A bi-fidelity method for the multiscale Boltzmann equation with random parameters [*]

Liu Liu[†], Xueyu Zhu [‡]

September 9, 2019

Abstract

In this paper, we study the multiscale Boltzmann equation with multi-dimensional random parameters by a bi-fidelity stochastic collocation (SC) method developed in [53, 72, 73]. By choosing the compressible Euler system as the low-fidelity model, we adapt the bi-fidelity SC method to combine computational efficiency of the low-fidelity model with high accuracy of the high-fidelity (Boltzmann) model. With only a small number of high-fidelity asymptotic-preserving solver runs for the Boltzmann equation, the bi-fidelity approximation can capture well the macroscopic quantities of the solution to the Boltzmann equation in the random space. A priori estimate on the accuracy between the high- and bi-fidelity solutions together with a convergence analysis is established. Finally, we present extensive numerical experiments to verify the efficiency and accuracy of our proposed method.

Keywords. Boltzmann equation, uncertainty, bi-fidelity models, multiple scales, stochastic collocation

1   Introduction

Kinetic equations are widely used in classical fields such as rarefied gas, plasma physics, astrophysics, also in emerging areas such as semiconductor device modeling, social and biological sciences. They model the non-equilibrium dynamics of a large number of particles from a statistical point of view [13]. The Boltzmann equation, as one of the most fundamental kinetic equations [11], is an integro-differential equation describing the time evolution of probability density distribution of particles in rarefied gas.

There have been extensive studies on the Boltzmann and related kinetic models, both in theory and numerical computation [67, 21]. One of the main computational challenges is that the kinetic problems often encounter multiple temporal and spatial scales, characterized by the Knudsen number, the dimensionless mean free path, that may vary in orders of magnitude in the computational domain, covering the regimes from fluid, transition to rarefied. Asymptotic-preserving (AP) schemes, which preserve the asymptotic transition from one scale to another at the discrete level, have been shown to be an effective computational paradigm in the past two decades [38, 39]. They allow efficient numerical approximations in all regimes–coarse mesh and large time steps can be used even in the fluid dynamic regime, without numerically resolving the small Knudsen number. For the space inhomogeneous Boltzmann equation, AP schemes were first designed using BGK-operator penalty based method [26]. Other approaches include the exponential integrator based methods [20] or micro-macro decomposition [44, 28]. Despite the successful development of AP solvers in recent years, the complexity and memory cost for computing the collision term still remains challenging [58, 52, 29, 27].

Another challenge, which has been ignored in the community until recently, is the issue of *uncertainties in kinetic models.* Kinetic equations, derived from N-body Newton's equations via the mean-field limit [8], typically contain an integral operator modeling interactions between particles. Calculating the collision kernel from first principles is extremely complicated and not possible for complex particle systems, thus only empirical formulas are used for general particles [12], which means that the collision kernel contains many uncertainties. Other sources of uncertainties may also come from inaccurate measurements of the initial or boundary data, forcing or source terms. See for example [41, 23, 34, 50, 17, 48, 49] for recent efforts on uncertainty quantification for kinetic equations, in particular [64, 35], where numerical schemes for the Boltzmann equation with multi-dimensional random inputs have been studied.

One widely used method in uncertainty quantification is stochastic collocation (SC) method, especially in conjunction with the gPC expansion and high-performance grids such as sparse grids. There have been many works developed, for example [1, 36, 54, 56, 31, 68, 6]. One of the challenges central to collocation approaches is the simulation cost. For many complex systems, in particular, the multiscale Boltzmann equation with multi-dimensional random inputs we are studying, an accurate high-fidelity deterministic simulation can be so time-consuming and memory demanding that only a few high-fidelity simulations can be afforded. As many stochastic algorithms such as SC require repetitive implementations of the deterministic solver, the overall accurate stochastic simulation can be difficult and even computationally infeasible.

Fortunately, there usually exist some approximate, less complex low-fidelity models for practical problems. Compared to the high-fidelity models, these low-fidelity models usually contain simplified physics and/or are simulated on a coarser physical mesh, and consequently, own a cheaper computational cost. Although their accuracy may not be high, the low-fidelity models are designed in such a way that they can resolve or capture certain important features of the underlying problem

and produce reliable and qualitative predictions. Despite numerous multi-fidelity algorithms have been developed in different communities from different perspectives [30, 14, 71, 42, 66, 61, 60, 59, 55, 25, 53, 72, 73], there have not been many attempts for kinetic equations with uncertainty in the multi-fidelity setting, except for the recent work [22]. In [22], the authors take the steady state or approximated time-dependent solution with the asymptotic behavior close to the fluid limit as control variate models to accelerate the convergence of standard Monte Carlo methods.

