NUMERICAL MATHEMATICS: Theory, Methods and Applications
Numer. Math. Theor. Meth. Appl., Vol. **xx**, No. x, pp. 1-23 (2017)

A Stochastic Galerkin Method for the Boltzmann Equation with Multi-dimensional Random Inputs Using Sparse Wavelet Bases

Ruiwen Shu[1], Jingwei Hu[2] and Shi Jin[1,3,*]

[1] *Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA.*
[2] *Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA.*
[3] *Institute of Natural Sciences, School of Mathematical Science, MOELSEC and SHL-MAC, Shanghai Jiao Tong University, Shanghai 200240, China.*

In celebration of the eightieth birthday of Prof. Zhen-huan Teng

Abstract. We propose a stochastic Galerkin method using sparse wavelet bases for the Boltzmann equation with multi-dimensional random inputs. The method uses locally supported piecewise polynomials as an orthonormal basis of the random space. By a sparse approach, only a moderate number of basis functions is required to achieve good accuracy in multi-dimensional random spaces. We discover a sparse structure of a set of basis-related coefficients, which allows us to accelerate the computation of the collision operator. Regularity of the solution of the Boltzmann equation in the random space and an accuracy result of the stochastic Galerkin method are proved in multi-dimensional cases. The efficiency of the method is illustrated by numerical examples with uncertainties from the initial data, boundary data and collision kernel.

AMS subject classifications: 35Q20, 65L60

Key words: Uncertainty quantification, Boltzmann equation, stochastic Galerkin methods, sparse approaches.

1. Introduction

The Boltzmann equation plays an essential role in kinetic theory [9]. It describes the time evolution of the density distribution of dilute gases, where fluid dynamics equations, such as the Euler equations and the Navier-Stokes equations, fail to provide reliable information. It is an indispensable tool in fields concerning non-equilibrium statistical mechanics, such as rarefied gas dynamics and astronautical engineering.

For most applications of the Boltzmann equation, the initial and boundary data are given by physical measurements, which inevitably bring measurement errors. Furthermore,

[*] Corresponding author. *Email addresses:* rshu2@math.wisc.edu (R. Shu), jingweihu@purdue.edu (J. Hu), sjin@wisc.edu (S. Jin)

due to the difficulty of deriving the collision kernels from first principles, empirical collision kernels are often used. Such kernels contain adjustable parameters which are determined by matching with experimental data [5]. This procedure involves uncertainty on the parameters in the collision kernel. To understand the impact of these random inputs on the solution of the Boltzmann equation, it is imperative to incorporate the uncertainties into the equation, and design numerical methods to solve the resulting system [30]. A proper quantification of uncertainty will provide reliable predictions and a guidance for improving the models. Since the uncertainties of the Boltzmann equations come from many independent sources, it is necessary to use a multi-dimensional random space to incorporate all the uncertainties. Moreover, a Karhunen-Loeve expansion of a random field will result in a multi-dimensional random space.

Various numerical methods have been developed to solve the problem of uncertainty quantification (UQ) [12, 19, 30, 31]. Monte-Carlo methods [23] use statistical sampling in the random space, which give halfth order convergence in any dimension. Stochastic collocation methods [2, 4, 22] take sampling points on a well-designed grid, usually according to a quadrature rule, or take sampling points by least-square or compressed sensing approaches, and the statistical moments are computed by numerical quadratures or reconstructed generalized polynomial chaos expansions. Stochastic Galerkin methods [3, 4] use an orthonormal basis expansion in the random space. By a truncation of the expansion and Galerkin projection, one is led to a deterministic system of expansion coefficients. Both methods can achieve spectral accuracy in one-dimensional random space if the quadrature rule (orthonormal basis) is well chosen.

Hu and Jin [16] gave a first numerical method to solve the Boltzmann equation with uncertainty by a generalized polynomial chaos based stochastic Galerkin method. By a singular value decomposition on a set of basis related coefficients, together with the fast spectral method for the Boltzmann collision operator proposed by [21], the computational cost of the collision operator is decreased dramatically. However, their work focuses on low dimensional random spaces, and a direct extension of their method to multi-dimensional random spaces will suffer from the curse of dimensionality, which means K, the total number of basis functions, will grow like $K = \binom{K_1+d}{K_1}$, where K_1 is the number of basis in one dimension, and d is the dimension of the random space. This cost is not affordable if both K_1 and d are large. Monte-Carlo methods are feasible, but a halfth order convergence rate can be unsatisfactory in many applications. Therefore it is desirable to have an efficient and accurate method to solve the Boltzmann equation with multi-dimensional random inputs.

In this work, we adopt a sparse approach [8, 11] for the stochastic Galerkin method to circumvent the curse of dimensionality. The idea of sparse approaches traces back to Smolyak [28]. In recent years, sparse approaches have become a major way to break the curse of dimensionality in various contexts, for example in Galerkin finite element methods [8, 27, 33], finite difference methods [13, 14], high-dimensional stochastic differential equations [24, 32] and uncertainty quantification [20, 25]. The sparse approach we adopt was first proposed by Schwab et al. [26] for transport-dominated diffusion problems, and then applied to discontinuous Galerkin methods for elliptic equations by Wang et al. [29] and transport equations by Guo and Cheng [15]. Simply speaking, we start from a hierarchical basis in one dimension. To construct the sparse wavelet basis in multi-dimension, we take the tensor basis and discard those basis functions that are in deep levels in most dimensions. In this way only a small number of basis functions are kept, yet it can be proved that the accuracy is still as good as the corresponding tensor basis, if the function to approximate is smooth enough. With a hierarchical basis with N levels and piecewise polynomials of degree at most m, our method can achieve an accuracy of $O(N^{d-1}2^{-N(m+1)})$ with number of basis $K=O((m+1)^d 2^N N^{d-1})$ for d-dimensional random spaces. This accuracy is $O(K^{-(m+1)}(\log K)^{(m+2)(d-1)})$ in terms of K. It is algebraically accurate, but as d increases, the accuracy deteriorates very slowly. Furthermore, we discover a sparse structure of a set of basis related coefficients, S_{ijk}, which greatly reduces the cost of the expensive collision operator evaluation.

The rest of the paper is organized as follows: in Section 2 we introduce the Boltzmann equation with uncertainty and the framework of stochastic Galerkin (sG) method; in Section 3 we introduce our sparse method with multi-wavelet functions; in Section 4 we give an estimate of the sparsity of the coefficients S_{ijk}; in Section 5 we prove the random space regularity of the solution of the Boltzmann equation with uncertainty, as well as the accuracy of the sG method with sparse wavelet basis; in Section 6 we give some numerical results; the paper is concluded in Section 7.

2. The Boltzmann equation with uncertainty

The classical (deterministic) Boltzmann equation in its dimensionless form reads

$$\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f = \frac{1}{\mathrm{Kn}} Q(f,f), \tag{2.1}$$

where $f=f(t,\mathbf{x},\mathbf{v})$ is the density distribution function of a dilute gas at time $t\in\mathbb{R}^+$, position $\mathbf{x}\in\Omega\subset\mathbb{R}^{d_x}$, and with particle velocity $\mathbf{v}\in\mathbb{R}^{d_v}$. Kn is the Knudsen number, a dimensionless number defined as the ratio of the mean free path and a typical length scale, such as the size of the spatial domain. The collision operator $Q(f,f)$ is given by

$$Q(f,f)=\int_{\mathbb{R}^{d_v}}\int_{\mathbb{S}^{d_v-1}} B(\mathbf{v},\mathbf{v}_*,\sigma)\left[f(\mathbf{v}')f(\mathbf{v}'_*)-f(\mathbf{v})f(\mathbf{v}_*)\right]\,\mathrm{d}\sigma\,\mathrm{d}\mathbf{v}_*, \tag{2.2}$$

which is a quadratic integral operator modeling the binary elastic collision between particles. $(\mathbf{v},\mathbf{v}_*)$ and $(\mathbf{v}',\mathbf{v}'_*)$ are the particle velocities before and after a collision, which are given by

$$\begin{cases} \mathbf{v}'=\dfrac{\mathbf{v}+\mathbf{v}_*}{2}+\dfrac{|\mathbf{v}-\mathbf{v}_*|}{2}\sigma, \\ \mathbf{v}'_*=\dfrac{\mathbf{v}+\mathbf{v}_*}{2}-\dfrac{|\mathbf{v}-\mathbf{v}_*|}{2}\sigma, \end{cases} \tag{2.3}$$

with a vector σ varying on the unit sphere. The collision kernel B is a non-negative function of the form $B(\mathbf{v},\mathbf{v}_*,\sigma)=B(|\mathbf{v}-\mathbf{v}_*|,\cos\theta)$, where $\theta=\arccos\frac{\sigma\cdot(\mathbf{v}-\mathbf{v}_*)}{|\mathbf{v}-\mathbf{v}_*|}$ is the deviation angle. A commonly used model for the collision kernel is the variable hard sphere (VHS) model [5], which takes the form

$$B = b|\mathbf{v}-\mathbf{v}_*|^\lambda, \tag{2.4}$$

where b and λ are some constants whose values are usually determined by matching with the experimental data to reproduce the correct transport coefficients such as the viscosity.

