergence and conservation properties. Furthermore, thanks to the close relationship of the Chebyshev functions and the Fourier cosine series, we are able to construct a fast algorithm with the help of the nonuniform fast Fourier transform. We demonstrate the superior accuracy of the proposed method in comparison to the Fourier spectral method through a series of two-dimensional and three-dimensional examples.

Key words. Boltzmann equation, Petrov–Galerkin spectral method, mapped Chebyshev function, unbounded domain, NUFFT

AMS subject classifications. 35Q20, 65M70

DOI. 10.1137/21M1420721

1. Introduction.

In multiscale modeling, kinetic theory serves as a basic building block that bridges microscopic particle models and macroscopic continuum models. By tracking the probability density function, kinetic equations describe the nonequilibrium dynamics of the complex particle systems and have been widely used in disparate fields such as rarefied gas dynamics [10], plasma physics [6], nuclear reactor modeling [11], chemistry [20], biology, and socioeconomics [33].

In this paper, we consider the numerical approximation of the Boltzmann equation, one of the fundamental equations in kinetic theory [9, 42]. The complete equation includes both particle transport and collisions, which are often treated separately by operator splitting. Since the collision part is the main difficulty when numerically solving the equation, we focus on the following spatially homogeneous Boltzmann equation:

$$\partial_t f = Q(f,f), \quad t>0, \quad \boldsymbol{v} \in \mathbb{R}^d, \quad d=2,3, \tag{1.1}$$

[*]Submitted to the journal's Methods and Algorithms for Scientific Computing section May 18, 2021; accepted for publication (in revised form) February 10, 2022; published electronically June 13, 2022.
https://doi.org/10.1137/21M1420721
Funding: The first and second authors are partially supported under the NSF CAREER grant DMS-2153208 and NSF CDS&E grant CBET-1854829. The third author is partially supported by NSF DMS-2012585 and AFOSR FA9550-20-1-0309. The fourth author is partially supported by NSF CAREER grant DMS-1945029.

[†]Department of Applied Mathematics, University of Washington, Seattle, WA 98195 USA (hujw@uw.edu).
[‡]Department of Mathematics, Purdue University, West Lafayette, IN 47907 USA (huan1178@purdue.edu, shen7@purdue.edu, yang1863@purdue.edu).

where $f = f(t, \boldsymbol{v})$ is the probability density function at time t and velocity $\boldsymbol{v}$, and $Q(f, f)$ is the collision operator whose bilinear form is given by

$$Q(g, f)(\boldsymbol{v}) = \int_{\mathbb{R}^d} \int_{S^{d-1}} \mathcal{B}(\boldsymbol{v}, \boldsymbol{v}_*, \boldsymbol{\sigma}) \left[g(\boldsymbol{v}'_*) f(\boldsymbol{v}') - g(\boldsymbol{v}_*) f(\boldsymbol{v})\right] \mathrm{d}\boldsymbol{\sigma} \, \mathrm{d}\boldsymbol{v}_*, \tag{1.2}$$

where the postcollisional velocities $(\boldsymbol{v}', \boldsymbol{v}'_*)$ are defined in terms of the precollisional velocities $(\boldsymbol{v}, \boldsymbol{v}_*)$ as

$$\begin{cases} \boldsymbol{v}' = \frac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_*) + \frac{1}{2}|\boldsymbol{v} - \boldsymbol{v}_*|\boldsymbol{\sigma}, \\ \boldsymbol{v}'_* = \frac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_*) - \frac{1}{2}|\boldsymbol{v} - \boldsymbol{v}_*|\boldsymbol{\sigma} \end{cases} \tag{1.3}$$

with $\boldsymbol{\sigma}$ being a vector over the unit sphere S^{d-1}. The collision kernel $\mathcal{B}$ takes the form

$$\mathcal{B}(\boldsymbol{v}, \boldsymbol{v}_*, \boldsymbol{\sigma}) = B(|\boldsymbol{v} - \boldsymbol{v}_*|, \cos\theta), \quad \cos\theta = \left\langle \frac{\boldsymbol{v} - \boldsymbol{v}_*}{|\boldsymbol{v} - \boldsymbol{v}_*|}, \boldsymbol{\sigma} \right\rangle, \tag{1.4}$$

i.e., it is a function depending only on the relative velocity $|\boldsymbol{v} - \boldsymbol{v}_*|$ and cosine of the scattering angle. The collision operator $Q(f, f)$ satisfies many important physical properties, including conservation of mass, momentum, and energy,

$$\int_{\mathbb{R}^d} Q(f, f) \, \mathrm{d}\boldsymbol{v} = \int_{\mathbb{R}^d} Q(f, f)\boldsymbol{v} \, \mathrm{d}\boldsymbol{v} = \int_{\mathbb{R}^d} Q(f, f)|\boldsymbol{v}|^2 \, \mathrm{d}\boldsymbol{v} = 0, \tag{1.5}$$

and the Boltzmann's H-theorem,

$$\int_{\mathbb{R}^d} Q(f, f) \log f \, \mathrm{d}\boldsymbol{v} \le 0. \tag{1.6}$$

In the physically relevant case ($d = 3$), the collision operator is a fivefold quadratic integral whose numerical approximation can be extremely challenging. The stochastic methods, such as the direct simulation Monte Carlo (DSMC) methods proposed by Nanbu [34] and Bird [4], have been historically popular due to their simplicity and efficiency. However, like any Monte Carlo method, they suffer from slow convergence and high statistical noise, especially for low-speed and unsteady flows. In the past two decades, the deterministic methods have undergone extensive development largely due to the advance in computing powers; see [12] for a recent review.

Among the deterministic methods for the Boltzmann equation, the Fourier–Galerkin spectral method stands out for its relatively high accuracy and the possibility of being accelerated by the fast Fourier transform. Some relevant works in this direction explore the structure of the collision operator in the Fourier domain [8, 7]. The method was formally formulated in [35, 36]; see [31, 17, 16] for major algorithmic development, [37, 13, 1, 23] for stability and convergence analysis, and [14, 19, 27] for extension to the spatially inhomogeneous case. Although being a method with reasonable efficiency and accuracy tradeoff, the Fourier spectral method requires a domain truncation which is unphysical since the original collision operator is defined in $\mathbb{R}^d$. This truncation changes the structure of the equation and often comes with an accuracy loss.

Inspired by the recent work [24] of the two authors here, where a spectral method was introduced for the one-dimensional (1D) inelastic Boltzmann equation, we develop in this paper a Petrov–Galerkin spectral method for the Boltzmann equation (1.1)[1]

[1] Unlike the inelastic Boltzmann equation, which has a nontrivial solution in one dimension, the classical Boltzmann equation (1.1) must be considered at least for $d \ge 2$.

using mapped Chebyshev functions in $\mathbb{R}^d$. Both the test functions and trial functions are carefully chosen to obtain desired approximation properties. Furthermore, thanks to the close relationship between the Chebyshev functions and the Fourier cosine series, we are able to construct a fast algorithm with the help of the nonuniform fast Fourier transform (NUFFT). This speedup is critical as the direct implementation of the proposed method would require excessive storage for precomputation and a significant online computational cost that soon becomes a bottleneck for larger N (the number of spectral modes). Extensive numerical tests in two and three dimensions are performed to demonstrate the accuracy and efficiency of the proposed method. In particular, since the new method does not require a truncation in the velocity space, it offers much better accuracy and conservation properties compared to the Fourier spectral method in [16].

Finally, we mention some recent spectral methods for the Boltzmann equation that use other orthogonal polynomial bases in $\mathbb{R}^d$ (e.g., Hermite polynomials, Burnett polynomials, Laguerre polynomials, spherical harmonics, or a combination of them) [41, 18, 15, 22, 28, 25, 26]. However, the numerical realization of the high-dimensional problems is quite challenging for these methods and some numerical strategies are proposed to balance the workload and accuracy. In comparison to these methods, our method is the first to use the mapped Chebyshev functions and take advantage of their connection to the fast Fourier transform to come up with a fast algorithm. Furthermore, to our knowledge, we are also the first to provide a consistency analysis of the method on the unbounded domain.

The rest of this paper is organized as follows. In section 2, we introduce the mapped Chebyshev functions in $\mathbb{R}^d$ along with their approximation properties. In section 3, we construct the Petrov–Galerkin spectral method for the Boltzmann equation using the mapped Chebyshev functions as trial and test functions. The approximation properties for the collision operator and moments are proved as well. The numerical realization including the fast algorithm is described in detail in section 4. In section 5, several numerical tests in two and three dimensions are performed to demonstrate the accuracy and efficiency of the proposed method. The paper is concluded in section 6.

2. Multidimensional mapped Chebyshev functions.

In this section, we introduce the mapped Chebyshev functions in $\mathbb{R}^d$ and discuss their approximation properties. These functions are an extension of the 1D mapped Chebyshev functions introduced in [24] based on tensor product formulation [38, 40]. Later in section 3, they will serve as the trial functions and test functions in the Petrov–Galerkin spectral method for the Boltzmann equation.

2.1. Mapped Chebyshev functions in $\mathbb{R}^d$.

To define the mapped Chebyshev functions in $\mathbb{R}^d$, we start with the 1D Chebyshev polynomials on the interval $I = (-1, 1)$:

$$T_0(\xi) = 1, \quad T_1(\xi) = \xi, \quad T_{k+1}(\xi) = 2\xi T_k(\xi) - T_{k-1}(\xi), \quad k \geq 1. \tag{2.1}$$

It is well-known that these polynomials are related to trigonometric functions:

$$T_k(\xi) = \cos(k \arccos(\xi)) \quad \forall \xi \in I. \tag{2.2}$$

Define the inner product $(\cdot, \cdot)_\omega$ as

$$(F, G)_\omega := \int_I F(\xi) G(\xi) \omega(\xi) \, \mathrm{d}\xi, \quad \omega(\xi) = (1 - \xi^2)^{-\frac{1}{2}}, \tag{2.3}$$

and then $\{T_k(\xi)\}_{k\geq 0}$ satisfy the orthogonality condition

$$(T_k, T_l)_\omega = c_k \delta_{k,l} \quad \forall\ k,l \geq 0, \tag{2.4}$$

where $c_0 = \pi$ and $c_k = \pi/2$ for $k \geq 1$.

Notice that Chebyshev functions are defined on a bounded domain. To deal with the unbounded domain problem, we introduce a one-to-one mapping $\xi \to v(\xi)$ (its inverse is denoted as $v \to \xi(v)$) from I to $\mathbb{R}$ such that

$$\frac{\mathrm{d}v}{\mathrm{d}\xi} = \frac{S}{(1-\xi^2)^{1+\frac{r}{2}}} := \frac{\omega(\xi)}{[\mu(\xi)]^2}, \quad v(\pm 1) = \pm\infty, \tag{2.5}$$

where $S > 0$ is a scaling parameter, $r \geq 0$ is the tail parameter, and the function μ is given by

$$\mu(\xi) = \frac{(1-\xi^2)^{\frac{1+r}{4}}}{\sqrt{S}}. \tag{2.6}$$

Using this mapping, the unbounded integral in the collision operator (1.2) can be transformed into a bounded domain. Here the scaling parameter S will play a key role in our method since it determines the distribution of mapped quadrature points. Its influence on the numerical accuracy will be investigated in section 5.

With this mapping we define two sets of mapped Chebyshev functions in $\mathbb{R}$ as

$$\widetilde{T}_k(v) := \frac{[\mu(\xi(v))]^4}{\sqrt{c_k}} T_k(\xi(v)), \quad \widehat{T}_k(v) := \frac{[\mu(\xi(v))]^{-2}}{\sqrt{c_k}} T_k(\xi(v)). \tag{2.7}$$

Define the inner product $(\cdot,\cdot)_{\mathbb{R}}$ as

$$(f,g)_{\mathbb{R}} := \int_{\mathbb{R}} f(v) g(v)\,\mathrm{d}v, \tag{2.8}$$

and then it is easy to check that $\{\widetilde{T}_k(v)\}_{k\geq 0}$ and $\{\widehat{T}_k(v)\}_{k\geq 0}$ satisfy the orthonormal condition:

$$(\widetilde{T}_k, \widehat{T}_l)_{\mathbb{R}} = \delta_{k,l} \quad \forall\ k,l \geq 0. \tag{2.9}$$

Remark 2.1. Here we present two one-to-one mappings between I and $\mathbb{R}$ that are of the above type:

- *logarithmic mapping* ($r = 0$):

$$v = \frac{S}{2} \ln\left(\frac{1+\xi}{1-\xi}\right), \quad \xi = \tanh\left(\frac{v}{S}\right), \quad \mu(\xi) = \frac{1}{\sqrt{S}}(1-\xi^2)^{\frac{1}{4}}, \tag{2.10}$$

- *algebraic mapping* ($r = 1$):

$$v = \frac{S\xi}{\sqrt{1-\xi^2}}, \quad \xi = \frac{v}{\sqrt{S^2+v^2}}, \quad \mu(\xi) = \frac{1}{\sqrt{S}}(1-\xi^2)^{\frac{1}{2}}. \tag{2.11}$$

These two mappings are commonly used ones to transform between I and $\mathbb{R}$. Several other mappings are discussed in Chapter 7.5 of [38].

In the multidimensional case, we denote the multivector as $\boldsymbol{v} = (v_1, \ldots, v_d)$ and multi-index as $\boldsymbol{k} = (k_1, \ldots, k_d)$, where k_j is a nonnegative integer for each $j = 1, \ldots, d$; $0 \le \boldsymbol{k} \le N$ means $0 \le k_j \le N$ for each $j = 1, \ldots, d$. We define the mapped Chebyshev functions in $\mathbb{R}^d$ using (2.7) via the tensor product as

$$\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}) := \prod_{j=1}^{d} \widetilde{T}_{k_j}(v_j), \quad \widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}) := \prod_{j=1}^{d} \widehat{T}_{k_j}(v_j). \tag{2.12}$$

The inner products $(\cdot,\cdot)_{\boldsymbol{\omega}}$ in $I^d = (-1,1)^d$ and $(\cdot,\cdot)_{\mathbb{R}^d}$ in $\mathbb{R}^d$ are defined, respectively, by

$$(F,G)_{\boldsymbol{\omega}} := \int_{I^d} F(\boldsymbol{\xi}) G(\boldsymbol{\xi}) \boldsymbol{\omega}(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}, \quad (f,g)_{\mathbb{R}^d} := \int_{\mathbb{R}^d} f(\boldsymbol{v}) g(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} \tag{2.13}$$

with the weight function $\boldsymbol{\omega}(\boldsymbol{\xi}) := \prod_{j=1}^{d} \omega(\xi_j)$. Then we still have the orthogonality

$$\left(\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}, \widehat{\boldsymbol{T}}_{\boldsymbol{l}}\right)_{\mathbb{R}^d} = \prod_{j=1}^{d} \left(\widetilde{T}_{k_j}, \widehat{T}_{l_j}\right)_{\mathbb{R}} = \prod_{j=1}^{d} \delta_{k_j, l_j} =: \boldsymbol{\delta}_{\boldsymbol{k},\boldsymbol{l}}. \tag{2.14}$$

Suppose that a d-variate function $f(\boldsymbol{v})$ can be expanded by $\{\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\}_{\boldsymbol{k}\ge 0}$ as

$$f(\boldsymbol{v}) = \sum_{\boldsymbol{k} \ge 0} \widetilde{f}_{\boldsymbol{k}} \widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}) = \sum_{\boldsymbol{k} \ge 0} \widetilde{f}_{\boldsymbol{k}} \frac{[\boldsymbol{\mu}(\boldsymbol{\xi})]^4}{\sqrt{\boldsymbol{c}_{\boldsymbol{k}}}} \boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi}), \tag{2.15}$$

and the expansion coefficients $\{\widetilde{f}_{\boldsymbol{k}}\}_{\boldsymbol{k}\ge 0}$ are determined by

$$\widetilde{f}_{\boldsymbol{k}} = \left(f, \widehat{\boldsymbol{T}}_{\boldsymbol{k}}\right)_{\mathbb{R}^d} = \frac{1}{\sqrt{\boldsymbol{c}_{\boldsymbol{k}}}} \left([\boldsymbol{\mu}(\boldsymbol{\xi})]^{-4} f(\boldsymbol{v}(\boldsymbol{\xi})), \boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi})\right)_{\boldsymbol{\omega}}, \tag{2.16}$$

where

$$\boldsymbol{\mu}(\boldsymbol{\xi}) := \prod_{j=1}^{d} \mu(\xi_j), \quad \boldsymbol{c}_{\boldsymbol{k}} := \prod_{j=1}^{d} c_{k_j}, \quad \boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi}) := \prod_{j=1}^{d} T_{k_j}(\xi_j), \tag{2.17}$$

and $\boldsymbol{v}(\boldsymbol{\xi})$ is the mapping from I^d to $\mathbb{R}^d$ such that each component ξ_j is mapped to v_j via the 1D mapping (2.5). The inverse mapping $\boldsymbol{\xi}(\boldsymbol{v})$ is understood similarly.