Despite numerous multi-fidelity algorithms have been developed for different applications [63, 62, 70, 69, 57, 45], the goal of our work is to adapt the bi-fidelity method developed in [53, 72, 73] to efficiently approximate the high-fidelity solutions of the Boltzmann equation with multi-dimensional random parameters and multiple scales. In rarefied gas dynamics, fluid models are derived when the mean free path of a particle is very small compared to the typical macroscopic length. One can perform a Hilbert or Chapman-Enskog expansion ([11]) of the solution to the Boltzmann equation in powers of the Knudsen number ε. At the leading order in ε, the distribution function approaches a local equilibrium – a Maxwellian whose parameters are the fluid variables (density, mean velocity and temperature) governed by the compressible Euler equations. When ε is small, the Euler equations provide us a good accuracy in the physical space and numerical efficiency. Motivated by the above observation, we take advantage of this multiscale nature of the kinetic problem and choose the low-fidelity model based on the Euler equations in this work. More specifically, we connect an ε-dependent microscopic model (Boltzmann equation) and its macroscopic model (Euler equations) through the corresponding macroscopic quantities of the Boltzmann equation. Our numerical experiments demonstrate that the bi-fidelity solutions can approximate the high-fidelity solutions well at a much reduced computational cost.

This paper is outlined as follows: In Section 2, we give an introduction of the Boltzmann equation and its macroscopic equations. Section 3 reviews the framework of SC method with multifidelity models and discusses how it is adapted to our problem under study. Section 4 establishes the convergence of the bi-fidelity approximation to the high-fidelity solution under suitable assumptions. In section 5, we provide extensive numerical experiments to illustrate the effectiveness and efficiency of our proposed method, where kinetic, fluid and mixed regimes are carefully examined. We conclude the paper in Section 6.

2   Introduction of the Boltzmann equation

2.1   The Boltzmann equation with uncertainty

We first give an introduction to the classical (deterministic) Boltzmann equation, known as one of the most celebrated kinetic equations for rarefied gas. A dimensionless form reads

$$\partial_t f + v \cdot \nabla_x f = \frac{1}{\varepsilon} \mathcal{Q}(f, f), \qquad (x, v) \in \Omega \times \mathbb{R}^{d_v}, \tag{2.1}$$

3

where $f(t,x,v)$ is the probability density distribution function, modeling the probability of finding a particle at time $t > 0$, at position x in a bounded domain $\Omega \subset \mathbb{R}^{d_x}$, with velocity $v \in \mathbb{R}^{d_v}$, where d_x and d_v are the dimensions of the x and v variables. Periodic boundary condition is considered. The parameter ε is the Knudsen number defined as the ratio of the mean free path over a typical length scale such as the size of the spatial domain. The collision operator $\mathcal{Q}$ is a quadratic integral operator modeling the binary elastic collision between particles, and is given by

$$\mathcal{Q}(f,f)(v) = \int_{\mathbb{R}^{d_v}} \int_{\mathbb{S}^{d_v-1}} B(|v - v_*|, \cos\theta) \left(f(v')f(v'_*) - f(v)f(v_*) \right) d\sigma dv_*. \quad (2.2)$$

(v, v_*) and (v', v'_*) are the velocity pairs before and after the collision, during which the momentum and energy are conserved; thus (v', v'_*) can be expressed in terms of (v, v_*) as follows:

$$\begin{cases} v' = \dfrac{v + v_*}{2} + \dfrac{|v - v_*|}{2}\sigma, \\ v'_* = \dfrac{v + v_*}{2} - \dfrac{|v - v_*|}{2}\sigma, \end{cases} \quad (2.3)$$

with the vector σ the scattering direction varying on the unit sphere $\mathbb{S}^{d_v-1}$. The collision kernel B is a non-negative function of the form $B(v, v_*, \sigma) = B(|v - v_*|, \cos\theta)$, where $\cos\theta = \frac{\sigma \cdot (v - v_*)}{|v - v_*|}$ is the deviation angle. We consider the variable hard sphere (VHS) model [7], with a commonly used form for the collision kernel:

$$B(v, v_*, \sigma) = b\,|v - v_*|^\gamma, \qquad -d_v < \gamma \le 1, \quad (2.4)$$

where b is a positive constant, $\gamma > 0$ corresponds to the hard potential, and $\gamma < 0$ is the soft potential. Denote the vector

$$m(v) = \left(1, v, \frac{|v|^2}{2} \right)^\top, \quad (2.5)$$

then

$$\int_{\mathbb{R}^{d_v}} \mathcal{Q}(f,f)m(v)\,dv = 0, \quad (2.6)$$

which correspond to the conservation of mass, momentum and kinetic energy of the collision operator.