The Boltzmann collision operator satisfies the conservation laws

$$\int_{\mathbb{R}^{d_v}} Q(f,f)\begin{pmatrix} 1 \\ \mathbf{v} \\ |\mathbf{v}|^2 \end{pmatrix} \mathrm{d}\mathbf{v} = 0, \tag{2.5}$$

as well as the H-theorem

$$-\int_{\mathbb{R}^{d_v}} Q(f,f)\ln f \,\mathrm{d}\mathbf{v} \geq 0. \tag{2.6}$$

The equality is achieved if and only if f takes the form

$$M(\mathbf{v})_{(\rho,\mathbf{u},T)} = \frac{\rho}{(2\pi T)^{d_v/2}} e^{-\frac{(\mathbf{v}-\mathbf{u})^2}{2T}}, \tag{2.7}$$

which is called the Maxwellian. ρ, $\mathbf{u}$ and T are the density, bulk velocity and temperature, given by

$$\rho = \int_{\mathbb{R}^{d_v}} f \,\mathrm{d}\mathbf{v}, \quad \mathbf{u} = \frac{1}{\rho}\int_{\mathbb{R}^{d_v}} f\mathbf{v}\mathrm{d}\mathbf{v}, \quad T = \frac{1}{d_v\rho}\int_{\mathbb{R}^{d_v}} f|\mathbf{v}-\mathbf{u}|^2\,\mathrm{d}\mathbf{v}. \tag{2.8}$$

The initial condition of the Boltzmann equation is given by

$$f(0,\mathbf{x},\mathbf{v}) = f^0(\mathbf{x},\mathbf{v}), \tag{2.9}$$

and a boundary condition is needed if the spatial domain Ω is a proper subset of $\mathbb{R}^{d_x}$. We adopt the Maxwell boundary condition, which takes the form

$$f(t,\mathbf{x},\mathbf{v}) = g(t,\mathbf{x},\mathbf{v}), \quad \mathbf{x}\in\partial\Omega, \quad \mathbf{v}\cdot\mathbf{n} > 0, \tag{2.10}$$

with

$$\begin{aligned} g(t,\mathbf{x},\mathbf{v}) =&(1-\alpha)f(t,\mathbf{x},\mathbf{v}-2(\mathbf{v}\cdot\mathbf{n})\mathbf{n}) \\ &+\frac{\alpha}{(2\pi)^{(d_v-1)/2}T_w(\mathbf{x})^{(d_v+1)/2}} e^{-\frac{|\mathbf{v}|^2}{2T_w(\mathbf{x})}} \int_{\mathbf{v}\cdot\mathbf{n}<0} f(t,\mathbf{x},\mathbf{v})|\mathbf{v}\cdot\mathbf{n}|\,\mathrm{d}\mathbf{v}, \end{aligned} \tag{2.11}$$

where T_w is the temperature of the wall, and $\mathbf{n}$ is the inner normal unit vector of the wall. The first term is the specular reflective part, and the second term is the diffusive part. α is the accommodation coefficient. $\alpha = 1$ implies purely diffusive boundary, while $\alpha = 0$ implies purely reflective boundary. For simplicity we only consider the case where the wall is static.

As mentioned before, there are many sources of uncertainties in the Boltzmann equation, such as the initial data, boundary data, and collision kernel. To quantify these uncertainties we introduce the Boltzmann equation with uncertainty

$$\begin{cases} \partial_t f(t,\mathbf{x},\mathbf{v},\mathbf{z}) + \mathbf{v}\cdot\nabla_{\mathbf{x}} f(t,\mathbf{x},\mathbf{v},\mathbf{z}) = \dfrac{1}{\mathrm{Kn}} Q_{\mathbf{z}}(f,f), \quad t\in\mathbb{R}_+,\ \mathbf{x}\in\Omega\subset\mathbb{R}^{d_x},\ \mathbf{v}\in\mathbb{R}^{d_v},\ \mathbf{z}\in I_{\mathbf{z}}\subset\mathbb{R}^{d}, \\ f(0,\mathbf{x},\mathbf{v},\mathbf{z}) = f^0(\mathbf{x},\mathbf{v},\mathbf{z}), \quad \mathbf{x}\in\Omega,\ \mathbf{v}\in\mathbb{R}^{d_v},\ \mathbf{z}\in I_{\mathbf{z}}, \\ f(t,\mathbf{x},\mathbf{v},\mathbf{z}) = g(t,\mathbf{x},\mathbf{v},\mathbf{z}), \quad t\in\mathbb{R}_+,\ \mathbf{x}\in\partial\Omega,\ \mathbf{v}\in\mathbb{R}^{d_v},\ \mathbf{z}\in I_{\mathbf{z}}. \end{cases} \tag{2.12}$$

Here $\mathbf{z}\in I_{\mathbf{z}}$ is a d-dimensional random vector with probability distribution $\pi(\mathbf{z})$ characterizing the uncertainty in the system. We assume that the collision kernel has the form

$$B(\mathbf{v},\mathbf{v}_*,\sigma,\mathbf{z}) = b(\mathbf{z})B_0(\mathbf{v},\mathbf{v}_*,\sigma),$$

which means that $Q_{\mathbf{z}}$ can be written as

$$Q_{\mathbf{z}}(f,f) = b(\mathbf{z})Q(f,f).$$

The Maxwell boundary data $g(t,\mathbf{x},\mathbf{v},\mathbf{z})$ is given by

$$\begin{aligned} g(t,\mathbf{x},\mathbf{v},\mathbf{z}) =& (1-\alpha(\mathbf{z}))f(t,\mathbf{x},\mathbf{v}-2(\mathbf{v}\cdot\mathbf{n})\mathbf{n},\mathbf{z}) \\ &+ \frac{\alpha(\mathbf{z})}{(2\pi)^{(d_v-1)/2}T_w(\mathbf{x},\mathbf{z})^{(d_v+1)/2}} e^{-\frac{|\mathbf{v}|^2}{2T_w(\mathbf{x},\mathbf{z})}} \int_{\mathbf{v}\cdot\mathbf{n}<0} f(t,\mathbf{x},\mathbf{v},\mathbf{z})|\mathbf{v}\cdot\mathbf{n}|\,\mathrm{d}\mathbf{v}. \end{aligned} \tag{2.13}$$

To solve the stochastic system (2.12), Hu and Jin [16] proposed a stochastic Galerkin (sG) method. The idea is to approximate f by a truncated polynomial series:

$$f(t,\mathbf{x},\mathbf{v},\mathbf{z}) \approx f^K(t,\mathbf{x},\mathbf{v},\mathbf{z}) = \sum_{k=1}^{K} f_k(t,\mathbf{x},\mathbf{v})\Phi_k(\mathbf{z}), \tag{2.14}$$

where $\{\Phi_k(\mathbf{z})\}$ are an orthonormal polynomial basis, which satisfies

$$\int_{I_{\mathbf{z}}} \Phi_i(\mathbf{z})\Phi_j(\mathbf{z})\pi(\mathbf{z})\,\mathrm{d}\mathbf{z} = \delta_{ij}.$$

If one uses polynomials of degree at most K_1 in a d dimensional random space, then the number of basis functions is $K = \binom{K_1+d}{K_1}$. Substituting (2.14) into (2.12) and conducting a standard Galerkin projection, one gets

$$\partial_t f_k(t,\mathbf{x},\mathbf{v}) + \mathbf{v}\cdot\nabla_{\mathbf{x}} f_k(t,\mathbf{x},\mathbf{v}) = Q_k(f^K,f^K), \tag{2.15}$$

$$f_k(0,\mathbf{x},\mathbf{v}) = f_k^0(\mathbf{x},\mathbf{v}), \tag{2.16}$$

$$Q_k(f^K,f^K) = \sum_{i,j=1}^{K} S_{ijk} Q(f_i,f_j), \tag{2.17}$$

where

$$S_{ijk} = \int_{I_{\mathbf{z}}} b(\mathbf{z})\Phi_i(\mathbf{z})\Phi_j(\mathbf{z})\Phi_k(\mathbf{z})\pi(\mathbf{z})\,d\mathbf{z}. \tag{2.18}$$

The boundary condition is given by

$$\begin{aligned} g_k =& \sum_{j=1}^{K}\int_{I_{\mathbf{z}}}(1-\alpha(\mathbf{z}))\Phi_k(\mathbf{z})\Phi_j(\mathbf{z})\pi(\mathbf{z})\,d\mathbf{z} f_j(t,\mathbf{x},\mathbf{v}-2(\mathbf{v}\cdot\mathbf{n})\mathbf{n}) \\ &+\sum_{j=1}^{K} D_{kj}(\mathbf{x},\mathbf{v})\int_{\mathbf{v}\cdot\mathbf{n}<0} f_j(t,\mathbf{x},\mathbf{v},\mathbf{z})|\mathbf{v}\cdot\mathbf{n}|\,d\mathbf{v}, \end{aligned} \tag{2.19}$$

where

$$D_{kj}(\mathbf{x},\mathbf{v}) = \int_{I_{\mathbf{z}}} \frac{\alpha(\mathbf{z})}{(2\pi)^{(d_v-1)/2}T_w(\mathbf{x},\mathbf{z})^{(d_v+1)/2}} e^{-\frac{|\mathbf{v}|^2}{2T_w(\mathbf{x},\mathbf{z})}}\Phi_k(\mathbf{z})\Phi_j(\mathbf{z})\pi(\mathbf{z})\,d\mathbf{z} \tag{2.20}$$

is a matrix that is time independent hence can be pre-computed.

This gPC-sG method works well for low dimensional random inputs, but for high dimensional ones, it might require a very large number of basis functions (K large) to approximate f to a given accuracy. If one takes K_1 basis functions in each dimension of a d-dimensional random space, then a direct extension of the gPC-sG method will require $K = \binom{K_1+d}{K_1}$ basis functions, which is prohibitively expensive if both K_1 and d are large. Furthermore, since the computation of Q_k typically requires $O(K^2)$ times evaluation of the deterministic collision operator, one has to choose a relatively small K in order to afford the computation. Also, [16] uses the singular value decomposition of a size K matrix as pre-computation for the collision operator, which reduces the computational cost by one order of magnitude, but this pre-computation can be prohibitively expensive if K is large. In the following sections we propose a stochastic Galerkin method with sparse grid basis functions, which requires much fewer basis functions for multi-dimensional random spaces.

3. A sparse approach with multi-wavelet basis functions

3.1. The sparse wavelet basis construction

For simplicity we restrict to the case $I_{\mathbf{z}} = [-1,1]^d$, and $\pi(\mathbf{z}) = \frac{1}{2^d}$ is the uniform distribution. We follow the notation by Guo and Cheng [15]. We start by constructing a hierarchical decomposition of the space consisting of piecewise polynomials of degree at most m. Let $P^m(a,b)$ be the space of polynomials of degree at most m on the interval (a,b), and for every $N \geq 0$,

$$V_N^m = \{\phi : \phi \in P^m(-1+2^{-N+1}j, -1+2^{-N+1}(j+1)), j = 0,1,\dots,2^N-1\}. \tag{3.1}$$

Then define the wavelet space $W_N^m, N = 1,2,\dots$ as the orthogonal complement of V_{N-1}^m inside V_N^m. For convenience we define $W_0^m = V_0^m$. Then one obtains the hierarchical decomposition $V_N^m = \oplus_{0\leq j\leq N} W_j^m$.