In section 3, we will introduce the Petrov–Galerkin spectral method for the Boltzmann equation in $\mathbb{R}^d$, where the trial function space and test function space are chosen, respectively, as

$$\widetilde{\mathbb{T}}_N^d := \{\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\}_{0 \le \boldsymbol{k} \le N}, \quad \widehat{\mathbb{T}}_N^d := \{\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\}_{0 \le \boldsymbol{k} \le N}. \tag{2.18}$$

The choice of these functions is motivated by their decay/growth properties at large $|\boldsymbol{v}|$. The following result is a straightforward extension of the 1D result in [24]; a detailed proof is provided in the appendix.

LEMMA 2.2. *For any $\boldsymbol{k} \ge 0$ and $|\boldsymbol{v}| \gg 1$, we have*

$$\left|\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\right| \sim \begin{cases} e^{-\frac{2}{S}(\sum_{j=1}^{d}|v_j|)}, & r=0, \\ \prod_{j=1}^{d} |v_j|^{-4}, & r=1; \end{cases} \quad \left|\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\right| \sim \begin{cases} e^{\frac{1}{S}(\sum_{j=1}^{d}|v_j|)}, & r=0, \\ \prod_{j=1}^{d} |v_j|^{2}, & r=1, \end{cases} \tag{2.19}$$

where $r = 0$ corresponds to the logarithmic mapping (2.10) *and $r = 1$ to the algebraic mapping* (2.11).

2.2. Approximation properties. We describe below some approximation properties of the mapped Chebyshev functions in $\mathbb{R}^d$.

For a function $f(\boldsymbol{v})$ defined in $\mathbb{R}^d$, the transform $\boldsymbol{v}(\boldsymbol{\xi})$ maps it to a function in I^d. Hence, we introduce the linked function pair (f, F) such that $f(\boldsymbol{v}) = f(\boldsymbol{v}(\boldsymbol{\xi})) \equiv F(\boldsymbol{\xi})$. In addition, we introduce another function pair $(\widehat{f}^\alpha, \widehat{F}^\alpha)$ as

$$\widehat{f}^\alpha(\boldsymbol{v}) := f(\boldsymbol{v})[\boldsymbol{\mu}(\boldsymbol{\xi}(\boldsymbol{v}))]^{-\alpha} = F(\boldsymbol{\xi})[\boldsymbol{\mu}(\boldsymbol{\xi})]^{-\alpha} =: \widehat{F}^\alpha(\boldsymbol{\xi}). \tag{2.20}$$

We define the approximation space in $\mathbb{R}^d$ with a parameter α as

$$\mathbb{V}_N^{\alpha,d}(\mathbb{R}^d) := \operatorname{span}\left\{\boldsymbol{T}_{\boldsymbol{k}}^\alpha(\boldsymbol{v}) := [\boldsymbol{\mu}(\boldsymbol{\xi}(\boldsymbol{v}))]^\alpha \boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi}(\boldsymbol{v})),\ 0 \le \boldsymbol{k} \le N\right\}. \tag{2.21}$$

Therefore, the trial function space $\widetilde{\mathbb{T}}_N^d$ and test function space $\widehat{\mathbb{T}}_N^d$ introduced in the previous section correspond to $\mathbb{V}_N^{4,d}$ and $\mathbb{V}_N^{-2,d}$, respectively.

In the following, the L^2 space with a given weight $\boldsymbol{w}$ is equipped with norm

$$\|f\|_{L^2_{\boldsymbol{w}}(I^d)} = \left(\int_{I^d} |f(\boldsymbol{\xi})|^2 \boldsymbol{w}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\right)^{1/2} \quad \text{or} \quad \|f\|_{L^2_{\boldsymbol{w}}(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |f(\boldsymbol{v})|^2 \boldsymbol{w}(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}\right)^{1/2}, \tag{2.22}$$

depending on the domain of interest.

Let $\mathbb{P}_N^d(I^d)$ denote the set of d-variate polynomials in I^d with degree $\le N$ in each direction, and $\Pi_N^d : L^2_{\boldsymbol{\omega}}(I^d) \to \mathbb{P}_N^d(I^d)$ be the Chebyshev orthogonal projection operator such that

$$\left(\Pi_N^d F - F, \phi\right)_{\boldsymbol{\omega}} = 0 \quad \forall \phi \in \mathbb{P}_N^d(I^d). \tag{2.23}$$

Then we define another projection operator $\pi_N^{\alpha,d} : L^2_{\boldsymbol{\mu}^{2-2\alpha}}(\mathbb{R}^d) \to \mathbb{V}_N^{\alpha,d}(\mathbb{R}^d)$ by

$$\pi_N^{\alpha,d} f := \boldsymbol{\mu}^\alpha \Pi_N^d(F\boldsymbol{\mu}^{-\alpha}) = \boldsymbol{\mu}^\alpha \Pi_N^d \widehat{F}^\alpha. \tag{2.24}$$

One can verify using the definition that

$$\begin{aligned}\left(\pi_N^{\alpha,d} f - f, \boldsymbol{\mu}^{2-2\alpha}\boldsymbol{T}_{\boldsymbol{k}}^\alpha\right)_{\mathbb{R}^d} &= \int_{\mathbb{R}^d} (\pi_N^{\alpha,d} f - f)\boldsymbol{\mu}^{2-\alpha}\boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi}(\boldsymbol{v}))\,\mathrm{d}\boldsymbol{v}\\ &= \int_{I^d}\left[\boldsymbol{\mu}^\alpha \Pi_N^d \widehat{F}^\alpha - \boldsymbol{\mu}^\alpha \widehat{F}^\alpha\right]\boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi})\boldsymbol{\mu}^{2-\alpha}\frac{\boldsymbol{\omega}(\boldsymbol{\xi})}{\boldsymbol{\mu}^2}\,\mathrm{d}\boldsymbol{\xi}\\ &= \left(\Pi_N^d \widehat{F}^\alpha - \widehat{F}^\alpha, \boldsymbol{T}_{\boldsymbol{k}}\right)_{\boldsymbol{\omega}} = 0 \quad \forall\ 0 \le \boldsymbol{k} \le N.\end{aligned} \tag{2.25}$$

Next, we introduce the function space $\boldsymbol{B}_\alpha^m(\mathbb{R}^d)$ equipped with the norm

$$\|f\|_{\boldsymbol{B}_\alpha^m(\mathbb{R}^d)} = \left(\sum_{0 \le \boldsymbol{k} \le m} \left\|\boldsymbol{D}_{\alpha,\boldsymbol{v}}^{\boldsymbol{k}} f\right\|^2_{L^2_{\boldsymbol{\varpi}^{\boldsymbol{k}+\frac{1+r}{2}\mathbf{1}}}(\mathbb{R}^d)}\right)^{1/2} \tag{2.26}$$

and seminorm

$$|f|_{\boldsymbol{B}_\alpha^m(\mathbb{R}^d)} = \left(\sum_{j=1}^d \left\|D_{\alpha,v_j}^m f\right\|^2_{L^2_{\boldsymbol{\varpi}^{m\boldsymbol{e}_j+\frac{1+r}{2}\mathbf{1}}}(\mathbb{R}^d)}\right)^{1/2}, \tag{2.27}$$

where $\mathbf{1}$ is an all-one vector, $\boldsymbol{e}_j = (0, \ldots, 1, \ldots, 0)$ with 1 in the jth position and 0 elsewhere, and

$$\boldsymbol{D}^{\boldsymbol{k}}_{\alpha,\boldsymbol{v}} f := D^{k_1}_{\alpha,v_1} \cdots D^{k_d}_{\alpha,v_d} f, \quad \boldsymbol{\varpi}^{\boldsymbol{k}} := \prod_{j=1}^{d} (1 - \xi(v_j)^2)^{k_j} \tag{2.28}$$

with

$$D^{k_j}_{\alpha,v_j} f := \underbrace{a(v_j) \frac{\partial}{\partial v_j} \left(a(v_j) \frac{\partial}{\partial v_j} \left(\ldots \left(a(v_j) \frac{\partial \widehat{f}^{\alpha}}{\partial v_j} \right) \ldots \right) \right)}_{k_j \text{ times derivatives}} = \frac{\partial^{k_j} \widehat{F}^{\alpha}}{\partial \xi_j}, \tag{2.29}$$

where $a(v_j) := \frac{\mathrm{d}v_j}{\mathrm{d}\xi_j}$ is determined by the mapping.

We have the following approximation result.

Theorem 2.3. *Let $\alpha \in \mathbb{R}$, $r \geq 0$. If $f \in \boldsymbol{B}^m_{\alpha}(\mathbb{R}^d)$, we have*

$$\left\| \pi_N^{\alpha,d} f - f \right\|_{L^2_{\boldsymbol{\mu}^{2-2\alpha}}(\mathbb{R}^d)} \leq C N^{-m} \left| f \right|_{\boldsymbol{B}^m_{\alpha}(\mathbb{R}^d)} . \tag{2.30}$$

Proof. Note that

$$\begin{aligned}
\left\| \pi_N^{\alpha,d} f - f \right\|^2_{L^2_{\boldsymbol{\mu}^{2-2\alpha}}(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} (\pi_N^{\alpha,d} f - f)^2 \boldsymbol{\mu}^{2-2\alpha} \, \mathrm{d}\boldsymbol{v} \\
&= \int_{I^d} \left[\boldsymbol{\mu}^{\alpha} \Pi_N^d \widehat{F}^{\alpha} - \boldsymbol{\mu}^{\alpha} \widehat{F}^{\alpha} \right]^2 \boldsymbol{\mu}^{2-2\alpha} \frac{\boldsymbol{\omega}(\boldsymbol{\xi})}{\boldsymbol{\mu}^2} \, \mathrm{d}\boldsymbol{\xi} \\
&= \left\| \Pi_N^d \widehat{F}^{\alpha} - \widehat{F}^{\alpha} \right\|^2_{L^2_{\boldsymbol{\omega}}(I^d)} .
\end{aligned}$$

By the multivariate (full tensor product) Chebyshev approximation result (Theorem 2.1 in [39]), we know

$$\left\| \Pi_N^d \widehat{F}^{\alpha} - \widehat{F}^{\alpha} \right\|_{L^2_{\boldsymbol{\omega}}(I^d)} \leq C N^{-m} \left(\sum_{j=1}^{d} \left\| \partial^m_{\xi_j} \widehat{F}^{\alpha} \right\|^2_{L^2_{\boldsymbol{\varpi}^{m\boldsymbol{e}_j - \frac{1}{2}\mathbf{1}}}(I^d)} \right)^{1/2} .$$

Hence,

$$\begin{aligned}
&\left\| \pi_N^{\alpha,d} f - f \right\|_{L^2_{\boldsymbol{\mu}^{2-2\alpha}}(\mathbb{R}^d)} = \left\| \Pi_N^d \widehat{F}^{\alpha} - \widehat{F}^{\alpha} \right\|_{L^2_{\boldsymbol{\omega}}(I^d)} \\
&\leq C N^{-m} \left(\sum_{j=1}^{d} \left\| \partial^m_{\xi_j} \widehat{F}^{\alpha} \right\|^2_{L^2_{\boldsymbol{\varpi}^{m\boldsymbol{e}_j - \frac{1}{2}\mathbf{1}}}(I^d)} \right)^{1/2} \\
&\leq C N^{-m} \left(\sum_{j=1}^{d} \left\| D^m_{\alpha,v_j} f \right\|^2_{L^2_{\boldsymbol{\varpi}^{m\boldsymbol{e}_j + \frac{1+r}{2}\mathbf{1}}}(\mathbb{R}^d)} \right)^{1/2} \\
&= C N^{-m} \left| f \right|_{\boldsymbol{B}^m_{\alpha}(\mathbb{R}^d)} .
\end{aligned}$$

□

3. A Petrov–Galerkin spectral method for the Boltzmann equation.

We consider the initial value problem

$$\begin{cases} \partial_t f(t,\boldsymbol{v}) = Q(f,f), \quad t>0, \quad \boldsymbol{v}\in\mathbb{R}^d, \\ f(0,\boldsymbol{v}) = f^0(\boldsymbol{v}), \end{cases} \tag{3.1}$$

where $Q(f,f)$, in a strong form, is given by (1.2). To construct the Petrov–Galerkin spectral method, the following weak form of the collision operator is more convenient:

$$\begin{aligned} (Q(f,f),\phi)_{\mathbb{R}^d} &= \int_{\mathbb{R}^d} Q(f,f)(\boldsymbol{v})\phi(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v} \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma}) f(\boldsymbol{v}) f(\boldsymbol{v}_*)[\phi(\boldsymbol{v}') - \phi(\boldsymbol{v})]\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_*, \end{aligned} \tag{3.2}$$

where $\phi(\boldsymbol{v})$ is a test function.