Then a standard sparse trick can be applied. For simplicity we introduce the following vector notations:

If $\mathbf{i} = (i_1, \ldots, i_d)$, $\quad \mathbf{j} = (j_1, \ldots, j_d)$ then

$$\mathbf{i} \le \mathbf{j} \text{ means } i_1 \le j_1, \ldots, i_d \le j_d,$$

$$\binom{\mathbf{j}}{\mathbf{i}} := \binom{j_1}{i_1} \times \cdots \times \binom{j_d}{i_d},$$

$$\mathbf{1_m} \text{ is the vector with 1 at } m\text{-th component and 0 elsewhere,}$$

$$|\mathbf{i}|_\infty = \max_{1 \le m \le d} \{|i_m|\}, \quad |\mathbf{i}|_1 = |i_1| + \cdots + |i_d|.$$

Define the d-fold tensor product of V_N^m by

$$\mathbf{V}_{N,\mathbf{z}}^m = V_{N,z_1}^m \times \cdots \times V_{N,z_d}^m. \tag{3.2}$$

Similarly define the d-fold tensor product of $W_{\mathbf{j}}^m$ by

$$\mathbf{W}_{\mathbf{j},\mathbf{z}}^m = W_{j_1,z_1}^m \times \cdots \times W_{j_d,z_d}^m. \tag{3.3}$$

Then

$$\mathbf{V}_{N,\mathbf{z}}^m = \oplus_{0 \le |\mathbf{j}|_\infty \le N} \mathbf{W}_{\mathbf{j},\mathbf{z}}^m.$$

The sparse trick is to replace the l^∞ norm on $\mathbf{j}$ by the l^1 norm. In this way we define the sparse wavelet space

$$\hat{\mathbf{V}}_{N,\mathbf{z}}^m = \oplus_{0 \le |\mathbf{j}|_1 \le N} \mathbf{W}_{\mathbf{j},\mathbf{z}}^m. \tag{3.4}$$

From now on we will omit the subscript $\mathbf{z}$ for these spaces.

3.2. Construction of the basis functions

We adopt the basis functions of W_j^m constructed by Alpert [1]. The basis functions of W_j^m are denoted by $\psi_{j,l}^{m'}$, $m' = 0, 1, \ldots, m$, $l = 0, 1, \ldots, 2^{j-1}-1$ for $j \ge 1$ and $l = 0$ for $j = 0$. $\psi_{0,0}^{m'}$ are the orthonormal Legendre polynomials of degree m' on $[-1,1]$, and $\psi_{1,0}^{m'}$ are piecewise polynomials on $[-1,0]$ and $[0,1]$ that are orthogonal to those Legendre polynomials, which can be constructed by a procedure similar to the Gram-Schmidt orthogonalization. Other $\psi_{j,l}^{m'}$ are defined by dilation and translation of $\psi_{1,0}^{m'}$:

$$\psi_{j,l}^{m'}(y) = 2^{(j-1)/2} \psi_{1,0}^{m'}(2^{j-1} y + 2^{j-1} - 1 - 2l), \quad j = 2, 3, \ldots, \quad l = 0, 1, \ldots, 2^{j-1} - 1,$$

which has support on the interval $[-1 + 2^{2-j} l, -1 + 2^{2-j}(l+1)]$.

The basis functions of $\mathbf{W}_{\mathbf{j}}^m$ are tensor products of the one dimensional basis functions:

$$\psi_{\mathbf{j},\mathbf{l}}^{\mathbf{m}'}(\mathbf{z}) = \psi_{j_1,l_1}^{m'_1}(z_1) \times \cdots \times \psi_{j_d,l_d}^{m'_d}(z_d), \quad 0 \le |\mathbf{m}'|_\infty \le m, 0 \le l_1 \le 2^{j_1-1}-1, \ldots, 0 \le l_d \le 2^{j_d-1}-1,$$

and the basis functions of $\hat{\mathbf{V}}_N^m$ consist of all the above functions for $0 \le |\mathbf{j}|_1 \le N$. By reordering the basis functions for $\hat{\mathbf{V}}_N^m$ we make them $\Phi_1(\mathbf{z}), \ldots, \Phi_K(\mathbf{z})$, where $K = K(m, N, d)$ is the total number of basis functions. It is proved in Lemma 2.3 of [29] that

$$K = O((m+1)^d 2^N N^{d-1}). \tag{3.5}$$

4. Estimate of the Sparsity of S_{ijk}

Recall the triple product tensor S_{ijk} defined in (2.18). Due to the local support of the sparse wavelet basis functions Φ_k, this tensor is sparse, especially when N and d are large. Due to this sparsity, when one computes $Q_k = \sum_{i,j=1}^{K} S_{ijk} Q(f_i, f_j)$, one only needs to compute those $Q(f_i, f_j)$ where there is at least one k with $S_{ijk} \neq 0$. Now we prove some results on its sparsity. We focus on the dependence on N, so every $O(\cdot)$ notation means multiplication by a constant that may depend on d.

Recall that when one takes the sparse wavelet space $\hat{\mathbf{V}}_N^m$, the basis functions are

$$\begin{aligned} \psi_{\mathbf{j},\mathbf{l}}^{\mathbf{m}'}(\mathbf{z}) = \psi_{j_1,l_1}^{m'_1}(z_1) \times \cdots \times \psi_{j_d,l_d}^{m'_d}(z_d), \quad 0 \le |\mathbf{m}'|_\infty \le m, \\ 0 \le l_1 \le 2^{j_1 - 1} - 1, \ldots, 0 \le l_d \le 2^{j_d - 1} - 1, \quad |\mathbf{j}|_1 \le N. \end{aligned} \tag{4.1}$$

The function $\psi_{j,l}^{m'}(z)$ is supported on the interval $[-1+2^{2-j}l, -1+2^{2-j}(l+1)]$ for $j \ge 1$. Since this support is independent of m', we omit the m' index in the following consideration. If $\psi_{\mathbf{j}^1,\mathbf{l}^1}$ and $\psi_{\mathbf{j}^2,\mathbf{l}^2}$ have non-intersecting supports, then

$$\int_{I_\mathbf{z}} b(\mathbf{z}) \psi_{\mathbf{j}^1,\mathbf{l}^1}(\mathbf{z}) \psi_{\mathbf{j}^2,\mathbf{l}^2}(\mathbf{z}) \psi_{\mathbf{j}^3,\mathbf{l}^3}(\mathbf{z}) \pi(\mathbf{z})\, d\mathbf{z} = 0, \quad \forall \mathbf{j}_3, \mathbf{l}_3.$$

Recall that the number of basis functions, in $\hat{\mathbf{V}}_N^m$, which includes those $\psi_{\mathbf{j},\mathbf{l}}$ with $|\mathbf{j}|_1 \le N$ and $0 \le l_1 \le 2^{j_1-1} - 1, \ldots, 0 \le l_d \le 2^{j_d - 1} - 1$, is $O((m+1)^d 2^N N^{d-1})$. Thus the number of the pairs of such functions is $O((m+1)^{2d} 2^{2N} N^{2d-2})$. Now we state our result:

Theorem 4.1. *The pairs of basis functions of $\hat{\mathbf{V}}_N^m$ with intersecting supports have a total number at most $O((m+1)^{2d} 2^{2N} N^{d+1})$.*

Proof. The number of $\phi_{j,l}$ for a fixed j is $(m+1)2^{j-1}$ for $j \ge 1$, and $m+1$ if $j = 0$. Thus it is less than or equal to $(m+1)2^j$ for all j. For fixed j^1, j^2, suppose $j^1 \ge j^2$, then ϕ_{j^1,l^1} and ϕ_{j^2,l^2} have intersecting supports if and only if the support of ϕ_{j^1,l^1} is a subinterval of the support of ϕ_{j^2,l^2}. For every l^1, there is one and only one such l^2. Thus the number of pairs l^1, l^2 such that ϕ_{j^1,l^1} and ϕ_{j^2,l^2} have intersecting supports is 2^{j^1}, which is $2^{\max\{j^1,j^2\}}$ in general.

Thus the desired number is

$$S = (m+1)^{2d} \sum_{0 \le |\mathbf{j}^1|_1 \le N, 0 \le |\mathbf{j}^2|_1 \le N} 2^{\max\{j_1^1, j_1^2\} + \cdots + \max\{j_d^1, j_d^2\}}. \tag{4.2}$$

Let $\mathbf{k}^1 = \max\{\mathbf{j}^1, \mathbf{j}^2\}$, where the maximum acts on each component of vectors. Similarly let $\mathbf{k}^2 = \min\{\mathbf{j}^1, \mathbf{j}^2\}$. Then $|\mathbf{k}^1 + \mathbf{k}^2|_1 = |\mathbf{j}^1 + \mathbf{j}^2|_1 = |\mathbf{j}^1|_1 + |\mathbf{j}^2|_1 \le 2N$, and for each fixed $\mathbf{k}^1, \mathbf{k}^2$, there are at most 2^d pairs of $\mathbf{j}^1, \mathbf{j}^2$ satisfying the conditions $\mathbf{k}^1 = \max\{\mathbf{j}^1, \mathbf{j}^2\}$ and

$\mathbf{k}^2 = \min\{\mathbf{j}^1, \mathbf{j}^2\}$. Thus

$$\begin{aligned} S &\leq C(d)(m+1)^{2d} \sum_{0 \leq |\mathbf{k}^1|_1 + |\mathbf{k}^2|_1 \leq 2N} 2^{|\mathbf{k}^1|_1} \\ &= C(d)(m+1)^{2d} \sum_{k=0}^{2N} 2^k \binom{k+d-1}{d-1} \sum_{l=0}^{2N-k} \binom{l+d-1}{d-1} \\ &\leq C(d)(m+1)^{2d} N \sum_{k=0}^{2N} 2^k (k+1)^{d-1} (2N-k+1)^{d-1}. \end{aligned}$$

The first equality is because there are $\binom{k+d-1}{d-1}$ choices of $\mathbf{k}^1$ with $|\mathbf{k}^1|_1 = k$, and similarly for $\mathbf{k}^2$. The second inequality is because $\binom{k+d-1}{d-1} = \frac{k+1}{1}\frac{k+2}{2}\cdots\frac{k+d-1}{d-1} \leq (k+1)^{d-1}$, and taking the largest term in the l summation.

Then by taking derivative with respect of k, it is easy to see that the previous summation is optimized at $k_{max} = 2N - O(d)$. Thus

$$\begin{aligned} S &\leq C(d)(m+1)^{2d} N^2 2^{k_{max}} (k_{max}+1)^{d-1} (2N - k_{max} + 1)^{d-1} \\ &\leq C(d)(m+1)^{2d} 2^{2N} N^{d+1}, \end{aligned}$$

which finishes the proof.

Remark 4.1. *When $d \geq 4$, one has $2^{2N} N^{2d-2} > 2^{2N} N^{d+1}$, thus in this case the number of $Q(f_i, f_j)$ needed to be computed is much less than the total number of pairs of f_i, f_j. And the bigger d is, the more saving one will gain.*

Numerically, we observe this sparsity result even in the cases $d = 2, 3$ (see Section 6.1.3), and for a fixed d, the percentage of $Q(f_i, f_j)$ needed to be computed decreases exponentially as N increases, which is better than what one expects from the above theorem (where the percentage is $O(\frac{1}{N^{d-3}})$). This suggests that the above theorem is not sharp.

5. Regularity and accuracy

In this section, we prove the regularity of the solution to the Boltzmann equation in the random space, and the accuracy of the stochastic Galerkin method using sparse wavelet basis. These are straightforward multi-dimensional extensions of the corresponding results in [16]. We assume that the random collision kernel depends linearly on $\mathbf{z}$. This is a reasonable assumption because when one uses the Karhunen-Loeve expansion to approximate a random field, the resulting dependence on $\mathbf{z}$ is linear.

We consider the spatially homogeneous Boltzmann equation

$$\frac{\partial f}{\partial t} = Q(f, f), \tag{5.1}$$

subject to random initial data and random collision kernel

$$f(0, \mathbf{v}, \mathbf{z}) = f^0(\mathbf{v}, \mathbf{z}), \quad B = B(\mathbf{v}, \mathbf{v}_*, \sigma, \mathbf{z}), \quad \mathbf{z} \in I_{\mathbf{z}}.$$

5.1. Regularity in the random space for the Boltzmann equation

We define the norms and operators:

$$\|f(t,\cdot,\mathbf{z})\|_{L^p_{\mathbf{v}}} = \left(\int_{\mathbb{R}^{d_v}} |f(t,\mathbf{v},\mathbf{z})|^p \,\mathrm{d}\mathbf{v}\right)^{1/p}, \quad \|f(t,\mathbf{v},\cdot)\|_{L^2_{\mathbf{z}}} = \left(\int_{I_{\mathbf{z}}} f(t,\mathbf{v},\mathbf{z})^2 \pi(\mathbf{z})\,\mathrm{d}\mathbf{z}\right)^{1/2},$$

$$|||f(t,\cdot,\cdot)|||_k = \sup_{\mathbf{z}\in I_{\mathbf{z}}} \left(\sum_{|\mathbf{l}|=0}^{k} \|\partial_{\mathbf{z}}^{\mathbf{l}} f(t,\mathbf{v},\mathbf{z})\|^2_{L^2_{\mathbf{v}}}\right)^{1/2},$$

$$Q(g,h)(\mathbf{v}) = \int_{\mathbb{R}^{d_v}} \int_{\mathbb{S}^{d_v-1}} B(\mathbf{v},\mathbf{v}_*,\sigma,\mathbf{z})\left[g(\mathbf{v}')h(\mathbf{v}'_*) - g(\mathbf{v})h(\mathbf{v}_*)\right] \,\mathrm{d}\sigma\,\mathrm{d}\mathbf{v}_*,$$

$$Q_{1,j}(g,h)(\mathbf{v}) = \int_{\mathbb{R}^{d_v}} \int_{\mathbb{S}^{d_v-1}} \partial_{z_j} B(\mathbf{v},\mathbf{v}_*,\sigma,\mathbf{z})\left[g(\mathbf{v}')h(\mathbf{v}'_*) - g(\mathbf{v})h(\mathbf{v}_*)\right] \,\mathrm{d}\sigma\,\mathrm{d}\mathbf{v}_*.$$

We first state the following estimates of $Q(g,h)$ and $Q_{1,j}(g,h)$, which are standard results proved in [7,18] and its extension to the uncertain case is straightforward:

Lemma 5.1. *Assume the collision kernel B depends on $\mathbf{z}$ linearly, B and $\partial_{\mathbf{z}}B$ are locally integrable and bounded in $\mathbf{z}$. If $g,h \in L^1_{\mathbf{v}} \cap L^2_{\mathbf{v}}$, then*

$$\|Q(g,h)\|_{L^2_{\mathbf{v}}},\ \|Q_{1,j}(g,h)\|_{L^2_{\mathbf{v}}} \le C_B \|g\|_{L^1_{\mathbf{v}}} \|h\|_{L^2_{\mathbf{v}}}, \tag{5.2}$$

$$\|Q(g,h)\|_{L^2_{\mathbf{v}}},\ \|Q_{1,j}(g,h)\|_{L^2_{\mathbf{v}}} \le C_B \|g\|_{L^2_{\mathbf{v}}} \|h\|_{L^2_{\mathbf{v}}}, \tag{5.3}$$

where the constant $C_B > 0$ depends only on B and $\partial_{z_j}B$, $j = 1,\dots,d$.

Now we state our estimate on $|||f|||_k$.

Theorem 5.1. *Assume that B satisfies the assumption in Lemma 5.1, and $\sup_{\mathbf{z}\in I_{\mathbf{z}}} \|f^0\|_{L^1_{\mathbf{v}}} \le M$, $|||f^0|||_k < \infty$ for some integer $k \ge 0$. Then there exists a constant $C_k > 0$, depending only on C_B, M, T, and $|||f^0|||_k$ such that*

$$|||f|||_k \le C_k, \qquad \textit{for any} \quad t \in [0,T]. \tag{5.4}$$

The proof of the theorem is provided in the Appendix.

5.2. Accuracy analysis

In this subsection, we will prove the convergence rate of the stochastic Galerkin method using the previously established regularity. As in section 5.1, we will still restrict to the spatially homogeneous equation (5.1).

We use the sparse wavelet space $\hat{\mathbf{V}}_N^m$ with parameters m,N. For this space, the number of basis functions $K = O((m+1)^d 2^N N^{d-1})$.

Define the space $\mathscr{H}^m(I_{\mathbf{z}})$ by

$$\|f\|_{\mathscr{H}^m(I_{\mathbf{z}})} = \max \sum_{0\le m_{i_1},\dots,m_{i_r}\le m} \ \sum_{0\le m_{j_1},\dots,m_{j_{d-r}}\le 1} \|\partial_{z_{i_1}}^{m_{i_1}} \cdots \partial_{z_{i_r}}^{m_{i_r}} \partial_{z_{j_1}}^{m_{j_1}} \cdots \partial_{z_{j_{d-r}}}^{m_{j_{d-r}}} f\|_{L^2(I_{\mathbf{z}})},$$

where the maximum is taken over all non-empty subsets $\{i_1,\dots,i_r\} \subset \{1,\dots,d\}$, and $\{j_1,\dots,j_{d-r}\}$ is the complement of $\{i_1,\dots,i_r\}$. Using the orthonormal basis $\{\Phi_k(z)\}$, the solution f to (5.1) can be represented as

$$f(t,\mathbf{v},\mathbf{z}) = \sum_{k=1}^{\infty} f_k(t,\mathbf{v})\Phi_k(\mathbf{z}), \quad \text{where} \quad f_k(t,\mathbf{v}) = \int_{I_\mathbf{z}} f(t,\mathbf{v},\mathbf{z})\Phi_k(\mathbf{z})\pi(\mathbf{z})\,\mathrm{d}\mathbf{z}. \tag{5.5}$$

Let P_K be the projection operator defined as

$$P_K f(t,\mathbf{v},\mathbf{z}) = \sum_{k=1}^{K} f_k(t,\mathbf{v})\Phi_k(\mathbf{z}).$$

Then one has the following projection error estimate (Theorem 5.1 in [26]):

Lemma 5.2. *For any $f \in \mathscr{H}^{m+1}(I_\mathbf{z})$, $N \geq 1$, we have*

$$\|P_K f - f\|_{L^2(I_\mathbf{z})} \leq C(m,d)N^{d-1}\,2^{-N(m+1)}\|f\|_{\mathscr{H}^{m+1}(I_\mathbf{z})}. \tag{5.6}$$

This lemma implies that the projection error

$$\|P_K f - f\|_{L^2(I_\mathbf{z})} \leq C(m,d)K^{-(m+1)}(\log K)^{(m+2)(d-1)}\|f\|_{\mathscr{H}^{m+1}(I_\mathbf{z})}. \tag{5.7}$$

Define the norms

$$\|f(t,\cdot,\cdot)\|_{L^2_{\mathbf{v},\mathbf{z}}} = \left(\int_{I_\mathbf{z}}\int_{\mathbb{R}^d} f(t,\mathbf{v},\mathbf{z})^2\,\mathrm{d}\mathbf{v}\pi(\mathbf{z})\,\mathrm{d}\mathbf{z}\right)^{1/2}, \tag{5.8}$$

then we have the following:

Lemma 5.3. *Assume $\mathbf{z}$ obeys the uniform distribution, i.e., $\mathbf{z} \in I_\mathbf{z} = [-1,1]^d$ and $\pi(\mathbf{z}) = 1/2^d$. If $|||f^0|||_{d(m+1)}$ is bounded, then*

$$\|P_K f - f\|_{L^2_{\mathbf{v},z}} \leq C(m,d)K^{-(m+1)}(\log K)^{(m+2)(d-1)}, \tag{5.9}$$

where $C(m,d)$ is a constant depending on m and d.

Given the gPC approximation of f:

$$f^K(t,\mathbf{v},\mathbf{z}) = \sum_{k=1}^{K} \hat{f}_k(t,\mathbf{x},\mathbf{v})\Phi_k(\mathbf{z}), \tag{5.10}$$

we now define the error function

$$e^K(t,\mathbf{v},\mathbf{z}) = P_K f(t,\mathbf{v},\mathbf{z}) - f^K(t,\mathbf{v},\mathbf{z}) := \sum_{k=1}^{K} e_k(t,\mathbf{v})\Phi_k(\mathbf{z}),$$

where $e_k = \hat{f}_k - f_k$. Then we have

Theorem 5.2. *Assume the random variable z and initial data f^0 satisfy the assumption in Lemma 5.3, and the gPC approximation f^K is uniformly bounded in K, then*

$$\|f-f^K\|_{L^2_{\mathbf{v},\mathbf{z}}} \le C(t)\left\{C(m,d)K^{-(m+1)}(\log K)^{(m+2)(d-1)} + \|e^K(0)\|_{L^2_{\mathbf{v},\mathbf{z}}}\right\}.$$

The proof of Lemma 5.4 and Theorem 5.5 can be proved in the same way as Section 4.2 in Hu and Jin [16], in view of Lemma 5.3. We omit the details.