We look for an approximation of f in the trial function space $\widetilde{\mathbb{T}}_N^d$ as

$$f(t,\boldsymbol{v}) \approx f_N(t,\boldsymbol{v}) = \sum_{0\le\boldsymbol{k}\le N} \widetilde{f}_{\boldsymbol{k}}(t)\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}) \in \widetilde{\mathbb{T}}_N^d. \tag{3.3}$$

Substituting f_N into (3.1) and requiring the residue of the equation to be orthogonal to the test function space $\widehat{\mathbb{T}}_N^d$, we obtain

$$\left(\partial_t f_N - Q(f_N,f_N), \widehat{\boldsymbol{T}}_{\boldsymbol{k}}\right)_{\mathbb{R}^d} = 0 \quad \forall \widehat{\boldsymbol{T}}_{\boldsymbol{k}} \in \widehat{\mathbb{T}}_N^d. \tag{3.4}$$

By the orthogonality condition (2.14), we find that the coefficients $\{\widetilde{f}_{\boldsymbol{k}}(t)\}$ satisfy the following ODE system:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{f}_{\boldsymbol{k}}(t) = \mathcal{Q}_{\boldsymbol{k}}^N, \\ \widetilde{f}_{\boldsymbol{k}}(0) = \widetilde{f}_{\boldsymbol{k}}^0, \end{cases} \qquad 0\le\boldsymbol{k}\le N, \tag{3.5}$$

where

$$\begin{aligned} \mathcal{Q}_{\boldsymbol{k}}^N &:= \left(Q(f_N,f_N), \widehat{\boldsymbol{T}}_{\boldsymbol{k}}\right)_{\mathbb{R}^d} \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma}) f_N(\boldsymbol{v}) f_N(\boldsymbol{v}_*) \left[\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}') - \widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\right] \mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_* \end{aligned} \tag{3.6}$$

and

$$\widetilde{f}_{\boldsymbol{k}}^0 := \left(f^0, \widehat{\boldsymbol{T}}_{\boldsymbol{k}}\right)_{\mathbb{R}^d} = \frac{1}{\sqrt{\boldsymbol{c}_{\boldsymbol{k}}}}\left([\boldsymbol{\mu}(\boldsymbol{\xi})]^{-4} f^0(\boldsymbol{v}(\boldsymbol{\xi})), \boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi})\right)_{\boldsymbol{\omega}}. \tag{3.7}$$

Note that we used the weak form (3.2) in (3.6).

Remark 3.1. An equivalent way of writing the ODE system (3.5) is

$$\begin{cases} \partial_t f_N(t,\boldsymbol{v}) = \pi_N^{4,d} Q(f_N,f_N), \\ f_N(0,\boldsymbol{v}) = \pi_N^{4,d} f^0(\boldsymbol{v}), \end{cases} \tag{3.8}$$

where $\pi_N^{4,d}$ is the projection operator defined in (2.24) (with $\alpha=4$). Indeed, for any $f\in L^2_{\boldsymbol{\mu}^{-6}}(\mathbb{R}^d)$,

$$\pi_N^{4,d} f = \sum_{0\le\boldsymbol{k}\le N} \left(f, \widehat{\boldsymbol{T}}_{\boldsymbol{k}}\right)_{\mathbb{R}^d} \widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}) \in \mathbb{V}_N^{4,d}(\mathbb{R}^d) = \widetilde{\mathbb{T}}_N^d. \tag{3.9}$$

3.1. Approximation property for the collision operator. In this subsection, we establish a consistency result of the spectral approximation for the collision operator. We will show that if f and $Q(f,f)$ have certain regularity, the proposed approximation of the collision operator $\pi_N^{4,d}Q(\pi_N^{4,d}f, \pi_N^{4,d}f)$ enjoys spectral accuracy.

We will only prove this result under the algebraic mapping (2.11) with $S = 1$, that is, in one dimension,

$$v = \frac{\xi}{\sqrt{1-\xi^2}}, \quad \xi = \frac{v}{\sqrt{1+v^2}}, \quad \mu = \sqrt{1-\xi^2} = \frac{1}{\sqrt{1+v^2}}. \tag{3.10}$$

The reason for this choice is strongly motivated by the existing regularity result of the Boltzmann collision operator under a polynomially weighted Lebesgue norm:

$$\|f\|_{\mathcal{L}_k^p(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |f(\boldsymbol{v})|^p (1+|\boldsymbol{v}|^2)^{kp/2}\,\mathrm{d}\boldsymbol{v}\right)^{1/p}, \quad k \in \mathbb{R},\ 1 \le p < \infty. \tag{3.11}$$

Specifically, we write the collision operator (1.2) as $Q(g,f) = Q^+(g,f) - Q^-(g,f)$, where the gain part and loss part are given by

$$\begin{aligned} Q^+(g,f)(\boldsymbol{v}) &= \int_{\mathbb{R}^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma}) g(\boldsymbol{v}_*') f(\boldsymbol{v}')\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_*, \\ Q^-(g,f)(\boldsymbol{v}) &= \int_{\mathbb{R}^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma}) g(\boldsymbol{v}_*) f(\boldsymbol{v})\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_*. \end{aligned} \tag{3.12}$$

Then we have the following regularity result for the gain operator $Q^+(g,f)$ established in [32], with the additional cut-off assumption: no frontal collision should occur, i.e., $b(\cos\theta)$ should vanish for θ close to π:

$$\exists \theta_b > 0; \quad \operatorname{supp} b(\cos\theta) \subset \{\theta \,|\, 0 \le \theta \le \pi - \theta_b\}. \tag{3.13}$$

THEOREM 3.2 (Theorem 2.1 in [32]). *Let $k, \eta \in \mathbb{R}$, $1 \le p < \infty$, and let the collision kernel $\mathcal{B}$ satisfy the cut-off assumption* (3.13). *Then the following estimate holds:*

$$\|Q^+(g,f)\|_{\mathcal{L}_\eta^p(\mathbb{R}^d)} \le C_{k,\eta,p}(\mathcal{B}) \|g\|_{\mathcal{L}^1_{|k+\eta|+|\eta|}(\mathbb{R}^d)} \|f\|_{\mathcal{L}^p_{k+\eta}(\mathbb{R}^d)}, \tag{3.14}$$

where $C_{k,\eta,p}(\mathcal{B})$ is a constant that depends only on the kernel $\mathcal{B}$ and k, η, and p.

To obtain a similar estimate for the loss operator $Q^-(g,f)$, we restrict ourselves to the variable hard sphere collision model [5], with a little modification to satisfy the cut-off assumption (3.13):

$$\mathcal{B} = b(\cos\theta)|\boldsymbol{v} - \boldsymbol{v}_*|^\lambda, \quad b(\cos\theta) = \begin{cases} C_\lambda, & 0 \le \theta \le \pi - \theta_b, \\ 0, & \text{otherwise}, \end{cases} \tag{3.15}$$

where $0 \le \lambda \le 1$ and C_λ is a positive constant.

Then we have the following result.

PROPOSITION 3.2. *Let $\eta \in \mathbb{R}$, $1 \le p < \infty$, and let the collision kernel take the form* (3.15). *Then the following estimate holds:*

$$\|Q^-(g,f)\|_{\mathcal{L}_\eta^p(\mathbb{R}^d)} \le C_\lambda \|g\|_{\mathcal{L}^1_\lambda(\mathbb{R}^d)} \|f\|_{L^p_{\lambda+\eta}(\mathbb{R}^d)}. \tag{3.16}$$

Proof. Note that

$$|\boldsymbol{v}-\boldsymbol{v}_*| \le |\boldsymbol{v}|+|\boldsymbol{v}_*| = (|\boldsymbol{v}|^2+|\boldsymbol{v}_*|^2+2|\boldsymbol{v}||\boldsymbol{v}_*|)^{1/2} \le \left(1+|\boldsymbol{v}|^2\right)^{1/2}\left(1+|\boldsymbol{v}_*|^2\right)^{1/2}.$$

Then

$$\begin{aligned} Q^-(g,f)(\boldsymbol{v}) &\le C_\lambda f(\boldsymbol{v})\int_{\mathbb{R}^d} g(\boldsymbol{v}_*)|\boldsymbol{v}-\boldsymbol{v}_*|^\lambda \,\mathrm{d}\boldsymbol{v}_* \\ &\le C_\lambda f(\boldsymbol{v})\left(1+|\boldsymbol{v}|^2\right)^{\lambda/2}\left[\int_{\mathbb{R}^d} g(\boldsymbol{v}_*)\left(1+|\boldsymbol{v}_*|^2\right)^{\lambda/2}\,\mathrm{d}\boldsymbol{v}_*\right] \\ &= C_\lambda f(\boldsymbol{v})\left(1+|\boldsymbol{v}|^2\right)^{\lambda/2}\|g\|_{L^1_\lambda(\mathbb{R}^d)}. \end{aligned}$$

Therefore, for any $\eta\in\mathbb{R}$, $1\le p<\infty$,

$$\|Q^-(g,f)\|_{\mathcal{L}^p_\eta(\mathbb{R}^d)} \le C_\lambda\|g\|_{\mathcal{L}^1_\lambda(\mathbb{R}^d)}\|f\|_{L^p_{\lambda+\eta}(\mathbb{R}^d)}. \qquad \square$$

Combining the previous two results, we can obtain the following theorem.

THEOREM 3.3. *Let the collision kernel take the form* (3.15)*, and then the collision operator $Q(g,f)$ satisfies*

$$\|Q(g,f)\|_{\mathcal{L}^2_{3d}(\mathbb{R}^d)} \le C_d(\mathcal{B})\|g\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}, \tag{3.17}$$

where $C_d(\mathcal{B})$ is a constant that depends only on the kernel $\mathcal{B}$ and the dimension d.

Proof. Choosing $k=\lambda$, $\eta=3d$, $p=2$ in (3.14), we have

$$\|Q^+(g,f)\|_{\mathcal{L}^2_{3d}(\mathbb{R}^d)} \le C_{\lambda,3d,2}(\mathcal{B})\|g\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}.$$

Choosing $\eta=3d$, $p=2$ in (3.16), we have

$$\|Q^-(g,f)\|_{\mathcal{L}^2_{3d}(\mathbb{R}^d)} \le C_\lambda\|g\|_{\mathcal{L}^1_\lambda(\mathbb{R}^d)}\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}.$$

Combining both, we obtain

$$\|Q(g,f)\|_{\mathcal{L}^2_{3d}(\mathbb{R}^d)} \le C_d(\mathcal{B})\|g\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}. \qquad \square$$

Before we proceed to the consistency proof, we need the following lemmas.

LEMMA 3.4. *Under the algebraic mapping* (2.11) *with $S=1$, we have*

$$\|f\|_{\mathcal{L}^2_\eta(\mathbb{R}^d)} \le \|f\|_{L^2_{\boldsymbol{\mu}^{-2\eta}}(\mathbb{R}^d)} \le \|f\|_{\mathcal{L}^2_{d\eta}(\mathbb{R}^d)} \quad \textit{for any } \eta\ge 0. \tag{3.18}$$

Proof. Note that

$$1+\sum_{j=1}^d |v_j|^2 \le \prod_{j=1}^d(1+|v_j|^2) \le (1+\sum_{j=1}^d |v_j|^2)^d.$$

Then we have

$$\begin{aligned} \|f\|_{L^2_{\boldsymbol{\mu}^{-2\eta}}(\mathbb{R}^d)} &= \left(\int_{\mathbb{R}^d}|f(\boldsymbol{v})|^2\boldsymbol{\mu}^{-2\eta}\,\mathrm{d}\boldsymbol{v}\right)^{1/2} \\ &= \left(\int_{\mathbb{R}^d}|f(\boldsymbol{v})|^2\prod_{j=1}^d(1+|v_j|^2)^\eta\,\mathrm{d}\boldsymbol{v}\right)^{1/2} \\ &\ge \left(\int_{\mathbb{R}^d}|f(\boldsymbol{v})|^2\left(1+\sum_{j=1}^d|v_j|^2\right)^\eta\,\mathrm{d}\boldsymbol{v}\right)^{1/2} = \|f\|_{\mathcal{L}^2_\eta(\mathbb{R}^d)}. \end{aligned}$$

Also,

$$\begin{aligned}\|f\|_{L^2_{\boldsymbol{\mu}^{-2\eta}}(\mathbb{R}^d)} &= \left(\int_{\mathbb{R}^d} |f(\boldsymbol{v})|^2 \prod_{j=1}^d (1+|v_j|^2)^\eta \,\mathrm{d}\boldsymbol{v}\right)^{1/2} \\ &\le \left(\int_{\mathbb{R}^d} |f(\boldsymbol{v})|^2 \left(1+\sum_{j=1}^d |v_j|^2\right)^{d\eta} \mathrm{d}\boldsymbol{v}\right)^{1/2} = \|f\|_{\mathcal{L}^2_{d\eta}(\mathbb{R}^d)}. \qquad \square\end{aligned}$$

LEMMA 3.5. *For any $\eta \ge 0$, there exist some $\epsilon > d-1$ and $C_\epsilon > 0$ such that*

$$\|f\|_{\mathcal{L}^1_\eta(\mathbb{R}^d)} \le C_\epsilon \|f\|_{\mathcal{L}^2_{\eta+\frac{1+\epsilon}{2}}(\mathbb{R}^d)}. \tag{3.19}$$

Proof. Note that

$$\begin{aligned}\|f\|^2_{\mathcal{L}^1_\eta(\mathbb{R}^d)} &= \left(\int_{\mathbb{R}^d} |f(\boldsymbol{v})|(1+|\boldsymbol{v}|^2)^{\frac{\eta}{2}} \,\mathrm{d}\boldsymbol{v}\right)^2 \\ &\le \int_{\mathbb{R}^d} |f(\boldsymbol{v})|^2 (1+|\boldsymbol{v}|^2)^{\eta+\frac{1+\epsilon}{2}} \,\mathrm{d}\boldsymbol{v} \int_{\mathbb{R}^d} (1+|\boldsymbol{v}|^2)^{-\frac{1+\epsilon}{2}} \,\mathrm{d}\boldsymbol{v} \\ &\le C_\epsilon \int_{\mathbb{R}^d} |f(\boldsymbol{v})|^2 (1+|\boldsymbol{v}|^2)^{\eta+\frac{1+\epsilon}{2}} \,\mathrm{d}\boldsymbol{v} \\ &= C_\epsilon \|f\|^2_{\mathcal{L}^2_{\eta+\frac{1+\epsilon}{2}}(\mathbb{R}^d)},\end{aligned}$$

where we used the Cauchy–Schwarz inequality. □

We are ready to present a consistency result.

THEOREM 3.6. *Let the collision kernel take the form* (3.15), *and then under the algebraic mapping* (2.11) *with $S=1$, we have*

$$\begin{aligned}&\left\|Q(f,f) - \pi_N^{4,d} Q(\pi_N^{4,d} f, \pi_N^{4,d} f)\right\|_{\mathcal{L}^2_3(\mathbb{R}^d)} \\ &\quad\le C_{d,\epsilon}(\mathcal{B}) N^{-m} \Bigg(|f|_{\mathbf{B}^m_{\lambda+6d+\frac{3+\epsilon}{2}}(\mathbb{R}^d)} \|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)} \\ &\qquad + |f|_{\mathbf{B}^m_{\lambda+3d+1}(\mathbb{R}^d)} \|f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)} + |Q(f,f)|_{\boldsymbol{B}^m_4(\mathbb{R}^d)} \Bigg),\end{aligned} \tag{3.20}$$

where m is a positive integer, d is the dimension, $\epsilon > d-1$ is a constant, and $C_{d,\epsilon}(\mathcal{B})$ is a constant depending only on the kernel $\mathcal{B}$, d, and ϵ.

Proof. By the triangle inequality

$$\begin{aligned}&\|Q(f,f) - \pi_N^{4,d} Q(\pi_N^{4,d} f, \pi_N^{4,d} f)\|_{\mathcal{L}^2_3(\mathbb{R}^d)} \\ &\le \|Q(f,f) - \pi_N^{4,d} Q(f,f)\|_{\mathcal{L}^2_3(\mathbb{R}^d)} + \|\pi_N^{4,d} Q(f,f) - \pi_N^{4,d} Q(\pi_N^{4,d} f, \pi_N^{4,d} f)\|_{\mathcal{L}^2_3(\mathbb{R}^d)}.\end{aligned}$$

For the first term, by Lemma 3.4 and Theorem 2.3, we have

$$\begin{aligned}\|Q(f,f) - \pi_N^{4,d} Q(f,f)\|_{\mathcal{L}^2_3(\mathbb{R}^d)} &\le \left\|Q(f,f) - \pi_N^{4,d} Q(f,f)\right\|_{L^2_{\boldsymbol{\mu}^{-6}}(\mathbb{R}^d)} \\ &\le C N^{-m} |Q(f,f)|_{\boldsymbol{B}^m_4(\mathbb{R}^d)}.\end{aligned}$$

For the second term, using again Lemma 3.4, Theorem 2.3 and Lemma 3.5, Theorem 3.3, we have

$$
\begin{aligned}
&\|\pi_N^{4,d}Q(f,f)-\pi_N^{4,d}Q(\pi_N^{4,d}f,\pi_N^{4,d}f)\|_{\mathcal{L}_3^2(\mathbb{R}^d)} \le \|\pi_N^{4,d}Q(f,f)-\pi_N^{4,d}Q(\pi_N^{4,d}f,\pi_N^{4,d}f)\|_{L^2_{\boldsymbol{\mu}^{-6}}(\mathbb{R}^d)}\\
&\le \|Q(f,f)-Q(\pi_N^{4,d}f,\pi_N^{4,d}f)\|_{L^2_{\boldsymbol{\mu}^{-6}}(\mathbb{R}^d)} \le \|Q(f,f)-Q(\pi_N^{4,d}f,\pi_N^{4,d}f)\|_{\mathcal{L}^2_{3d}(\mathbb{R}^d)}\\
&\le \|Q(f-\pi_N^{4,d}f,f)\|_{\mathcal{L}^2_{3d}(\mathbb{R}^d)} + \|Q(\pi_N^{4,d}f,f-\pi_N^{4,d}f)\|_{\mathcal{L}^2_{3d}(\mathbb{R}^d)}\\
&\le C_d(\mathcal{B})\left(\|f-\pi_N^{4,d}f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)} + \|\pi_N^{4,d}f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f-\pi_N^{4,d}f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}\right)\\
&\le C_d(\mathcal{B})\Big(\|f-\pi_N^{4,d}f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}(\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)} + \|f-\pi_N^{4,d}f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)})\\
&\quad + \|f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f-\pi_N^{4,d}f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}\Big)\\
&\le C_{d,\epsilon}(\mathcal{B})\|f-\pi_N^{4,d}f\|_{\mathcal{L}^2_{\lambda+6d+\frac{1+\epsilon}{2}}(\mathbb{R}^d)}\left(\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)} + \|f-\pi_N^{4,d}f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}\right)\\
&\quad + \ C_d(\mathcal{B})\|f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f-\pi_N^{4,d}f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)}\\
&\le C_{d,\epsilon}(\mathcal{B})\|f-\pi_N^{4,d}f\|_{L^2_{\boldsymbol{\mu}^{-2(\lambda+6d+\frac{1+\epsilon}{2})}}(\mathbb{R}^d)}(\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)} + \|f-\pi_N^{4,d}f\|_{L^2_{\boldsymbol{\mu}^{-2(\lambda+3d)}}(\mathbb{R}^d)})\\
&\quad + \ C_d(\mathcal{B})\|f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\|f-\pi_N^{4,d}f\|_{L^2_{\boldsymbol{\mu}^{-2(\lambda+3d)}}(\mathbb{R}^d)}\\
&\le C_{d,\epsilon}(\mathcal{B})N^{-m}\big|f\big|_{\mathbf{B}^m_{\lambda+6d+\frac{3+\epsilon}{2}}(\mathbb{R}^d)}\left(\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)} + CN^{-m}\big|f\big|_{\mathbf{B}^m_{\lambda+3d+1}(\mathbb{R}^d)}\right)\\
&\quad + \ C_d(\mathcal{B})N^{-m}\|f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\big|f\big|_{\mathbf{B}^m_{\lambda+3d+1}(\mathbb{R}^d)}\\
&\le C_{d,\epsilon}(\mathcal{B})N^{-m}\left(\big|f\big|_{\mathbf{B}^m_{\lambda+6d+\frac{3+\epsilon}{2}}(\mathbb{R}^d)}\|f\|_{\mathcal{L}^2_{\lambda+3d}(\mathbb{R}^d)} + \big|f\big|_{\mathbf{B}^m_{\lambda+3d+1}(\mathbb{R}^d)}\|f\|_{\mathcal{L}^1_{\lambda+6d}(\mathbb{R}^d)}\right).
\end{aligned}
$$

Combining the above inequalities, we arrive at the desired result. □

3.2. Approximation property for the moments. For the Boltzmann equation, the moments or macroscopic observables are important physical quantities. Still, under the algebraic mapping (2.11), we can show that the spectral method (3.5) preserves mass and energy.

THEOREM 3.7. *If using the algebraic mapping* (2.11) *with $N \ge 2$, the spectral method* (3.5) *preserves mass and energy, i.e., $\rho(t)$ and $E(t)$ defined by*

$$
\rho(t) := \int_{\mathbb{R}^d} f_N(t,\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}, \quad E(t) = \int_{\mathbb{R}^d} f_N(t,\boldsymbol{v})|\boldsymbol{v}|^2\,\mathrm{d}\boldsymbol{v} \tag{3.21}
$$

remain constant in time. Furthermore,

$$
\rho(t) \equiv \int_{\mathbb{R}^d} f^0(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}, \quad E(t) \equiv \int_{\mathbb{R}^d} f^0(\boldsymbol{v})|\boldsymbol{v}|^2\,\mathrm{d}\boldsymbol{v}. \tag{3.22}
$$

Proof. In one dimension, the first few Chebyshev polynomials read

$$
T_0(\xi) = 1, \quad T_1(\xi) = \xi, \quad T_2(\xi) = 2\xi^2 - 1.
$$

With the algebraic mapping (2.11), we have

$$
T_0(\xi(v)) = 1, \quad T_1(\xi(v)) = \frac{v}{\sqrt{v^2+S^2}}, \quad T_2(\xi(v)) = \frac{v^2-S^2}{v^2+S^2}, \quad \mu(\xi(v)) = \frac{\sqrt{S}}{\sqrt{v^2+S^2}}.
$$

Then

$$\widehat{T}_k(v) = \frac{[\mu(\xi(v))]^{-2}}{\sqrt{c_k}} T_k(\xi(v)) = \frac{v^2+S^2}{\sqrt{c_k}S} T_k(\xi(v)).$$

Specifically,

$$\widehat{T}_0(v) = \frac{v^2+S^2}{\sqrt{c_0}S}, \quad \widehat{T}_1(v) = \frac{v\sqrt{v^2+S^2}}{\sqrt{c_1}S}, \quad \widehat{T}_2(v) = \frac{v^2-S^2}{\sqrt{c_2}S}.$$

Therefore,

$$1 = \frac{\sqrt{c_0}}{2S}\widehat{T}_0(v) - \frac{\sqrt{c_2}}{2S}\widehat{T}_2(v), \quad v^2 = \frac{\sqrt{c_0}S}{2}\widehat{T}_0(v) + \frac{\sqrt{c_2}S}{2}\widehat{T}_2(v).$$

Hence we can replace $(\widehat{T}_0(v), \widehat{T}_2(v))$ by $(1, v^2)$ as basis functions, namely,

$$\widehat{\mathbb{T}}_N^1 = \operatorname{span}\{1, \widehat{T}_1, v^2, \widehat{T}_3, \widehat{T}_4, \cdots, \widehat{T}_N\}.$$

In d dimensions, it is easy to see

$$1,\ v_1^2,\ v_2^2,\ \ldots,\ v_d^2 \in \widehat{\mathbb{T}}_N^d \quad \text{for } N \geq 2.$$

In other words, we have shown that $1, |\boldsymbol{v}|^2 \in \widehat{\mathbb{T}}_N^d$ for $N \geq 2$.

On the other hand, by (3.21) and (3.4), we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) =& (\partial_t f_N(t,\boldsymbol{v}), 1)_{\mathbb{R}^d} = (Q(f_N,f_N),1)_{\mathbb{R}^d} = 0;\\ \frac{\mathrm{d}}{\mathrm{d}t}E(t) =& (\partial_t f_N(t,\boldsymbol{v}), |\boldsymbol{v}|^2)_{\mathbb{R}^d} = \big(Q(f_N,f_N),|\boldsymbol{v}|^2\big)_{\mathbb{R}^d} = 0,\end{aligned}$$

where in the last equality we used the conservation property (1.5) of the collision operator.

It remains to show

$$\int_{\mathbb{R}^d} f_N(0,\boldsymbol{v})\,\mathrm{d}\boldsymbol{v} = \int_{\mathbb{R}^d} f^0(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}, \quad \int_{\mathbb{R}^d} f_N(0,\boldsymbol{v})|\boldsymbol{v}|^2\,\mathrm{d}\boldsymbol{v} = \int_{\mathbb{R}^d} f^0(\boldsymbol{v})|\boldsymbol{v}|^2\,\mathrm{d}\boldsymbol{v}.$$

Noting that $f_N(0,\boldsymbol{v}) = \pi^{4,d}f^0$, it suffices to show

$$(\pi^{4,d}f^0 - f^0, 1)_{\mathbb{R}^d} = (\pi^{4,d}f^0 - f^0, |\boldsymbol{v}|^2)_{\mathbb{R}^d} = 0,$$

which is true by (2.25) (with $\alpha = 4$). □

4. Numerical realization.

To implement the proposed spectral method, one needs to solve the ODE system (3.5). For time discretization, one can just use the explicit Runge–Kutta methods. Hence, the key is the efficient evaluation of $\mathcal{Q}_{\boldsymbol{k}}^N$ as defined in (3.6).

In this section, we introduce two algorithms to compute $\mathcal{Q}_{\boldsymbol{k}}^N$. The first one is a direct algorithm that treats $\mathcal{Q}_{\boldsymbol{k}}^N$ as a matrix/tensor-vector multiplication. Since the weight matrix/tensor does not depend on the numerical solution f_N, it can be precomputed and stored for repeated use. This approach is simple but will soon meet a bottleneck when N increases since the memory requirement as well as the online computational cost can get extremely high. To alleviate this, we propose a fast algorithm, where the key idea is to recognize the gain term of the collision operator as a nonuniform discrete Fourier cosine transform to be accelerated by the NUFFT. Note that this is possible because we are using the mapped Chebyshev functions as a basis, which is related to the Fourier cosine series.

4.1. A direct algorithm. To derive the direct algorithm, we substitute (3.3) into (3.6) to obtain

$$\begin{aligned}\mathcal{Q}^N_{\boldsymbol{k}} &= \sum_{0\le \boldsymbol{i},\boldsymbol{j}\le N} \widetilde{f}_{\boldsymbol{i}}\widetilde{f}_{\boldsymbol{j}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma})\, \widetilde{\boldsymbol{T}}_{\boldsymbol{i}}(\boldsymbol{v})\widetilde{\boldsymbol{T}}_{\boldsymbol{j}}(\boldsymbol{v}_*)[\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}') - \widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})]\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_* \\ &= \sum_{0\le \boldsymbol{i},\boldsymbol{j}\le N} \widetilde{f}_{\boldsymbol{i}}\widetilde{f}_{\boldsymbol{j}} \left[\widetilde{I}_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) - \widetilde{I}_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})\right], \quad 0 \le \boldsymbol{k} \le N, \end{aligned} \tag{4.1}$$

where

$$\widetilde{I}_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma})\, \widetilde{\boldsymbol{T}}_{\boldsymbol{i}}(\boldsymbol{v})\widetilde{\boldsymbol{T}}_{\boldsymbol{j}}(\boldsymbol{v}_*)\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}')\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_*, \tag{4.2}$$

$$\widetilde{I}_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma})\, \widetilde{\boldsymbol{T}}_{\boldsymbol{i}}(\boldsymbol{v})\widetilde{\boldsymbol{T}}_{\boldsymbol{j}}(\boldsymbol{v}_*)\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_*. \tag{4.3}$$

Since the tensors $\widetilde{I}_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$ and $\widetilde{I}_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$ do not depend on coefficients $\{\widetilde{f}_{\boldsymbol{k}}\}_{0\le\boldsymbol{k}\le N}$, a straightforward way to evaluate $\mathcal{Q}^N_{\boldsymbol{k}}$ is to precompute $\widetilde{I}_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$ and $\widetilde{I}_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$, and then evaluate the sum in (4.1) directly in the online computation. This is what we refer to as the direct algorithm. We observe that this algorithm requires $\mathcal{O}(N^{3d})$ memory to store the tensors $\widetilde{I}_1$ and $\widetilde{I}_2$; and to evaluate (4.1), it requires $\mathcal{O}(N^{3d})$ operations. Both the memory requirement and online computational cost can be quite demanding, especially for $d=3$ and large N.