Remark 5.1. *In general, wavelet bases are used for functions with low regularity. Here we briefly explain the reason why we use them for smooth functions. For low dimensional random spaces ($d \le 4$), by choosing a large m (i.e., $m \ge 2$) one can obtain a good accuracy order (almost $(m+1)$-th order) with the wavelet basis. However, due to the factor $(m+1)^d$ in the number of basis functions K (see (3.5)), m cannot be large for higher dimensional random spaces ($d \ge 5$). Thus for such random spaces one has to sacrifice the accuracy order a little and take $m = 0, 1$ in order to make the number of basis functions K affordable.*

6. Numerical results

In this section we give some numerical results of the stochastic Galerkin method with sparse technique. We first demonstrate the efficiency of the sparse wavelet basis, and then show its application to the Boltzmann equation with uncertainty.

The random space is taken as $[0,1]^d$ with the uniform distribution. For the Boltzmann equation with uncertainty, the physical space is taken as $[0,1]$, and the velocity space is truncated as $[-R_v, R_v]^2$. The physical space is discretized into N_x grid points

$$x_i = (i+\frac{1}{2})\Delta x, \quad i = 0,1,\ldots,N_x-1, \tag{6.1}$$

where $\Delta_x = \frac{1}{N_x}$. The velocity space is discretized into N_v grid points in each dimension:

$$v_{i,j} = (-R_v + (i+\frac{1}{2})\Delta v, -R_v + (j+\frac{1}{2})\Delta v), \quad i,j = 0,1,\ldots,N_v-1, \tag{6.2}$$

where $\Delta v = \frac{2R_v}{N_v}$.

The flux term $\mathbf{v}\cdot\nabla_{\mathbf{x}} f_k$ in (2.16) is discretized by the second order upwind scheme with the minmod slope limiter. The collision operator is computed by the fast spectral method [21]. The time discretization is given by the second order Runge-Kutta scheme.

6.1. The sparse wavelet basis

6.1.1. Number of basis functions

We first give a comparison of number of basis functions between our sparse wavelet function space $\hat{\mathbf{V}}_N^m$ and the tensor basis $\mathbf{V}_N^m$. The result is shown in Table 1. It is clear that the sparse technique saves a great number of basis functions, especially in multi-dimensional random spaces.

(a) $m=0$

	$N=3$	$N=4$	$N=5$
$d=1$	8,8	16,16	32,32
$d=2$	20,64	48,256	112,1024
$d=3$	38,512	104,4096	272,32768
$d=4$	63,4096	192,65536	552,1048576

(b) $m=1$

	$N=3$	$N=4$	$N=5$
$d=1$	16,16	32,32	64,64
$d=2$	80,256	192,1024	448,4096
$d=3$	304,4096	832,32768	2176,262144

Table 1: Comparison of number of basis functions: m is the maximal degree of polynomials. d is the dimension; in each cell, the left number is the number of basis of functions of $\hat{\mathbf{V}}_N^m$; the right number is the number of basis of functions of $\mathbf{V}_N^m$.

6.1.2. Efficiency of the sparse wavelet function space

We give a comparison of the L^2 approximation error of $\hat{\mathbf{V}}_N^m$ and $\mathbf{V}_N^m$. For each random dimension $d=2,3,4$ we pick a smooth test function as follows:

$$f(\mathbf{z})=\frac{1}{2\pi\mathscr{K}(\mathbf{z})^2}\exp\left(-\frac{1}{2\mathscr{K}(\mathbf{z})}\right)\left(2\mathscr{K}(\mathbf{z})-1+\frac{1-\mathscr{K}(\mathbf{z})}{2\mathscr{K}(\mathbf{z})}\right), \tag{6.3}$$

where

$$\begin{aligned}
&\mathscr{K}_{d=2}(\mathbf{z})=1-0.5(0.5+0.1\sin(z_1)+0.1\sin(2z_2)),\\
&\mathscr{K}_{d=3}(\mathbf{z})=1-0.5(0.5+0.1\sin(z_1)+0.1\sin(2z_2)+0.1\cos(z_3)),\\
&\mathscr{K}_{d=4}(\mathbf{z})=1-0.5(0.5+0.1\sin(z_1)+0.1\sin(2z_2)+0.1\cos(z_3)+0.1\cos(2z_4)).
\end{aligned} \tag{6.4}$$

We use the function spaces $\hat{\mathbf{V}}_N^m$ and $\mathbf{V}_N^m$ with different m, N values to approximate these functions, and compute their relative L^2 error $\frac{\|f-P_K f\|_{L^2}}{\|f\|_{L^2}}$, where P_K is the projection operator onto the corresponding function space. The result is shown in Figure 1. It can be seen that the sparse wavelet method performs much better than the tensor method.

6.1.3. Sparsity of S_{ijk}

We give a test of the sparsity of the tensor S_{ijk}, as well as the number of $Q(f_i,f_j)$ needed to compute. We take a random collision kernel $b(\mathbf{z})=1+0.2z_1$. For simplicity we only show the results with $m=0$, since the sparsity of S_{ijk} with larger m is similar. The result is shown in Figure 2. One can clearly see an exponential decay of the percentage of nonzeros in S_{ijk}, as well as the percentage of $Q(f_i,f_j)$ needed to compute, as N or d increase. This is even better than what we have proved.

To further demonstrate the sparsity of S_{ijk} we give a graph of nonzero elements of S_{ijk} for $m=0, N=4, d=3$, shown in Figure 3. The points in the first graph represent nonzero elements in S_{ijk}. The second graph is the projection of the first graph onto i,j coordinates, and the points in it represent those $Q(f_i,f_j)$ needed to compute.

6.2. Application to the Boltzmann equation with uncertainty

In this subsection, the velocity space is assumed to be two-dimensional and its discretization is always given by $N_v=32$. The time discretization is given by 0.8 times the

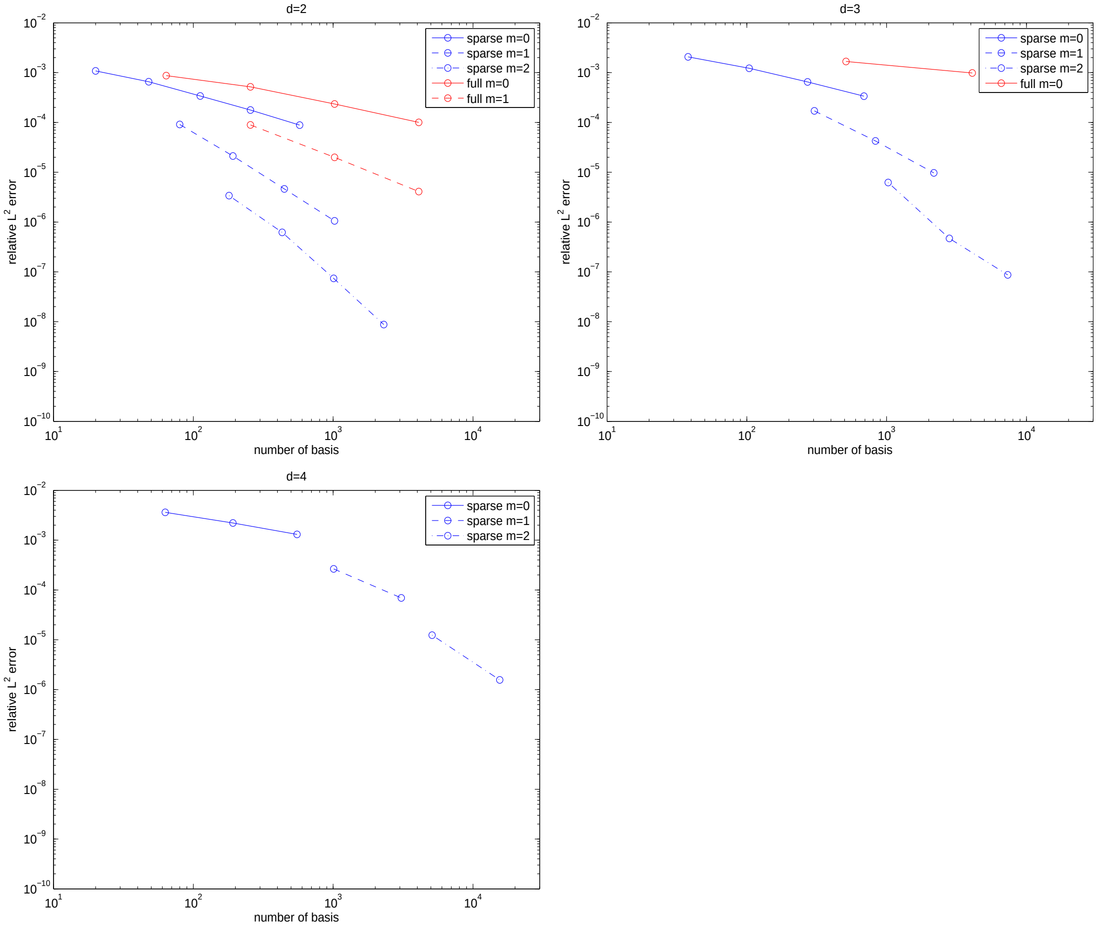

Figure 1: Comparison of approximation error of both sparse basis and full tensor basis for $d = 2,3,4$. For $d = 4$ we do not give the result by tensor basis because the number of basis functions is too large.

CFL condition for spatial inhomogeneous problems.