We give some details on how to approximate $\widetilde{I}_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$ and $\widetilde{I}_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$, though this step can be completed in advance and does not take the actual computational time. We first perform a change of variables $(\boldsymbol{v},\boldsymbol{v}_*) \to (\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta}))$ to transform the integrals of $(\boldsymbol{v},\boldsymbol{v_*}) \in \mathbb{R}^d\times\mathbb{R}^d$ into integrals of $(\boldsymbol{\xi},\boldsymbol{\eta}) \in I^d \times I^d$, using the mapping introduced in section 2.1:

$$\widetilde{I}_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) = \int_{I^d}\int_{I^d} G_{\boldsymbol{k}}(\boldsymbol{\xi},\boldsymbol{\eta})\frac{\boldsymbol{T}_{\boldsymbol{i}}(\boldsymbol{\xi})\boldsymbol{T}_{\boldsymbol{j}}(\boldsymbol{\eta})}{\sqrt{\boldsymbol{c}_{\boldsymbol{i}}\boldsymbol{c}_{\boldsymbol{j}}}}\boldsymbol{\omega}(\boldsymbol{\xi})\boldsymbol{\omega}(\boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\eta}, \tag{4.4}$$

$$\widetilde{I}_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) = \int_{I^d}\int_{I^d} L_{\boldsymbol{k}}(\boldsymbol{\xi},\boldsymbol{\eta})\frac{\boldsymbol{T}_{\boldsymbol{i}}(\boldsymbol{\xi})\boldsymbol{T}_{\boldsymbol{j}}(\boldsymbol{\eta})}{\sqrt{\boldsymbol{c}_{\boldsymbol{i}}\boldsymbol{c}_{\boldsymbol{j}}}}\boldsymbol{\omega}(\boldsymbol{\xi})\boldsymbol{\omega}(\boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\eta}, \tag{4.5}$$

where

$$G_{\boldsymbol{k}}(\boldsymbol{\xi},\boldsymbol{\eta}) := [\boldsymbol{\mu}(\boldsymbol{\xi})\boldsymbol{\mu}(\boldsymbol{\eta})]^2 \int_{S^{d-1}} \mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta}),\boldsymbol{\sigma})\frac{\boldsymbol{T}_{\boldsymbol{k}}\left(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\sigma})}\right)}{\sqrt{\boldsymbol{c}_k}[\boldsymbol{\mu}\left(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\sigma})}\right)]^2}\,\mathrm{d}\boldsymbol{\sigma}, \tag{4.6}$$

$$L_{\boldsymbol{k}}(\boldsymbol{\xi},\boldsymbol{\eta}) := \frac{\boldsymbol{T}_{\boldsymbol{k}}(\boldsymbol{\xi})\,[\boldsymbol{\mu}(\boldsymbol{\eta})]^2}{\sqrt{\boldsymbol{c}_{\boldsymbol{k}}}}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta}),\boldsymbol{\sigma})\,\mathrm{d}\boldsymbol{\sigma}. \tag{4.7}$$

Notice that in (4.6), $\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\sigma})} \in I^d$ is the value transformed from

$$\boldsymbol{v}' = \frac{1}{2}(\boldsymbol{v}(\boldsymbol{\xi}) + \boldsymbol{v}_*(\boldsymbol{\eta})) + \frac{1}{2}|\boldsymbol{v}(\boldsymbol{\xi}) - \boldsymbol{v}_*(\boldsymbol{\eta})|\boldsymbol{\sigma} \in \mathbb{R}^d$$

under the same mapping. To approximate the above integrals in $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, and $\boldsymbol{\sigma}$, we choose $M_{\boldsymbol{v}}$ Chebyshev–Gauss–Lobatto quadrature points in each dimension of I^d for both $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, and $M_{\boldsymbol{\sigma}}$ quadrature points on the unit sphere S^{d-1} (for $d=2$, this can be the uniform points in polar angle; for $d=3$, this can be the Lebedev quadrature [29]). Therefore, for each fixed index $\boldsymbol{k}$, (4.4) and (4.5) are forward Chebyshev transforms of the functions $G_{\boldsymbol{k}}(\boldsymbol{\xi},\boldsymbol{\eta})$ and $L_{\boldsymbol{k}}(\boldsymbol{\xi},\boldsymbol{\eta})$, respectively. Thus, they can be evaluated efficiently using the fast Chebyshev transform.

4.2. A fast algorithm. To introduce the fast algorithm, we take the original form (3.6) and split $\mathcal{Q}_{\boldsymbol{k}}^{N}$ into a gain term and a loss term as $\mathcal{Q}_{\boldsymbol{k}}^{N}=\mathcal{Q}_{\boldsymbol{k}}^{N,+}-\mathcal{Q}_{\boldsymbol{k}}^{N,-}$ (under the cut-off assumption of the collision kernel, i.e., $\mathcal{B}$ is integrable in the angular direction), where

$$\mathcal{Q}_{\boldsymbol{k}}^{N,+}=\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\int_{S^{d-1}}\mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma})f_N(\boldsymbol{v}_*)\,\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v}')\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_*\right)f_N(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}, \tag{4.8}$$

$$\mathcal{Q}_{\boldsymbol{k}}^{N,-}=\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\int_{S^{d-1}}\mathcal{B}(\boldsymbol{v},\boldsymbol{v}_*,\boldsymbol{\sigma})f_N(\boldsymbol{v}_*)\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_*\right)f_N(\boldsymbol{v})\,\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}. \tag{4.9}$$

We propose to evaluate $\mathcal{Q}_{\boldsymbol{k}}^{N,+}$ and $\mathcal{Q}_{\boldsymbol{k}}^{N,-}$ following the above expressions. To this end, given the coefficients $\{\widetilde{f}_{\boldsymbol{k}}\}_{0\leq\boldsymbol{k}\leq N}$ at each time step, we first reconstruct f_N as in (3.3) at $M_{\boldsymbol{v}}$ Chebyshev–Gauss–Lobatto quadrature points in each dimension of $\boldsymbol{v}$ (for an accurate approximation we choose $M_{\boldsymbol{v}}=N+2$). This can be achieved by the fast Chebyshev transform in $\mathcal{O}(M_{\boldsymbol{v}}^d\log M_{\boldsymbol{v}})$ operations.

To evaluate the gain term $\mathcal{Q}_{k}^{N,+}$, we change the integrals of $(\boldsymbol{v},\boldsymbol{v_*})\in\mathbb{R}^d\times\mathbb{R}^d$ into integrals of $(\boldsymbol{\xi},\boldsymbol{\eta})\in I^d\times I^d$ in (4.8) (similarly as in the previous subsection for $\widetilde{I}_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$):

(4.10)

$$\begin{aligned}\mathcal{Q}_{\boldsymbol{k}}^{N,+}&=\int_{I^d}\left(\int_{I^d}\int_{S^{d-1}}\mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta}),\boldsymbol{\sigma})f_N(\boldsymbol{v}_*(\boldsymbol{\eta}))\frac{\boldsymbol{T}_{\boldsymbol{k}}\left(\zeta_{(\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\sigma})}\right)}{\sqrt{\boldsymbol{c}_{\boldsymbol{k}}}\left[\boldsymbol{\mu}\left(\zeta_{(\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\sigma})}\right)\right]^2}\frac{\boldsymbol{\omega}(\boldsymbol{\eta})}{[\boldsymbol{\mu}(\boldsymbol{\eta})]^2}\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{\eta}\right)\\&\qquad\times\ f_N(\boldsymbol{v}(\boldsymbol{\xi}))\frac{\boldsymbol{\omega}(\boldsymbol{\xi})}{[\boldsymbol{\mu}(\boldsymbol{\xi})]^2}\,\mathrm{d}\boldsymbol{\xi}\\&=\int_{I^d}\left(\int_{S^{d-1}}F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\sigma}\right)f_N(\boldsymbol{v}(\boldsymbol{\xi}))\frac{\boldsymbol{\omega}(\boldsymbol{\xi})}{[\boldsymbol{\mu}(\boldsymbol{\xi})]^2}\,\mathrm{d}\boldsymbol{\xi},\end{aligned}$$

where

$$F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi}):=\int_{I^d}\mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta}),\boldsymbol{\sigma})f_N(\boldsymbol{v}_*(\boldsymbol{\eta}))\frac{\boldsymbol{T}_{\boldsymbol{k}}\left(\zeta_{(\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\sigma})}\right)}{\sqrt{\boldsymbol{c}_{\boldsymbol{k}}}\left[\boldsymbol{\mu}\left(\zeta_{(\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\sigma})}\right)\right]^2}\frac{\boldsymbol{\omega}(\boldsymbol{\eta})}{[\boldsymbol{\mu}(\boldsymbol{\eta})]^2}\,\mathrm{d}\boldsymbol{\eta}. \tag{4.11}$$

Supposing $M_{\boldsymbol{v}}$ quadrature points are used in each dimension of $\boldsymbol{v}$ and $\boldsymbol{v}_*$ and $M_{\boldsymbol{\sigma}}$ points are used on the sphere S^{d-1}, a direct evaluation of (4.11) would require $\mathcal{O}(M_{\boldsymbol{\sigma}}M_{\boldsymbol{v}}^{2d}N^d)$ operations. Given $F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi})$, a direct evaluation of (4.10) would take $\mathcal{O}(M_{\boldsymbol{\sigma}}M_{\boldsymbol{v}}^{d}N^d)$ operations. Therefore, the major bottleneck is to compute $F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi})$, which is prohibitively expensive without a fast algorithm. Our main idea is to recognize (4.11) as a nonuniform discrete Fourier cosine transform so it can be evaluated by the NUFFT. We will see that the total complexity to evaluate $F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi})$ can be brought down to $\mathcal{O}(M_{\boldsymbol{\sigma}}M_{\boldsymbol{v}}^{2d}|\log\epsilon|+M_{\boldsymbol{\sigma}}M_{\boldsymbol{v}}^{d}N^d\log N)$, where ϵ is the requested precision in the NUFFT algorithm.

Applying the Chebyshev–Gauss–Lobatto quadrature $(\boldsymbol{\eta_j},w_{\boldsymbol{j}})_{1\leq\boldsymbol{j}\leq M_{\boldsymbol{v}}}$, (4.11) becomes

$$\begin{aligned} F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi}) &\approx \sum_{1\le \boldsymbol{j}\le M_{\boldsymbol{v}}} w_{\boldsymbol{j}} \frac{\mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta_j}),\boldsymbol{\sigma}) f_N(\boldsymbol{v}_*(\boldsymbol{\eta_j}))}{\sqrt{c_{\boldsymbol{k}}}[\boldsymbol{\mu}(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma})})]^2[\boldsymbol{\mu}(\boldsymbol{\eta_j})]^2} \boldsymbol{T_k}(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma})}) \\ &= \sum_{1\le \boldsymbol{j}\le M_{\boldsymbol{v}}} w_{\boldsymbol{j}} \frac{\mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta_j}),\boldsymbol{\sigma}) f_N(\boldsymbol{v}_*(\boldsymbol{\eta_j}))}{\sqrt{c_{\boldsymbol{k}}}[\boldsymbol{\mu}(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma})})]^2[\boldsymbol{\mu}(\boldsymbol{\eta_j})]^2} \prod_{l=1}^{d} \cos\left(k_l \arccos\big(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma}),l}\big)\right) \\ &= \frac{1}{\sqrt{c_{\boldsymbol{k}}}} \sum_{1\le \boldsymbol{j}\le M_{\boldsymbol{v}}} q_{\boldsymbol{j}} \prod_{l=1}^{d} \cos\left(k_l \boldsymbol{z}_{\boldsymbol{j},l}\right), \end{aligned} \tag{4.12}$$

where $\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma}),l}$ is the lth component of $\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma})}$, and

$$q_{\boldsymbol{j}} := w_{\boldsymbol{j}} \frac{\mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta_j}),\boldsymbol{\sigma}) f_N(\boldsymbol{v}_*(\boldsymbol{\eta_j}))}{[\boldsymbol{\mu}(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma})})]^2[\boldsymbol{\mu}(\boldsymbol{\eta_j})]^2}, \quad \boldsymbol{z}_{\boldsymbol{j},l} := \arccos\big(\boldsymbol{\zeta}_{(\boldsymbol{\xi},\boldsymbol{\eta_j},\boldsymbol{\sigma}),l}\big). \tag{4.13}$$

(4.12) is almost a nonuniform discrete Fourier cosine transform. Indeed, we propose to evaluate $F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi})$ in two steps. First, for each fixed $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$, we compute

$$\widetilde{F}_{\boldsymbol{K}} := \sum_{1\le \boldsymbol{j}\le M_{\boldsymbol{v}}} q_{\boldsymbol{j}} \exp(\mathbf{i}\boldsymbol{K}\cdot \boldsymbol{z_j}), \quad -N \le \boldsymbol{K} \le N, \tag{4.14}$$

which is a nonuniform discrete Fourier transform mapping samples $\boldsymbol{z_j} \in [0,\pi]^d$ into frequencies $\boldsymbol{K} \in [-N,N]^d$. This can be done efficiently using the NUFFT algorithm.

In recent years, various NUFFT algorithms and applications have been well developed. Greengard and Lee [21, 30] describe an extremely simple and efficient implementation of NUFFT, which uses a Gaussian kernel in the algorithm. The general idea is to apply an interpolation between nonuniform samples and an equispaced grid, and then perform the uniform FFT on the new grid. In our numerical realization, we employ the recent FINUFFT library [3, 2], which uses a new "exponential of semicircle" kernel

$$\phi(z) = e^{\beta\sqrt{1-z^2}}.$$

This algorithm only costs $\mathcal{O}(M_{\boldsymbol{v}}^d|\log\epsilon| + N^d\log N)$ operations to compute (4.14) with the requested precision ϵ.

Once we obtain $\widetilde{F}_{\boldsymbol{K}}$, $F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi})$ can be retrieved as

- in two dimensions

$$F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi}) = \frac{1}{\sqrt{c_{\boldsymbol{k}}}}\frac{1}{2}\mathbf{Re}\left(\widetilde{F}_{(k_1,k_2)} + \widetilde{F}_{(-k_1,k_2)}\right); \tag{4.15}$$

- in three dimensions

$$F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi}) = \frac{1}{\sqrt{c_{\boldsymbol{k}}}}\frac{1}{4}\mathbf{Re}\left(\widetilde{F}_{(k_1,k_2,k_3)} + \widetilde{F}_{(-k_1,k_2,k_3)} + \widetilde{F}_{(k_1,-k_2,k_3)} + \widetilde{F}_{(k_1,k_2,-k_3)}\right). \tag{4.16}$$

This procedure needs to be repeated for every $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$; hence the overall computational cost for getting $F_{\boldsymbol{k}}(\boldsymbol{\sigma},\boldsymbol{\xi})$ is $\mathcal{O}(M_{\boldsymbol{\sigma}}M_{\boldsymbol{v}}^{2d}|\log\epsilon| + M_{\boldsymbol{\sigma}}M_{\boldsymbol{v}}^d N^d \log N)$.

To evaluate the loss term $\mathcal{Q}_k^{N,-}$, we change the integrals of $(\boldsymbol{v},\boldsymbol{v_*}) \in \mathbb{R}^d\times\mathbb{R}^d$ into integrals of $(\boldsymbol{\xi},\boldsymbol{\eta}) \in I^d \times I^d$ in (4.9) (similarly as in the previous subsection for $\widetilde{I}_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$):

$$\begin{aligned} \mathcal{Q}_{\boldsymbol{k}}^{N,-} = \int_{I^d} &\left(\int_{I^d}\int_{S^{d-1}} \mathcal{B}(\boldsymbol{v}(\boldsymbol{\xi}),\boldsymbol{v}_*(\boldsymbol{\eta}),\boldsymbol{\sigma}) f_N(\boldsymbol{v}_*(\boldsymbol{\eta})) \frac{\boldsymbol{\omega}(\boldsymbol{\eta})}{[\boldsymbol{\mu}(\boldsymbol{\eta})]^2}\, \mathrm{d}\boldsymbol{\sigma}\, \mathrm{d}\boldsymbol{\eta}\right) \\ &\times f_N(\boldsymbol{v}(\boldsymbol{\xi}))\, \frac{\boldsymbol{T_k}(\boldsymbol{\xi})\boldsymbol{\omega}(\boldsymbol{\xi})}{\sqrt{\boldsymbol{c_k}}\,[\boldsymbol{\mu}(\boldsymbol{\xi})]^4}\, \mathrm{d}\boldsymbol{\xi}. \end{aligned} \tag{4.17}$$

Table 1

Storage requirement and (online) computational cost of the direct and fast algorithms. N is the number of spectral modes in each dimension of $\boldsymbol{v}$; $M_{\boldsymbol{v}} = \mathcal{O}(N)$ is the number of quadrature points in each dimension; $M_{\boldsymbol{\sigma}} \ll N^d$ is the number of quadrature points on the sphere S^{d-1}; and ϵ is the requested precision in the NUFFT algorithm. The proposed fast algorithm does not require extra storage other than that storing the computational target, e.g., the gain and loss terms.