6.2.1. Accuracy of the approximation of the collision operator

We first check the accuracy of the collision operator $Q(f,f)$ computed by the sparse stochastic Galerkin method. The function f is given by the Bobylev-Krook-Wu [6,17] solution with uncertainty:

$$f(\mathbf{v},\mathbf{z}) = \frac{1}{2\pi\mathscr{K}(\mathbf{z})^2}\exp\left(-\frac{|\mathbf{v}|^2}{2\mathscr{K}(\mathbf{z})}\right)\left(2\mathscr{K}(\mathbf{z})-1+\frac{1-\mathscr{K}(\mathbf{z})}{2\mathscr{K}(\mathbf{z})}\mathbf{v}^2\right), \tag{6.5}$$

where

$$\begin{aligned}
&\mathscr{K}_{d=2}(\mathbf{z}) = 1-0.5(0.5+0.1\sin(z_1)+0.1\sin(2z_2)),\\
&\mathscr{K}_{d=3}(\mathbf{z}) = 1-0.5(0.5+0.1\sin(z_1)+0.1\sin(2z_2)+0.1\cos(z_3)),\\
&\mathscr{K}_{d=4}(\mathbf{z}) = 1-0.5(0.5+0.1\sin(z_1)+0.1\sin(2z_2)+0.1\cos(z_3)+0.1\cos(2z_4)).
\end{aligned} \tag{6.6}$$

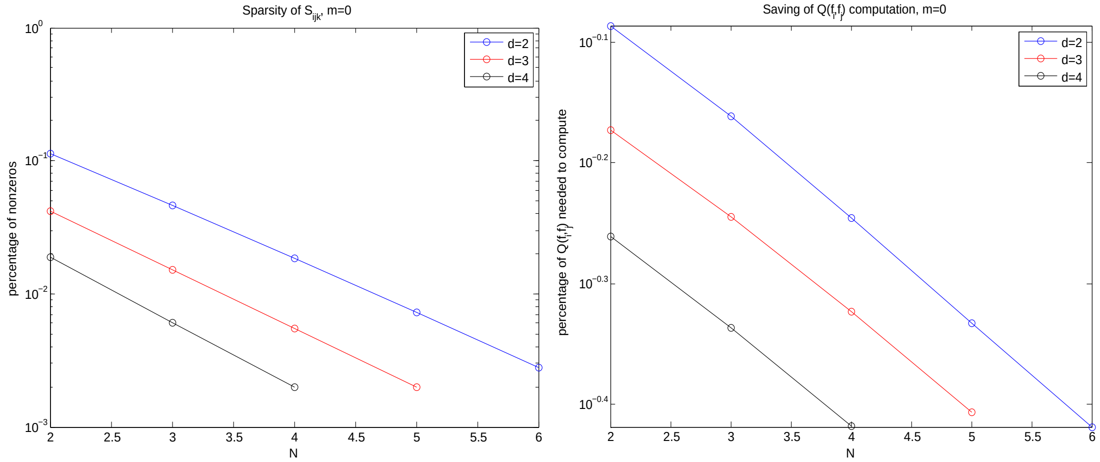

Figure 2: Sparsity of S_{ijk} and the number of $Q(f_i, f_j)$ needed to compute, $d = 2,3,4$, $m = 0$.

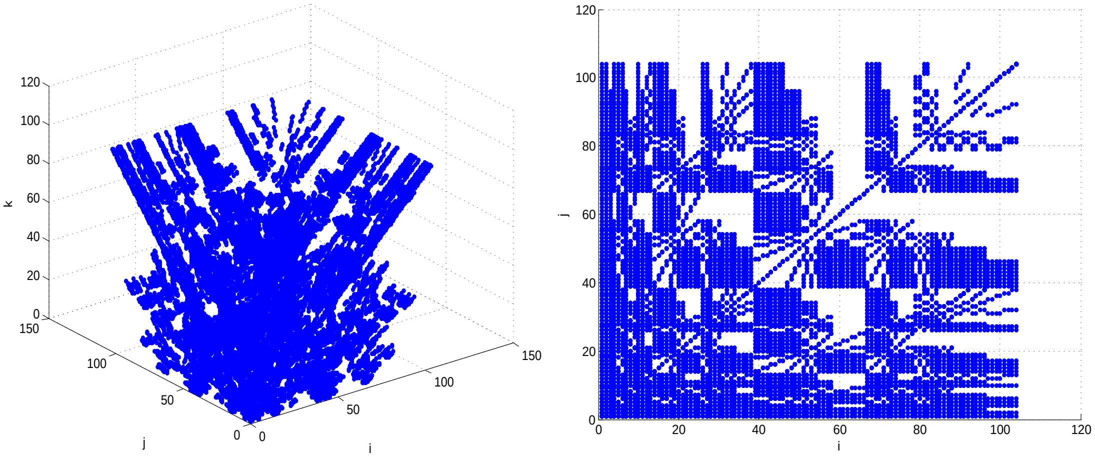

Figure 3: Demonstration of sparsity of S_{ijk}: $m = 0, N = 4, d = 3$. Left: blue points represent non-zeros terms of S_{ijk}. Right: blue points represent a pair (i, j) with $S_{ijk} \neq 0$ for some k.

For this f, $Q(f,f)$ with collision kernel $B = \frac{1}{2\pi}$ is given explicitly by

$$Q(f,f)(\mathbf{v},\mathbf{z}) = \Bigg(\bigg(-\frac{2}{\mathcal{K}(\mathbf{z})} + \frac{|\mathbf{v}|^2}{2\mathcal{K}(\mathbf{z})^2}\bigg) f \\ + \frac{1}{2\pi\mathcal{K}(\mathbf{z})^2} \exp\bigg(-\frac{|\mathbf{v}|^2}{2\mathcal{K}(\mathbf{z})}\bigg)\bigg(2 - \frac{1}{2\mathcal{K}(\mathbf{z})^2}|\mathbf{v}|^2\bigg)\Bigg) \frac{1-\mathcal{K}(\mathbf{z})}{8}. \tag{6.7}$$

The numerical solution is given by

$$\tilde{Q}(f,f)(\mathbf{v},\mathbf{z}) = \sum_{k=0}^{K} Q_k(\mathbf{v})\Phi_k(\mathbf{z}), \quad \text{where } Q_k(\mathbf{v}) = \sum_{i,j=0}^{K} S_{ijk} Q(f_i, f_j)(\mathbf{v}).$$

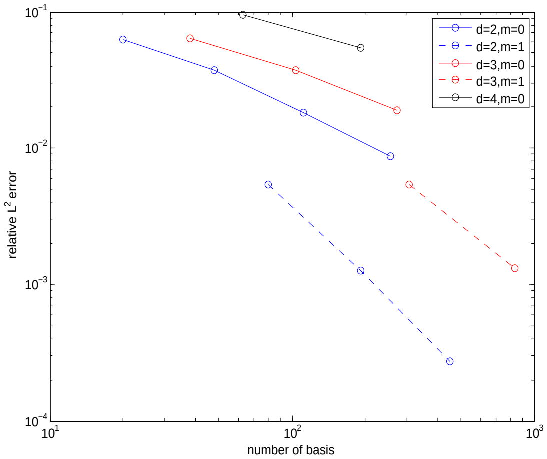

Figure 4: Accuracy of the approximation of the collision operator for $d = 2,3,4$.

We compare the relative L^2 error for $d = 2,3,4$ and sparse basis $\hat{\mathbf{V}}_N^m$ with different m, N. The result is shown in Figure 4. One can clearly see the error is a little worse than $O(K^{-(m+1)})$, and it becomes a little worse as d increases. This is caused by the $\log K$ factor in the error estimate.

6.2.2. The homogeneous Boltzmann equation with uncertainty on the collision kernel

We solve the homogeneous Boltzmann equation with deterministic initial data and a random collision kernel. We take the dimension of the random space $d = 2,3$, and the collision kernels are

$$\begin{aligned} b(\mathbf{z}) &= 1 + 0.2z_1 + 0.1z_2, \quad & d = 2, \\ b(\mathbf{z}) &= 1 + 0.2z_1 + 0.1z_2 + 0.07z_3, \quad & d = 3. \end{aligned} \tag{6.8}$$

The initial data is the BKW solution

$$f_0(\mathbf{v},\mathbf{z}) = \frac{1}{\pi}\exp(-|\mathbf{v}|^2)\frac{|\mathbf{v}|^2}{2}, \tag{6.9}$$

and the exact solution is given by

$$f(t,\mathbf{v},\mathbf{z}) = \frac{1}{2\pi\mathscr{K}^2}\exp\left(-\frac{|\mathbf{v}|^2}{2\mathscr{K}}\right)\left(2\mathscr{K} - 1 + \frac{1-\mathscr{K}}{2\mathscr{K}}|\mathbf{v}|^2\right), \tag{6.10}$$

with

$$\mathscr{K}(t,\mathbf{z}) = 1 - \exp(-b(\mathbf{z})t/8)/2. \tag{6.11}$$

We solve this equation by the sparse sG method with $m = 0$, time step $\Delta t = 0.01$ and final time $t = 1$, and check the relative L^2 error with the exact solution. The result is shown in Figure 5. The phenomenon is similar to the previous accuracy test.

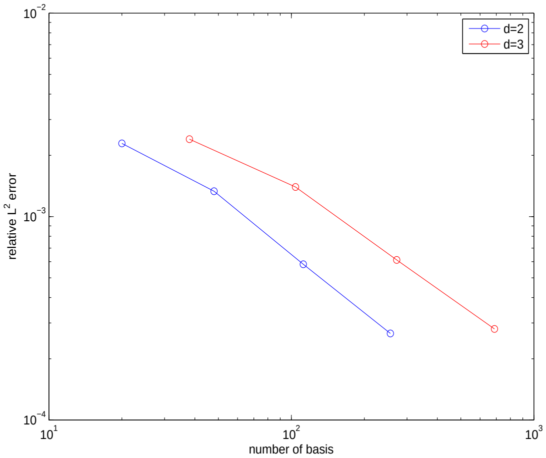

Figure 5: The homogeneous Boltzmann equation with a random collision kernel: accuracy result. $m=0$, $\Delta t=0.01$, $t=1$.

6.2.3. The Boltzmann equation with random initial data

We test our method on the (inhomogeneous) Boltzmann equation with uncertainty. The random space is 4-dimensional. We take the x-domain to be $[0,1]$ with the periodic boundary condition. We use the following random initial data to mimic the Karhunen-Loeve expansion

$$\begin{cases} \rho_0 = \frac{1}{3}\big(2+\sin(2\pi x)+\sin(4\pi x)z_1/2+\sin(6\pi x)z_2/4+\sin(8\pi x)z_3/6+\sin(10\pi x)z_4/7\big), \\ \mathbf{u}_0=(0.2,0), \\ T_0 = \frac{1}{4}\big(3+\cos(2\pi x)+\cos(4\pi x)z_1/2+\cos(6\pi x)z_2/4+\cos(8\pi x)z_3/6+\cos(10\pi x)z_4/7\big), \\ f = \frac{\rho_0}{4\pi T_0}\left(\exp(-\frac{|\mathbf{v}-\mathbf{u}_0|^2}{2T_0})+\exp(-\frac{|\mathbf{v}+\mathbf{u}_0|^2}{2T_0})\right). \end{cases} \tag{6.12}$$

The x-domain is discretized into $N_x=50$ mesh points, and we compare the solution by the sparse stochastic Galerkin method with $m=0, N=3$ and a stochastic collocation method with full tensor basis in random space at time $t=0.1$. The collocation method is implemented by solving the deterministic problem at points of the form $\mathbf{z}=(z_1,\dots,z_d)$ where each z_i is one of the $M_z=8$ Gauss-Legendre quadrature points (thus one needs to solve M_z^d deterministic problems). And then the mean and standard deviation are computed by numerical quadrature. The comparison result is shown in Figure 6. We see the results by the two methods agree well.