	Direct algorithm		Fast algorithm
	Storage	(Online) operation	(Online) operation
Gain term	$\mathcal{O}(N^{3d})$	$\mathcal{O}(N^{3d})$	$\mathcal{O}(M_{\boldsymbol{\sigma}} M_{\boldsymbol{v}}^{2d} \lvert\log \epsilon\rvert + M_{\boldsymbol{\sigma}} M_{\boldsymbol{v}}^d N^d \log N)$
Loss term	$\mathcal{O}(N^{3d})$	$\mathcal{O}(N^{3d})$	$\mathcal{O}(M_{\boldsymbol{\sigma}} M_{\boldsymbol{v}}^{2d})$

Then one can just evaluate the terms in the parentheses directly with complexity $\mathcal{O}(M_{\boldsymbol{v}}^{2d} M_{\boldsymbol{\sigma}})$. The outer integral in $\boldsymbol{\xi}$ can be viewed as the Chebyshev transform of some function and thus can be evaluated efficiently by the fast Chebyshev transform in $\mathcal{O}(M_{\boldsymbol{v}}^d \log M_{\boldsymbol{v}})$. In particular, if we consider the Maxwell kernel, i.e., $\mathcal{B}(\boldsymbol{v}, \boldsymbol{v}_*, \boldsymbol{\sigma}) \equiv$ constant, terms in the parentheses only require $\mathcal{O}(M_{\boldsymbol{v}}^d)$ complexity.

4.3. Comparison of direct and fast algorithms. To summarize, we list the storage requirement and (online) computational complexity for both the direct and fast algorithms in Table 1. Note that we only list the dominant complexity for each term. It is clear that the main cost of the fast algorithm comes from evaluating the gain term. Compared with the direct algorithm, the fast algorithm is generally faster as $M_{\boldsymbol{\sigma}}$ can be chosen much smaller than N^d in practice (see section 5). Most importantly, the fast algorithm does not require any precomputation with excessive storage requirement and everything can be computed on the fly.

5. Numerical examples.

5. Numerical examples. In this section, we perform extensive numerical tests to demonstrate the accuracy and efficiency of the proposed Petrov–Galerkin spectral method in both two and three dimensions.

Recall that the main motivation of the current work is to obtain better accuracy by considering approximations in an unbounded domain. To illustrate this point, we will compare three methods to solve the Boltzmann equation:

(1) *Fast Fourier–Galerkin spectral method proposed in* [16]: This method can achieve a good accuracy-efficiency tradeoff among the current deterministic methods for the Boltzmann equation. However, it requires the truncation of the domain to $[-L, L]^d$, where L is often chosen empirically such that the solution is close to zero at the boundary.

(2) *Fast Chebyshev-*0 *method:* The method proposed in the current paper using the logarithmic mapping (2.10), where $r = 0$ and the scaling parameter S needs to be properly chosen.

(3) *Fast Chebyshev-*1 *method:* The method proposed in the current paper using the algebraic mapping (2.11), where $r = 1$ and the scaling parameter S needs to be properly chosen.

In all three methods, the choice of truncation or mapping/scaling parameters has a great impact on numerical accuracy. In the following tests, we first determine L in the Fourier spectral method. Then for the two Chebyshev methods, we propose an adaptive strategy to determine the scaling parameter S: for example, in one dimension, the Chebyshev–Gauss–Lobatto quadrature points on the interval $[-1, 1]$ are given by

$$\xi_j = -\cos \frac{(j-1)\pi}{M-1}, \quad 1 \le j \le M,$$

such that $-1 = \xi_1 < \xi_2 < \cdots < \xi_M = 1$. We choose S such that the two quadrature points ξ_2 and ξ_{M-1} are mapped to the boundary of $[-L, L]$, i.e.,

$$v(\xi_1) = -\infty, \quad v(\xi_2) = -L, \quad v(\xi_{M-1}) = L, \quad v(\xi_M) = \infty.$$

More specifically, for given CGL quadrature, we let

- *logarithmic mapping* ($r = 0$):

$$S = \frac{2L}{\ln\left(\frac{1+\xi_{M-1}}{1-\xi_{M-1}}\right)}; \tag{5.1}$$

- *algebraic mapping* ($r = 1$):

$$S = \frac{L\sqrt{1-\xi_{M-1}^2}}{\xi_{M-1}}. \tag{5.2}$$

This S is adaptive in the sense that different M will correspond to different S.

5.1. 2D examples.

5.1.1. 2D BKW solution. We consider first the 2D (BKW) solution. This is one of the few known analytical solutions to the Boltzmann equation and a perfect example to verify the accuracy of a numerical method.

When $d = 2$ and the collision kernel $\mathcal{B} \equiv 1/(2\pi)$, the following is a solution to the initial value problem (3.1):

$$f_{\mathrm{BKW}}(t, \boldsymbol{v}) = \frac{1}{2\pi K^2} \exp\left(-\frac{\boldsymbol{v}^2}{2K}\right)\left(2K - 1 + \frac{1-K}{2K}\boldsymbol{v}^2\right), \tag{5.3}$$

where $K = 1 - \exp(-t/8)/2$. By taking the time derivative of $f_{\mathrm{BKW}}(t, \boldsymbol{v})$, we can obtain the exact collision operator as

$$Q_{\mathrm{BKW}}(f) = \left\{\left(-\frac{2}{K} + \frac{\boldsymbol{v}^2}{2K^2}\right) f_{\mathrm{BKW}} + \frac{1}{2\pi K^2} \exp\left(-\frac{\boldsymbol{v}^2}{2K}\right)\left(2 - \frac{1}{2K^2}\boldsymbol{v}^2\right)\right\} K', \tag{5.4}$$

where $K' = \exp(-t/8)/16$. This way we can apply the numerical method to compute $Q_{\mathrm{BKW}}(f)$ directly and check its error without worrying about the time discretization.

In the fast Fourier spectral method, we take $N_\rho = N$ quadrature points in the radial direction and $M_\sigma = 8$ quadrature points on the unit circle (see [16] for more details). In the fast Chebyshev methods, we take $M_{\boldsymbol{v}} = N + 2$ quadrature points for each dimension of $(\boldsymbol{v}, \boldsymbol{v}_*)$ and $M_{\boldsymbol{\sigma}} = N$ quadrature points on the unit circle. The precision in NUFFT is selected as $\epsilon = 1e - 14$. The numerical error of $Q_{\mathrm{BKW}}(f)$ is estimated on a 200×200 uniform grid in the rectangular domain $[-6.3, 6.3]^2$ at time $t = 2$.

Test 01. In this test, we examine thoroughly the numerical errors concerning different truncation parameters L in the Fourier method and scaling parameters S in the Chebyshev methods. Here we test different truncation domain $[-L, L]$ in the Fourier method, where $2.20 \le L \le 15.45$. We also test the fast Chebyshev methods with scaling parameter $1 \le S \le 9$. The L^2 errors of $Q_{\mathrm{BKW}}(f)$ for three methods are presented in Figure 1. Here the x-axes correspond to the parameters L and S. Each curve in these figures represents the approximation error of the corresponding method using N basis functions in the spectral method.

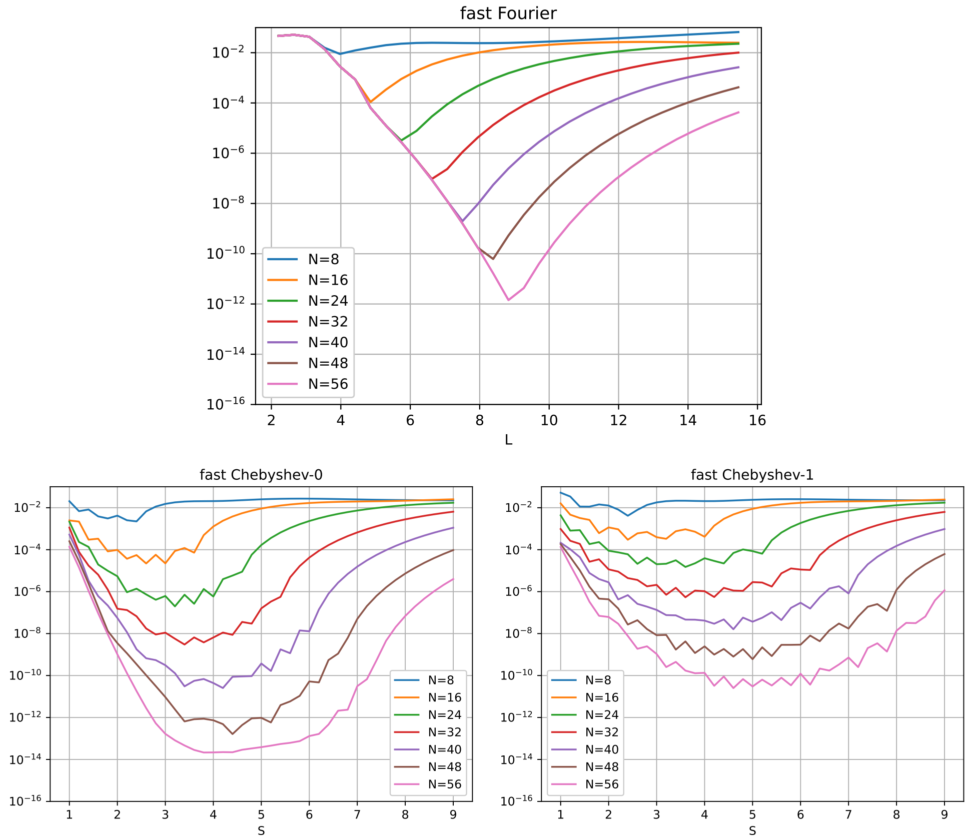

Fig. 1. *(2D BKW: Test 01) The L^2 error of $Q_{\text{BKW}}(f)$ at time $t = 2$. Top: fast Fourier method. Bottom: fast Chebyshev methods.*

Table 2

(2D BKW: Test 01) The best accuracy (L^2 error) of $Q_{\text{BKW}}(f)$ at time $t = 2$.

	Fast Fourier	Fast Chebyshev-0	Fast Chebyshev-1
$N = 8$	8.9223e-03	2.2289e-03	4.1388e-03
$N = 16$	1.0989e-04	2.2033e-05	2.9713e-04
$N = 24$	3.2104e-06	1.9843e-07	1.5004e-05
$N = 32$	9.4720e-08	3.0082e-09	5.3946e-07
$N = 40$	1.9836e-09	2.5434e-11	1.6120e-08
$N = 48$	6.1797e-11	1.6255e-13	6.0320e-10
$N = 56$	1.4315e-12	2.1482e-14	2.5213e-11

Clearly, the accuracy is not good when L and S are too small or too large. When L and S are chosen appropriately, the accuracy can be close to the machine precision. In Table 2, we record the best accuracy for a given N of each method, which is the smallest error of each curve in Figure 1. One can see that the fast Chebyshev-0 method can always achieve the best accuracy.

In practice, it would be difficult to choose the optimal L or S since the exact solution is not available. In most cases, the truncation domain L in the Fourier method is selected roughly so that it covers twice of the compact support of f. S in the Chebyshev method is then chosen adaptively as described before.

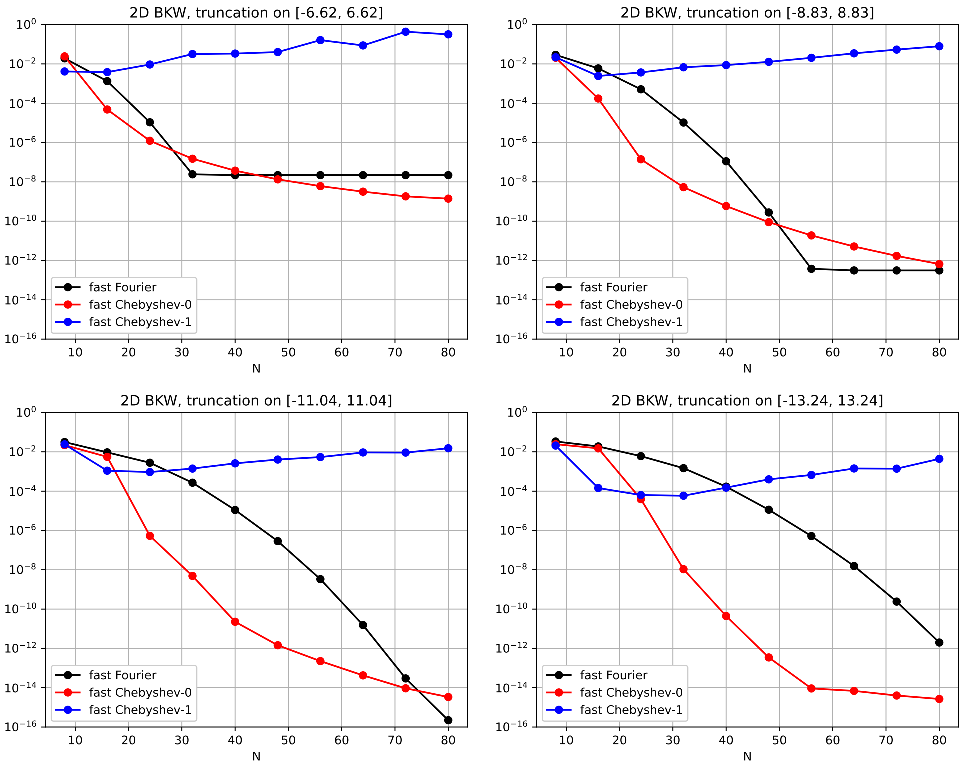

FIG. 2. *(2D BKW: Test* 02*) The L^∞ error of $Q_{\mathrm{BKW}}(f)$ at time $t = 2$ with different truncation domain.*

Test 02. In this test, we fix the computational domain and examine the numerical errors concerning different N. In the Fourier method, we test four different truncation parameters: $L = 6.62, 8.83, 11.04,$ and 13.24. In the Chebyshev methods, we use the same L to select the scaling parameter S accordingly. The L^∞ errors of $Q_{\mathrm{BKW}}(f)$ for three methods are presented in Figure 2. Among these three methods, the fast Chebyshev-0 method can achieve the best accuracy when N is small. The fast Chebyshev-1 method doesn't provide a good approximation. The selected quadrature points and function spaces might explain the poor performance of the fast Chebyshev-1 method:

1. The quadrature points in the Chebyshev-1 method are much more clustered near the origin compared to the Chebyshev-0 method. The quadrature points located far away from the origin also play an important role in the unbounded domain problem.
2. As described in Lemma 2.2, the trial functions in the Chebyshev-0 method decay exponentially as $|\boldsymbol{v}| \to \infty$, which mimics better the decay property of the BKW solution at infinity.

Test 03. In this test, we examine the numerical errors of the Chebyshev methods with a fixed scaling parameter: $S = 4$ in the fast Chebyshev-0 method; $S = 5$ in the fast Chebyshev-1 method. These two values are selected based on observation of the results in Figure 1. The L^∞ errors of $Q_{\mathrm{BKW}}(f)$ for both methods are presented in Figure 3. As a comparison, results of the fast Fourier method are also plotted

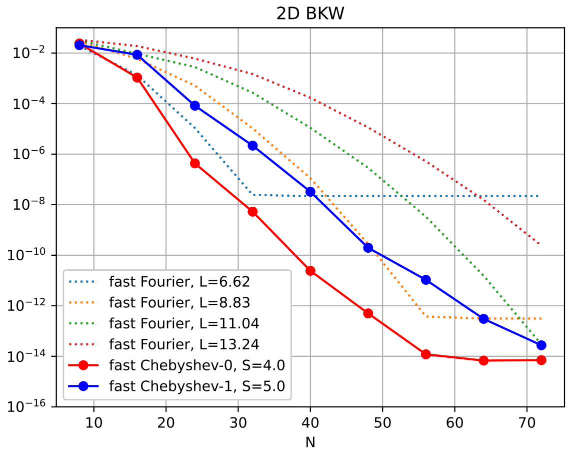

FIG. 3. *(2D BKW: Test 03) The L^∞ error of $Q_{\text{BKW}}(f)$ at time $t = 2$.*

TABLE 3
(2D BKW: Test 04) Running time in seconds for a single evaluation of the gain term.

	Direct Chebyshev algorithm		Fast Chebyshev algorithm
	Online (sec)	Precomputation (sec)	Online (sec)
$N = 8$	0.0047	1.5956	0.194955
$N = 16$	0.1207	62.7991	2.036821
$N = 32$	-	-	24.779492
$N = 64$	-	-	4.937722e+02
$N = 128$	-	-	1.331576e+04

with different L. In comparison to Test 02, it is easy to see the improvement of the Chebyshev-1 method by using different S. But for small N, the Chebyshev-0 method still has the best accuracy among all three methods.

Test 04. In this test, we report the computational time of the direct (Chebyshev) algorithm and the fast (Chebyshev) algorithm. The computations were done on an Intel Core i7-6700 CPU in a single thread. Table 3 shows the running time of the direct and fast algorithms concerning different N. Note that the direct algorithm is left out when $N \geq 32$ due to the memory constraint.

Test 05. We also compare the accuracy and computational time between the fast Fourier method and the fast Chebyshev-0 method. In the fast Fourier method, we choose $L = 13.24$ as the domain truncation, and $N_\rho = N$ quadrature points in the radial direction. In the fast Chebyshev-0 method, we take $M_{\boldsymbol{v}} = N + 2$ quadrature points for each dimension of $(\boldsymbol{v}, \boldsymbol{v}_*)$. The scaling parameter S is adaptively chosen as in (5.1). In the fast Fourier method, $M_{\boldsymbol{\sigma}} = 8$ is enough (we have tested that a larger value of $M_{\boldsymbol{\sigma}}$ would not further increase the accuracy) for the integral on the unit circle S^1. In the fast Chebyshev-0 method, more quadrature points are needed, but it is always less than N (again we have tested that a larger value of $M_{\boldsymbol{\sigma}}$ would not further increase the accuracy). One will also see a similar phenomenon in a 3D test. As shown in Table 4, for the same N, although the Fourier method can be much faster than the Chebyshev method (because it takes advantage of the convolutional structure of the collision operator in the Fourier domain, which is absent in any other spectral method) the accuracy it can achieve is much lower than the Chebyshev method mainly due to the domain truncation error.

TABLE 4
(2D BKW: Test 05) The L^∞ error of $Q_{BKW}(f)$ at time $t = 2$ and running time for a single evaluation of the gain term.

	Fast Fourier		Fast Chebyshev-0	
	Error	Online (sec)	Error	Online (sec)
$N = 8$	3.36e-02 ($M_{\boldsymbol{\sigma}} = 8$)	0.00288	2.45e-02 ($M_{\boldsymbol{\sigma}} = 6$)	0.316
$N = 16$	1.88e-02 ($M_{\boldsymbol{\sigma}} = 8$)	0.00824	1.51e-02 ($M_{\boldsymbol{\sigma}} = 6$)	1.1
$N = 24$	6.05e-03 ($M_{\boldsymbol{\sigma}} = 8$)	0.0352	4.27e-05 ($M_{\boldsymbol{\sigma}} = 12$)	5.35
$N = 32$	1.48e-03 ($M_{\boldsymbol{\sigma}} = 8$)	0.0946	1.07e-08 ($M_{\boldsymbol{\sigma}} = 24$)	23.9
$N = 40$	1.70e-04 ($M_{\boldsymbol{\sigma}} = 8$)	0.165	4.44e-11 ($M_{\boldsymbol{\sigma}} = 30$)	61.4
$N = 48$	1.14e-05 ($M_{\boldsymbol{\sigma}} = 8$)	0.263	3.47e-13 ($M_{\boldsymbol{\sigma}} = 36$)	108
$N = 56$	5.19e-07 ($M_{\boldsymbol{\sigma}} = 8$)	0.418	1.71e-14 ($M_{\boldsymbol{\sigma}} = 40$)	232

5.1.2. Computing the moments.

We next consider the time evolution problem and check the accuracy for moments approximation. Since the fast Chebyshev-0 method performs generally better than the fast Chebyshev-1 method, we will restrict it to the former in the following tests. The comparison with the fast Fourier method will still be considered.

In (3.1), we choose the collision kernel $\mathcal{B} \equiv 1/(2\pi)$ and the initial condition as

$$f^0(\boldsymbol{v}) = \frac{\rho_1}{2\pi T_1}\exp\left(-\frac{(\boldsymbol{v}-V_1)^2}{2T_1}\right) + \frac{\rho_2}{2\pi T_2}\exp\left(-\frac{(\boldsymbol{v}-V_2)^2}{2T_2}\right), \tag{5.5}$$

where $\rho_1 = \rho_2 = 1/2$, $T_1 = T_2 = 1$ and $V_1 = (x_1, y_1) = (-1, 2)$, $V_2 = (x_2, y_2) = (3, -3)$.

Then for the momentum flow and energy flow defined as

$$P_{ij} = \int_{\mathbb{R}^2} f v_i v_j \,\mathrm{d}\boldsymbol{v}, \quad (i, j = 1, 2), \quad q_i = \int_{\mathbb{R}^2} f v_i |\boldsymbol{v}|^2 \,\mathrm{d}\boldsymbol{v}, \quad (i = 1, 2), \tag{5.6}$$

we have their exact formulas

$$P_{11} = -\frac{9}{8}e^{-t/2} + \frac{57}{8}, \quad P_{12} = P_{21} = -5e^{-t/2} - \frac{1}{2}, \quad P_{22} = \frac{9}{8}e^{-t/2} + \frac{51}{8} \tag{5.7}$$

and

$$q_1 = \frac{1}{4}\left(11e^{-t/2} + 103\right), \quad q_2 = -\frac{1}{8}\left(89e^{-t/2} + 103\right). \tag{5.8}$$

For interested readers, the computation of these formulas is provided in the appendix.

In the fast Fourier method, we take $N_\rho = N$ quadrature points in the radial direction and $N_\theta = N$ quadrature points on the unit circle. The truncation domain $[-L, L]^2$ is selected as $L = 14.35$. In the fast Chebyshev-0 method, we take $M_{\boldsymbol{v}} = N+2$ quadrature points for each dimension of $(\boldsymbol{v}, \boldsymbol{v}_*)$ and $M_{\boldsymbol{\sigma}} = N$ quadrature points on the unit circle. The precision in NUFFT is selected as $\epsilon = 1e - 14$. The scaling parameter S is adaptively chosen based on L. For both methods, we use the fourth-order Runge–Kutta method with $\Delta t = 0.02$ for time integration.

The absolute errors of the moments are presented in Figures 4–8. The fast Chebyshev-0 method provides a better approximation in comparison to the Fourier method for fixed N.

5.2. 3D BKW solution.

We finally consider the 3D BKW solution. When $d = 3$ and the collision kernel $\mathcal{B} \equiv 1/(4\pi)$, the following is a solution to the initial value problem (3.1):

$$f_{\mathrm{BKW}}(t, \boldsymbol{v}) = \frac{1}{2(2\pi K)^{3/2}}\exp\left(-\frac{\boldsymbol{v}^2}{2K}\right)\left(\frac{5K-3}{K} + \frac{1-K}{K^2}\boldsymbol{v}^2\right), \tag{5.9}$$

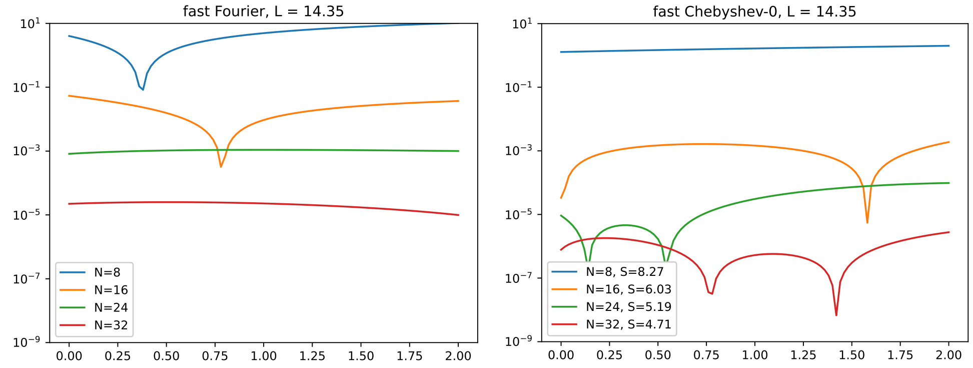

FIG. 4. *(2D moments) The time evolution for the absolute error of the momentum P_{11}. Left: the fast Fourier method. Right: the fast Chebyshev-0 method.*

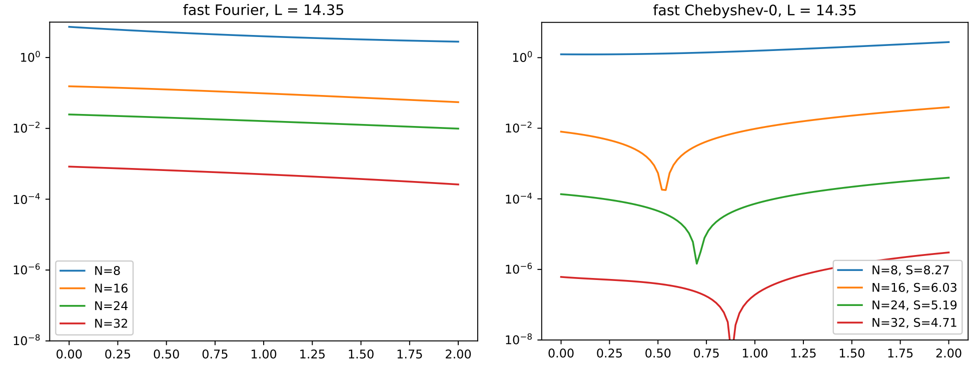

FIG. 5. *(2D moments) The time evolution for the absolute error of the momentum P_{12}. Left: the fast Fourier method. Right: the fast Chebyshev-0 method.*

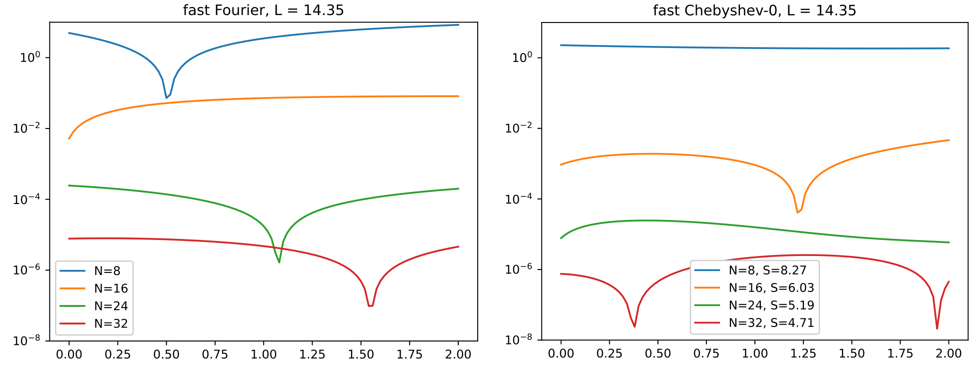

FIG. 6. *(2D moments) The time evolution for the absolute error of the momentum P_{22}. Left: the fast Fourier method. Right: the fast Chebyshev-0 method.*

where $K = 1 - \exp(-t/6)$. As in two dimensions, we can obtain the exact collision operator as

$$Q_{\mathrm{BKW}}(f) = \left\{ \left(-\frac{3}{2K} + \frac{\boldsymbol{v}^2}{2K^2} \right) f_{\mathrm{BKW}} + \frac{1}{2(2\pi K)^{3/2}} \exp\left(-\frac{\boldsymbol{v}^2}{2K} \right) \left(\frac{3}{K^2} + \frac{K-2}{K^3} \boldsymbol{v}^2 \right) \right\} K' \tag{5.10}$$

with $K' = \exp(-t/6)/6$.

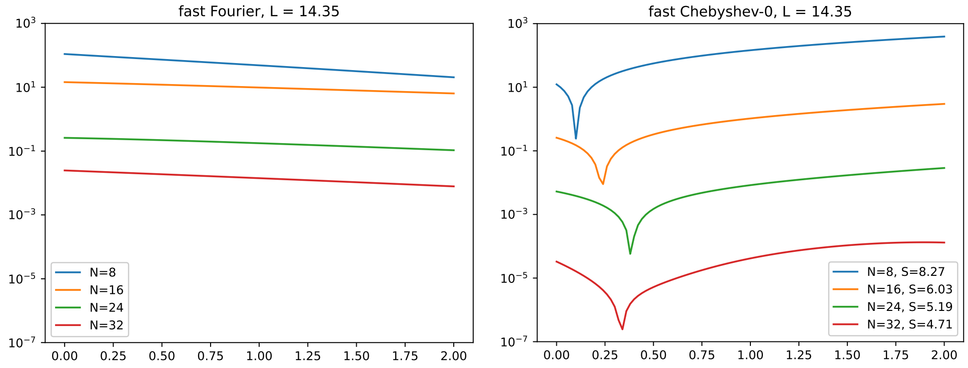

Fig. 7. *(2D moments) The time evolution for the absolute error of the momentum q_1. Left: the fast Fourier method. Right: the fast Chebyshev-0 method.*

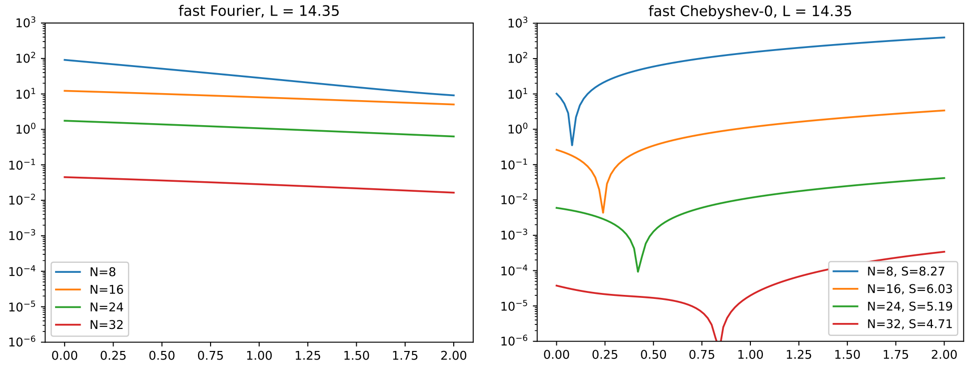

Fig. 8. *(2D moments) The time evolution for the absolute error of the momentum q_2. Left: the fast Fourier method. Right: the fast Chebyshev-0 method.*

Table 5
(3D BKW) The L^∞ error of $Q_{BKW}(f)$ at time $t = 6.5$.