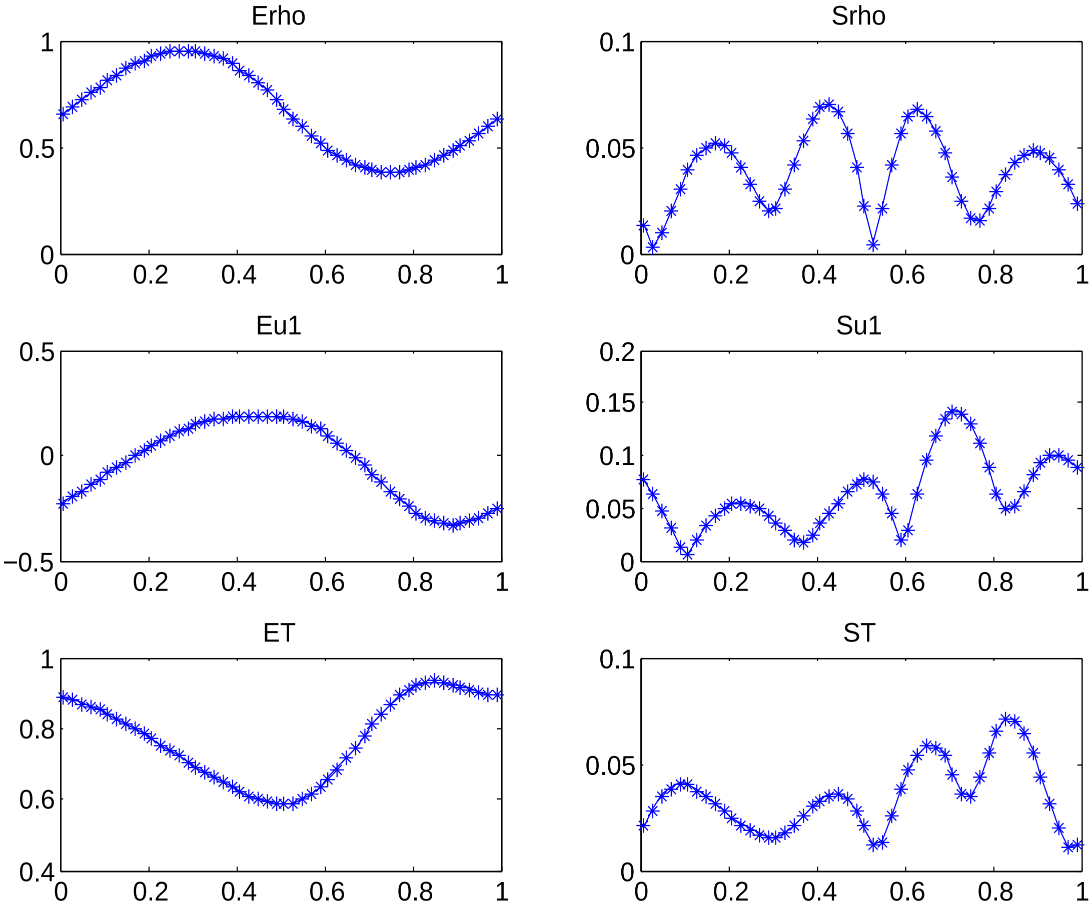

Figure 6: The Boltzmann equation with random initial data. $N_x = 50$, $t = 0.1$. Curve: collocation with $M_z = 8$; asterisks: Galerkin with $m = 0, N = 3$. Left column: mean of density, first component of bulk velocity, and temperature. Right column: standard deviation of density, first component of bulk velocity, and temperature.

6.2.4. The Boltzmann equation with randomness on initial data, boundary data, and collision kernel

We finally solve the inhomogeneous Boltzmann equation with uncertainty on initial data, boundary data, and collision kernel. The random domain is taken to be 6-dimensional. We take the initial data to be the equilibrium with

$$\rho(x,\mathbf{z}) = 1, \quad \mathbf{u}(x,\mathbf{z}) = 0, \quad T = 1+0.5(1+0.2z_2)\exp(-100(1+0.1z_3)(x-0.4-0.01z_1)^2), \tag{6.13}$$

and the boundary data is given by the Maxwellian boundary condition with random parameters

$$T_w = 1+0.2z_4, \quad \alpha = 0.5+0.3z_5. \tag{6.14}$$

The collision kernel is given by

$$b(\mathbf{z}) = 1+0.2z_6. \tag{6.15}$$

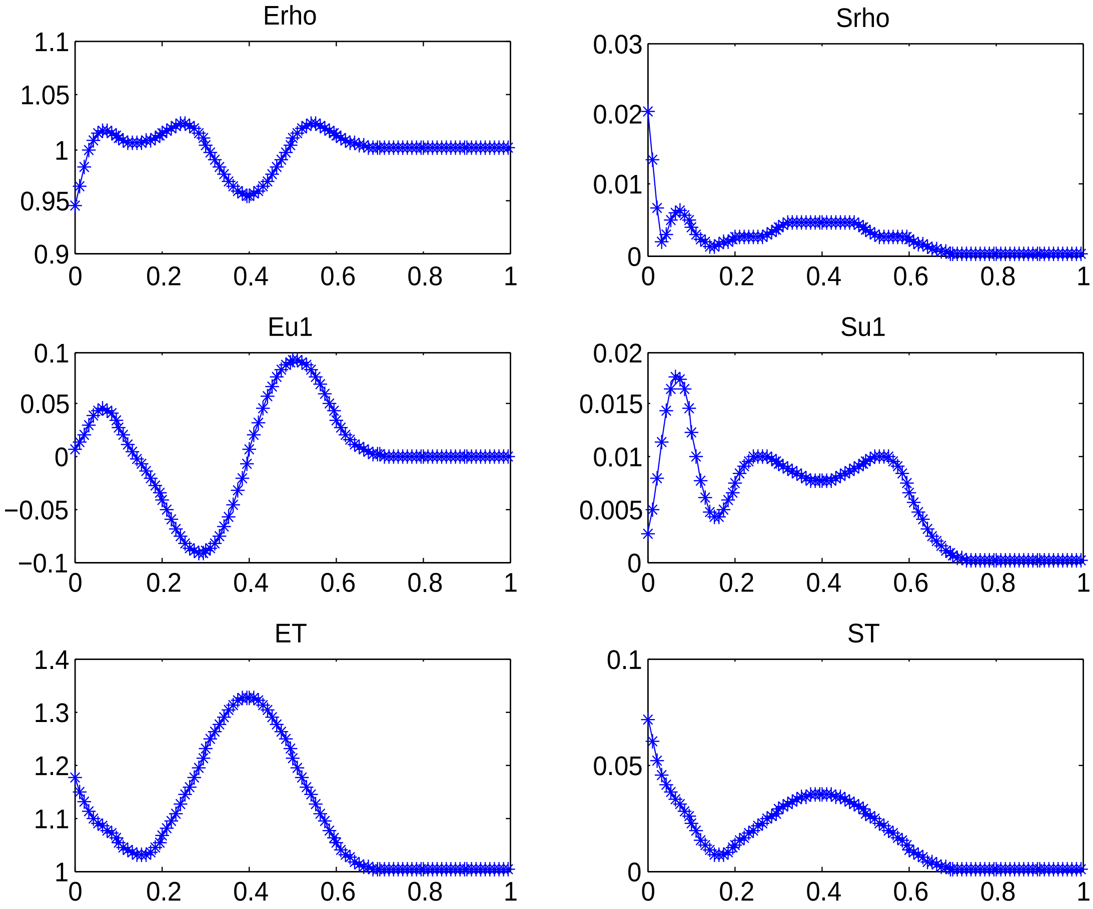

Figure 7: The Boltzmann equation with randomness on initial data, boundary data, and collision kernel ($d=6$). $N_x=100$, $t=0.04$. Curve: collocation with $M_z=4$; asterisks: Galerkin with $m=0, N=3$. Left column: mean of density, first component of bulk velocity, and temperature. Right column: standard deviation of density, first component of bulk velocity, and temperature.

The spatial discretization is given by $N_x = 100$ to better capture the details near the boundary. We compare the result by the stochastic Galerkin method with sparse technique with the stochastic collocation method with full grid at time $t = 0.04$. The Galerkin method has parameters $m = 0, N = 3$, and the collocation method is as described in the previous numerical result with $M_z = 4$ collocation points in each dimension. The result is shown in Figure 7. One can see that the two results agree well.

7. Conclusion

In this paper we developed a sparse wavelets based stochastic Galerkin method for the Boltzmann equation with uncertainty. The uncertainty could come from initial data, boundary data, and collision kernel. This method enables us to quantify the uncertainty from multi-dimensional random inputs, which is previously infeasible using the global gPC basis. We proved and numerically demonstrated the sparsity of the basis related coefficient, S_{ijk}, which allows us to dramatically accelerate the computation of the collision operator under the Galerkin projection. Regularity of the solution of the Boltzmann equation in the random space and an accuracy result of the stochastic Galerkin method are proved.

Many related problems are still open, for example, asymptotic-preserving schemes [10] for the Boltzmann equation with uncertainty, adaptive mesh techniques to capture discontinuities in the random space, quantification of nonlinear uncertainties on the collision kernel, etc.

Appendix: Proof of Theorem 5.2

Proof. First, from the conservation property of Q, one has

$$\|f(t,\cdot,\mathbf{z})\|_{L^1_\mathbf{v}} = \|f^0(\cdot,\mathbf{z})\|_{L^1_\mathbf{v}} \le M\,.$$

Then we use mathematical induction on k. For $k=0$, multiplying (5.1) by f and integrating on $\mathbf{v}$, by the Cauchy-Schwarz inequality and (5.2), one obtains

$$\frac{1}{2}\partial_t \int_{\mathbb{R}^d} f^2 \,\mathrm{d}\mathbf{v} = \int_{\mathbb{R}^d} f Q(f,f)\,\mathrm{d}\mathbf{v} \le \|f\|_{L^2_\mathbf{v}}\|Q(f,f)\|_{L^2_\mathbf{v}} \le C_B\|f\|_{L^1_\mathbf{v}}\|f\|^2_{L^2_\mathbf{v}} \le C_B M\|f\|^2_{L^2_\mathbf{v}}\,.$$

Now Gronwall's inequality implies that there is a positive constant C_0 such that (5.4) is true for $k=0$.

Now for some $k \ge 0$ assume (5.4) holds. Take any multi-index $\mathbf{j}$ with $|\mathbf{j}|_1 = k+1$. Taking $\mathbf{j}$-th derivative of z on (5.1) gives

$$\partial_t \partial^{\mathbf{j}}_{\mathbf{z}} f = \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{j}} \binom{\mathbf{j}}{\mathbf{l}} Q(\partial^{\mathbf{l}}_{\mathbf{z}} f, \partial^{\mathbf{j}-\mathbf{l}}_{\mathbf{z}} f) + \sum_{m=1}^{d} j_m \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{j}-\mathbf{1_m}} \binom{\mathbf{j}-\mathbf{1_m}}{\mathbf{l}} Q_{1,m}(\partial^{\mathbf{l}}_{\mathbf{z}} f, \partial^{\mathbf{j}-\mathbf{1_m}-\mathbf{l}}_{\mathbf{z}} f), \tag{A.1}$$

where we used the bilinearity of the collision operator and the assumption that B is linear in $\mathbf{z}$.

Multiplying (A.1) by $\partial^{\mathbf{j}}_{\mathbf{z}} f$ and integrating over $\mathbf{v}$ yields

$$\begin{aligned}
&\frac{1}{2}\partial_t \int_{\mathbb{R}^d} (\partial^{\mathbf{j}}_{\mathbf{z}} f)^2 \,\mathrm{d}\mathbf{v} \\
\le{}& \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{j}} \binom{\mathbf{j}}{\mathbf{l}} \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \|Q(\partial^{\mathbf{l}}_{\mathbf{z}} f, \partial^{\mathbf{j}-\mathbf{l}}_{\mathbf{z}} f)\|_{L^2_\mathbf{v}} + \sum_{m=1}^{d} j_m \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{j}-\mathbf{1_m}} \binom{\mathbf{j}-\mathbf{1_m}}{\mathbf{l}} \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \|Q_{1,m}(\partial^{\mathbf{l}}_{\mathbf{z}} f, \partial^{\mathbf{j}-\mathbf{1_m}-\mathbf{l}}_{\mathbf{z}} f)\|_{L^2_\mathbf{v}} \\
\le{}& \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{j}} \binom{\mathbf{j}}{\mathbf{l}} C_B \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \|\partial^{\mathbf{l}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \|\partial^{\mathbf{j}-\mathbf{l}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} + \sum_{m=1}^{d} j_m \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{j}-\mathbf{1_m}} \binom{\mathbf{j}-\mathbf{1_m}}{\mathbf{l}} C_B \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \|\partial^{\mathbf{l}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \|\partial^{\mathbf{j}-\mathbf{1_m}-\mathbf{l}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \\
\le{}& C_B C_k^2 \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \sum_{\mathbf{0}\le\mathbf{l}\le\mathbf{j},\mathbf{l}\ne\mathbf{0},\mathbf{j}} \binom{\mathbf{j}}{\mathbf{l}} + 2C_B C_0 \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|^2_{L^2_\mathbf{v}} + C_B C_k^2 \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} \sum_{m=1}^{d} j_m \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{j}-\mathbf{1_m}} \binom{\mathbf{j}-\mathbf{1_m}}{\mathbf{l}} \\
={}& (2^{k+1}-2) C_B C_k^2 \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}} + 2C_B C_0 \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|^2_{L^2_\mathbf{v}} + 2^k(k+1) C_B C_k^2 \|\partial^{\mathbf{j}}_{\mathbf{z}} f\|_{L^2_\mathbf{v}}.
\end{aligned} \tag{A.2}$$

In the first inequality we used the Cauchy-Schwarz inequality, and in the second inequality we used (5.3). In the third inequality the induction assumption is used for the second sum, since the indexes $\mathbf{l}$ and $\mathbf{j}-\mathbf{1_m}-\mathbf{l}$ appeared there have order less than or equal to k. Every term in the first sum can be treated similarly except terms corresponding to the cases of $\mathbf{l}=\mathbf{0}$ and $\mathbf{l}=\mathbf{j}$, which are treated separately. In the final equality, we used the identity $\sum_{l=0}^{L}\binom{L}{l}=(1+1)^L=2^L$.

Then we apply Gronwall's inequality to (A.2) and get the control

$$\sup_{\mathbf{z}\in I_{\mathbf{z}}}\left(\|\partial_{\mathbf{z}}^{\mathbf{j}}f(t,\mathbf{v},\mathbf{z})\|_{L_{\mathbf{v}}^2}^2\right)^{1/2}\le C_{k+1},$$

with a positive constant C_{k+1}. Sum over all $\mathbf{j}$ with $|\mathbf{j}|_1=k+1$ we get (5.4) for $k+1$. This completes the mathematical induction and the proof.

Acknowledgments

The second author's research was partially supported by NSF grant DMS-1620250 and a startup grant from Purdue University. The third author's research was supported by NSF grants DMS-1522184 and DMS-1107291: RNMS KI-Net, by NSFC grant No. 91330203, and by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison with funding from the Wisconsin Alumni Research Foundation. The authors would like to thank the referees for the helpful suggestions.

References

[1] B. Alpert, *A class of bases in L^2 for the sparse representation of integral operators*, SIAM Journal on Mathematical Analysis, 24 (1993), pp. 246–262.

[2] I. Babuska, F. Nobile, and R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 1005–1034.

[3] I. Babuska, R. Tempone, and G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM Journal on Numerical Analysis, 42 (2004), pp. 800–825.

[4] J. Back, F. Nobile, L. Tamellini, and R. Tempone, *Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison*, in Spectral and High Order Methods for Partial Differential Equations, E. M. R. J. S. Hesthaven, ed., Springer-Verlag Berlin Heidelberg, 2011.

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

[6] A. V. Bobylev, *One class of invariant solutions of the Boltzmann equation*, Akademiia Nauk SSSR, Doklady, 231 (1976), pp. 571–574.

[7] F. Bouchut and L. Desvillettes, *A proof of the smoothing properties of the positive part of Boltzmann's kernel*, Revista Matemática Iberoamericana, 14 (1998), pp. 47–61.

[8] H.-J. Bungartz and M. Griebel, *Sparse grids*, Acta Numerica, 13 (2004), pp. 147–269.

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

[10] F. FILBET AND S. JIN, *A class of asymptotic-preserving schemes for kinetic equations and related problems with stiff sources*, Journal of Computational Physics, 229 (2010), pp. 7625–7648.

[11] J. GARCKE AND M. GRIEBEL, *Sparse Grids and Applications*, Springer, 2013.

[12] R. G. GHANEM AND P. D. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.

[13] M. GRIEBEL, *Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences*, Computing, 61 (1998), pp. 151–179.

[14] M. GRIEBEL AND G. ZUMBUSCH, *Adaptive sparse grids for hyperbolic conservation laws*, in Hyperbolic Problems: Theory, Numerics, Applications, Springer, 1999, pp. 411–422.

[15] W. GUO AND Y. CHENG, *A sparse grid discontinuous Galerkin method for high-dimensional transport equations and its application to kinetic simulations*. SIAM Journal on Scientific Computing, accepted.

[16] J. HU AND S. JIN, *A stochastic Galerkin method for the Boltzmann equation with uncertainty*, Journal of Computational Physics, 315 (2016), pp. 150–168.

[17] M. KROOK AND T. T. WU, *Formation of Maxwellian tails,* Physics of Fluids, 20 (1977), pp. 1589–1595.

[18] P.-L. LIONS, *Compactness in Boltzmann's equation via Fourier integral operators and applications. I, II.*, Journal of Mathematics of Kyoto University, 34 (1994), pp. 391–427,429–461.

[19] O. P. L. MAÎTRE AND O. M. KNIO, *Spectral Methods for Uncertainty Quantification, Scientific Computation, with Applications to Computational Fluid Dynamics*, Springer, New York, 2010.

[20] O. P. L. MAÎTRE, H. N. NAJM, R. G. GHANEM, AND O. M. KNIO, *Multi-resolution analysis of Wiener-type uncertainty propagation schemes*, Journal of Computational Physics, 197 (2004), pp. 502–531.

[21] C. MOUHOT AND L. PARESCHI, *Fast algorithms for computing the Boltzmann collision operator*, Mathematics of Computation, 75 (2006), pp. 1833–1852.

[22] A. NARAYAN AND T. ZHOU, *Stochastic collocation on unstructured multivariate meshes*, Communications in Computational Physics, 18 (2015), pp. 1–36.

[23] H. NIEDERREITER, P. HELLEKALEK, G. LARCHER, AND P. ZINTERHOF, *Monte Carlo and Quasi-Monte Carlo Methods 1996*, Springer-Verlag, 1998.

[24] F. NOBILE, R. TEMPONE, AND C. WEBSTER, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM Journal on Numerical Analysis, 46 (2008), pp. 2309–2345.

[25] D. SCHIAVAZZI, A. DOOSTAN, AND G. IACCARINO, *Sparse multiresolution stochastic approximation for uncertainty quantification*, Recent Advances in Scientific Computing and Applications, 586 (2013), p. 295.

[26] C. SCHWAB, E. SÜLI, AND R. A. TODOR, *Sparse finite element approximation of high-dimensional transport-dominated diffusion problems*, ESAIM: Mathematical Modelling and Numerical Analysis, 42 (2008), pp. 777–819.

[27] J. SHEN AND H. YU, *Efficient spectral sparse grid methods and applications to high-dimensional elliptic problems*, SIAM Journal on Scientific Computing, 32 (2010), pp. 3228–3250.

[28] S. SMOLYAK, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Doklady Akademii Nauk SSSR, 4 (1963), pp. 240–243.

[29] Z. WANG, Q. TANG, W. GUO, AND Y. CHENG, *Sparse grid discontinuous Galerkin methods for high-dimensional elliptic equations*. Journal of Computational Physics, accepted.

[30] D. XIU, *Fast numerical methods for stochastic computations: a review,* Communications in Computational Physics, 5 (2009), pp. 242–272.

[31] ——, *Numerical Methods for Stochastic Computation*, Princeton University Press, Princeton, New Jersey, 2010.

[32] D. XIU AND J. HESTHAVEN, *High-order collocation methods for differential equations with random inputs*, SIAM Journal on Scientific Computing, 27 (2005), pp. 1118–1139.

[33] C. ZENGER, *Sparse grids*, in Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, vol. 31, 1990.