	Fast Fourier ($M_{\boldsymbol{\sigma}} = 38$)	Fast Chebyshev-0
$N = 12$	2.36e-03	1.61e-02 ($M_{\boldsymbol{\sigma}} = 14$)
$N = 16$	4.37e-04	2.72e-03 ($M_{\boldsymbol{\sigma}} = 38$)
$N = 20$	3.62e-05	3.08e-06 ($M_{\boldsymbol{\sigma}} = 86$)
$N = 24$	3.61e-06	3.10e-08 ($M_{\boldsymbol{\sigma}} = 146$)
$N = 28$	1.64e-07	1.58e-08 ($M_{\boldsymbol{\sigma}} = 170$)
$N = 32$	3.82e-08	7.16e-10 ($M_{\boldsymbol{\sigma}} = 230$)

Here we compare the fast Fourier spectral method with the fast Chebyshev-0 method. In the former, we take domain $L = 6.62$, $N_\rho = N$ quadrature points in the radial direction and $M_{\boldsymbol{\sigma}} = 38$ Lebedev quadrature points on the unit sphere. In the latter, we choose S adaptively based on L, $M_{\boldsymbol{v}} = N + 2$ quadrature points for each dimension of $(\boldsymbol{v}, \boldsymbol{v}_*)$ and $M_{\boldsymbol{\sigma}}$ Lebedev quadrature points on the unit sphere. The precision in NUFFT is selected as $\epsilon = 1e - 14$. The L^∞ error of $Q_{\mathrm{BKW}}(f)$ is estimated on a $30 \times 30 \times 30$ uniform grid in the rectangular domain $[-6.3, 6.3]^3$ at time $t = 6.5$.

The results are reported in Table 5. Unlike the Fourier method for which $M_{\boldsymbol{\sigma}} = 38$ is enough (a larger value of $M_{\boldsymbol{\sigma}}$ would not further increase the accuracy), we observe that more quadrature points on the sphere are needed to get the best accuracy in the

Chebyshev method. As soon as $N \geq 20$, the Chebyshev method can always obtain better accuracy than the Fourier method.

6. Conclusion and future work.

We introduced a Petrov–Galerkin spectral method for the spatially homogeneous Boltzmann equation in multidimensions. The mapped Chebyshev functions in $\mathbb{R}^d$ were carefully chosen to serve as the trial functions and test functions in the approximation. In the case of the algebraic mapping, we established a consistency result for approximation of the collision operator as well as the conservation property for the moments. Thanks to the close relationship between the Chebyshev functions and the Fourier cosine series, we proposed a fast algorithm to alleviate the memory constraint in the precomputation and accelerate the online computation in the direct implementation. Through a series of numerical examples in two and three dimensions, we demonstrated that the proposed method can provide better accuracy (at least one or two digits for small N) in comparison to the popular Fourier spectral method.

Finally, we mention that although our main focus in this paper is on the spatially homogeneous equation (1.1), the proposed method can be easily extended to the spatially inhomogeneous problem following a pseudospectral approach. Simply speaking, one just needs to transfer between spectral coefficients $\widetilde{f}_{\boldsymbol{k}}$ (or $\mathcal{Q}^N_{\boldsymbol{k}}$) and function point values f_N (or $\pi_N^{4,d} Q(f_N, f_N)$) at every spatial point, similarly as done in the Fourier spectral method (e.g., [27]). In other words, the proposed method is only used as a black box solver to evaluate the collision operator (take f_N at quadrature points as input and produce $\pi_N^{4,d} Q(f_N, f_N)$ at the same points as output) and can be coupled with other spatial discretizations to solve the full Boltzmann equation. Of course the computational cost would be a major concern even equipped with the fast algorithm. We leave it for future work.

Appendix A. Proof of Lemma 2.2.

Proof. By Lemma 3.1 in [24], we know that in a 1D case,

$$\lim_{|v|\to\infty} |\widetilde{T}_k(v)| \sim \lim_{|v|\to\infty} [\mu(\xi(v))]^4, \tag{A.1}$$

$$\lim_{|v|\to\infty} |\widehat{T}_k(v)| \sim \lim_{|v|\to\infty} [\mu(\xi(v))]^{-2} \tag{A.2}$$

and

$$r = 0, \quad \mu = \frac{1}{\sqrt{S}}(1-\xi(v)^2)^{1/4} = \sqrt{\frac{\operatorname{sech}(v/S)}{S}} \sim e^{-\frac{|v|}{2S}}, \tag{A.3}$$

$$r = 1, \quad \mu = \frac{1}{\sqrt{S}}(1-\xi(v)^2)^{1/2} = \sqrt{\frac{S}{S^2+v^2}} \sim |v|^{-1}. \tag{A.4}$$

Consider the multidimensional case with logarithmic mapping ($r = 0$), and we get

$$\left|\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\right| \sim \prod_{j=1}^d e^{-\frac{2|v_j|}{S}} = e^{-\frac{2}{S}\sum_{j=1}^d |v_j|}, \quad \left|\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\right| \sim \prod_{j=1}^d e^{\frac{|v_j|}{S}} = e^{\frac{1}{S}\sum_{j=1}^d |v_j|}. \tag{A.5}$$

Similarly, for algebraic mapping ($r = 1$), we have

$$\left|\widetilde{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\right| \sim \prod_{j=1}^d |v_j|^{-4}, \quad \left|\widehat{\boldsymbol{T}}_{\boldsymbol{k}}(\boldsymbol{v})\right| \sim \prod_{j=1}^d |v_j|^{2}. \tag{A.6}$$

□

Appendix B. The exact computation of momentum flow $\boldsymbol{P_{ij}}$. The macroscopic quantities of initial condition (5.5) can be computed directly:

$$\rho = 1, \quad m = (1, -\frac{1}{2}), \quad u = (1, -\frac{1}{2}). \tag{B.1}$$

The initial values of the momentum flow are

$$P_{11}(0) = \rho_1(x_1^2 + T_1) + \rho_2(x_2^2 + T_2) = 6, \tag{B.2}$$

$$P_{12}(0) = \rho_1 x_1 y_1 + \rho_2 x_2 y_2 = -\frac{11}{2}, \tag{B.3}$$

$$P_{22}(0) = \rho_1(y_1^2 + T_1) + \rho_2(y_2^2 + T_2) = \frac{15}{2}. \tag{B.4}$$

Consider the time derivative of P_{11}:

$$\begin{aligned}
\frac{\partial P_{11}}{\partial t} &= \int_{\mathbb{R}^2} f_t v_1^2 \mathrm{d}\boldsymbol{v} \\
&= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{S^1} \frac{1}{2\pi} f(\boldsymbol{v}) f(\boldsymbol{v}_*)[(v_1')^2 - (v_1)^2]\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_* \\
&= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f(\boldsymbol{v}) f(\boldsymbol{v}_*) \frac{[3v_1 + (v_*)_1][(v_*)_1 - v_1]}{4}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_* \\
&\quad + \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{S^1} \frac{1}{2\pi} f(\boldsymbol{v}) f(\boldsymbol{v}_*) \frac{|\boldsymbol{v} - \boldsymbol{v}_*|^2}{4} \cos^2\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}_* \\
&= \frac{1}{2}(m_1 - \rho P_{11}) + \frac{\rho(P_{11} + P_{22}) - (m_1^2 + m_2^2)}{4} \\
&= \frac{m_1^2 - m_2^2 + \rho(P_{22} - P_{11})}{4} = \frac{3}{16} + \frac{P_{22} - P_{11}}{4}.
\end{aligned} \tag{B.5}$$

Similarly, one can get

$$\frac{\partial P_{12}}{\partial t} = \frac{m_1 m_2 - \rho P_{12}}{2} = -\frac{1}{4} - \frac{P_{12}}{2}, \tag{B.6}$$

$$\frac{\partial P_{22}}{\partial t} = \frac{m_2^2 - m_1^2 + \rho(P_{11} - P_{22})}{4} = -\frac{3}{16} + \frac{P_{11} - P_{22}}{4}. \tag{B.7}$$

Combining with the initial conditions, we get the exact formulas as (5.7).

REFERENCES

[1] R. Alonso, I. Gamba, and S. Tharkabhushanam, *Convergence and error estimates for the Lagrangian-based conservative spectral method for Boltzmann equations*, SIAM J. Numer. Anal., 56 (2018), pp. 3534–3579.

[2] A. H. Barnett, *Aliasing error of the* $\exp\left(\beta\sqrt{1-z^2}\right)$ *kernel in the nonuniform fast Fourier transform*, Appl. Comput. Harmon. Anal., 51 (2021), pp. 1–16.

[3] A. H. Barnett, J. Magland, and L. af Klinteberg, *A parallel nonuniform fast Fourier transform library based on an "exponential of semicircle" kernel*, SIAM J. Sci. Comput., 41 (2019), pp. C479–C504.

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

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

[6] C. K. Birdsall and A. B. Langdon, *Plama Physics via Computer Simulation*, CRC Press, Boca Raton, FL, 2018.

[7] A. Bobylev and S. Rjasanow, *Fast deterministic method of solving the Boltzmann equation for hard spheres*, Eur. J. Mech. B Fluids, 18 (1999), pp. 869–887.

[8] A. V. Bobylev, *Fourier transform method in the theory of the Boltzmann equation for Maxwellian molecules*, Soviet Phys. Dokl., 20 (1976), pp. 820–822.

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

[10] C. Cercignani, *Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations*, Cambridge University Press, Cambridge, 2000.

[11] S. Chandrasekhar, *Radiative Transfer*, Dover, New York, 1960.

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

[13] F. Filbet and C. Mouhot, *Analysis of spectral methods for the homogeneous Boltzmann equation*, Trans. Amer. Math. Soc., 363 (2011), pp. 1947–1980.

[14] F. Filbet and G. Russo, *High order numerical methods for the space non-homogeneous Boltzmann equation*, J. Comput. Phys., 186 (2003), pp. 457–480.

[15] E. Fonn, P. Grohs, and R. Hiptmair, *Polar Spectral Scheme for the Spatially Homogeneous Boltzmann Equation*, Research report, ETH Zurich, 2014.

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

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

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

[19] I. M. Gamba and S. H. Tharkabhushanam, *Shock and boundary structure formation by spectral-Lagrangian methods for the inhomogeneous Boltzmann transport equation*, J. Comput. Math., 28 (2010), pp. 430–460.

[20] V. Giovangigli, *Multicomponent Flow Modeling*, Springer, New York, 1999.

[21] L. Greengard and J.-Y. Lee, *Accelerating the nonuniform fast Fourier transform*, SIAM Revi., 46 (2004), pp. 443–454.

[22] P. Grohs, R. Hiptmair, and S. Pintarelli, *Tensor-product discretization for the spatially inhomogeneous and transient Boltzmann equation in two dimensions*, SMAI J. Comput. Math., 3 (2017), pp. 219–248.

[23] J. Hu, K. Qi, and T. Yang, *A new stability and convergence proof of the Fourier-Galerkin spectral method for the spatially homogeneous Boltzmann equation*, SIAM J. Numer. Anal., 59 (2021), pp. 613–633.

[24] J. Hu, J. Shen, and Y. Wang, *A Petrov-Galerkin spectral method for the inelastic Boltzmann equation using mapped Chebyshev functions*, Kinet. Relat. Models, 13 (2020).

[25] Z. Hu and Z. Cai, *Burnett spectral method for high-speed rarefied gas flows*, SIAM J. Sci. Comput., 42 (2020), pp. B1193–B1226.

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

[27] S. Jaiswal, A. Alexeenko, and J. Hu, *A discontinuous Galerkin fast spectral method for the full Boltzmann equation with general collision kernels*, J. Comput. Phys., 378 (2019), pp. 178–208.

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

[29] V. Lebedev, *Quadratures on a sphere*, Comput. Math. Math. Phys., 16 (1976), pp. 10–24.

[30] J.-Y. Lee and L. Greengard, *The type* 3 *nonuniform FFT and its applications*, J. Comput. Phys., 206 (2005), pp. 1–5.

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

[32] C. Mouhot and C. Villani, *Regularity theory for the spatially homogeneous Boltzmann equation with cut-off*, Arch. Ration. Mech. Anal., 173 (2004), pp. 169–212.

[33] G. Naldi, L. Pareschi, and G. Toscani, eds., *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Birkhäuser, Basel, 2010.

[34] K. Nanbu, *Direct simulation scheme derived from the Boltzmann equation.* I. *Monocomponent gases*, J. Phys. Soc. Jpn., 49 (1980), pp. 2042–2049.

[35] L. Pareschi and B. Perthame, *A Fourier spectral method for homogeneous Boltzmann equations*, Transp. Theory Statist. Phys., 25 (1996), pp. 369–382.

[36] L. Pareschi and G. Russo, *Numerical solution of the Boltzmann equation* I: *Spectrally accurate approximation of the collision operator*, SIAM J. Numer. Anal., 37 (2000), pp. 1217–1245.

[37] L. Pareschi and G. Russo, *On the stability of spectral methods for the homogeneous Boltzmann equation*, Transp. Theory Statist. Phys., 29 (2000), pp. 431–447.

[38] J. Shen, T. Tang, and L.-L. Wang, *Spectral Methods: Algorithms, Analysis and Applications*, Springer Ser. Comput. Math. 41, Springer, New York, 2011.

[39] J. Shen and L.-L. Wang, *Sparse spectral approximations of high-dimensional problems based on hyperbolic cross*, SIAM J. Numer. Anal., 48 (2010), pp. 1087–1109.

[40] J. Shen, L.-L. Wang, and H. Yu, *Approximations by orthonormal mapped Chebyshev functions for higher-dimensional problems in unbounded domains*, J. Comput. Appl. Math., 265 (2014), pp. 264–275.

[41] M.-B. Tran, *Nonlinear Approximation Theory for the Homogeneous Boltzmann Equation*, preprint, arXiv:1305.1667, 2013.

[42] C. Villani, *A review of mathematical topics in collisional kinetic theory*, in Handbook of Mathematical Fluid Mechanics, Vol. I, S. Friedlander and D. Serre, eds., North-Holland, Amsterdam, 2002, pp. 71–305.