The celebrated Boltzmann's H-theorem gives the dissipation of entropy ([11]):

$$\int_{\mathbb{R}^d} \mathcal{Q}(f,f)\ln f\,dv \le 0.$$

Furthermore, the equality holds if and only if f reaches the equilibrium state

$$M(v)_{\rho,u,T} = \frac{\rho}{(2\pi T)^{\frac{d_v}{2}}} \exp\left(-\frac{|v - u|^2}{2T} \right), \quad (2.7)$$

which is known as the Maxwellian. Here ρ, u and T are the density, bulk velocity and temperature, respectively:

$$\rho = \int_{\mathbb{R}^{d_v}} f(v)\,dv, \qquad u = \frac{1}{\rho} \int_{\mathbb{R}^{d_v}} f(v)v\,dv, \qquad T = \frac{1}{d_v\rho} \int_{\mathbb{R}^{d_v}} f(v)|v - u|^2\,dv. \quad (2.8)$$

There are many sources of uncertainties in the Boltzmann equation, such as the initial data, boundary data, and collision kernel. We introduce the Boltzmann equation with uncertainty

$$
\begin{cases}
\partial_t f + v \cdot \nabla_x f = \dfrac{1}{\varepsilon} \mathcal{Q}_z(f, f), \\
f(0, x, v, z) = f_I(x, v, z), \qquad (x, v) \in \Omega \times \mathbb{R}^{d_v}, z \in I_z.
\end{cases}
\tag{2.9}
$$

Here $z \in I_z$ is a d-dimensional random parameter with probability distribution $\pi(z)$ known in priori characterizing the uncertainty in the system. Without loss of generality, we only consider periodic boundary condition in space throughout this paper.

2.2 The macroscopic fluid equations

When the Knudsen number $\varepsilon > 0$ becomes very small, the macroscopic fluid dynamics describing the evolution of averaged quantities such as the density ρ, momentum ρu and temperature T of the gas, namely, the compressible Euler or Navier-Stokes equations, become adequate [2, 8].

Multiplying (2.9) by $m(v)$ and integrating with respect to v, by using the conservation property of $\mathcal{Q}$ given in (2.6), one gets a non-closed system of conservation laws

$$
\partial_t \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix} + \nabla_x \cdot \begin{pmatrix} \rho u \\ \rho u \otimes u + \mathbb{P} \\ Eu + \mathbb{P}u + \mathbb{Q} \end{pmatrix} = 0,
\tag{2.10}
$$

where E is the total energy defined by

$$
E = \langle \frac{1}{2}|v|^2 f \rangle = \frac{1}{2}\rho|u|^2 + \frac{d_v}{2}\rho T,
\tag{2.11}
$$

with $\langle \cdot \rangle$ denoted as a velocity average of the argument,

$$
\langle g \rangle = \int_{\mathbb{R}^{d_v}} g(v)\, dv.
$$

Here $\mathbb{P} = \langle (v - u) \otimes (v - u)f \rangle$ is the pressure tensor, and $\mathbb{Q} = \frac{1}{2}\langle (v - u)|v - u|^2 f \rangle$ is the heat flux vector. Note that the variables ρ, u and E in (2.10) depend on the random parameter z.

When $\varepsilon \to 0$, $f \to M(v)_{\rho, u, T}$. We can approximate f by $M(v)_{\rho, u, T}$ and use the expression (2.7), $\mathbb{P}$ and $\mathbb{Q}$ become

$$
\mathbb{P} = p\,\mathbb{I}, \qquad \mathbb{Q} = 0,
$$

where $p = \rho T$ is the pressure, $\mathbb{I}$ is the identity matrix. Then (2.10) reduces to the compressible Euler equations of gas dynamics for a mono-atomic gas:

$$
\partial_t \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix} + \nabla_x \cdot \begin{pmatrix} \rho u \\ \rho u \otimes u + p\,\mathrm{I} \\ (E + p)u \end{pmatrix} = 0,
\tag{2.12}
$$

which is known as a first order approximation with respect to ε to the Boltzmann equation (2.9). By the Chapman-Enskog expansion, the compressible Navier-Stokes equations give a second order approximation in ε to the distribution function of the Boltzmann equation [11].

5

3   A stochastic collocation method with bi-fidelity models

In this section, we first briefly review an efficient bi-fidelity approximation to the high-fidelity solution studied in [53, 72], then we shall discuss the motivation of our choices of the low-fidelity model in our current study.

3.1   A bi-fidelity Algorithm

Assume we have access to the high-fidelity solutions $u^H(z)$ and low-fidelity solutions $u^L(z)$. Let M be the number of affordable low-fidelity simulation runs, which can be very large. N denotes the number of high-fidelity simulation runs that can be afforded and is often very small, i.e., $M \gg N$. Let $\gamma = \{z_1, \cdots, z_k\}$, $k \geq 1$ be a set of sample points in I_z. Denote the low-fidelity snapshot matrix $u^L(\gamma) = [u^L(z_1), \cdots, u^L(z_k)]$ and the corresponding low-fidelity approximation space

$$U^L(\gamma) = \mathrm{span}\{u^L(\gamma)\} = \mathrm{span}\{u^L(z_1), \cdots, u^L(z_k)\}.$$

Similarly, the high-fidelity snapshot matrix and the correponding high-fidelity approximation as follows:

$$u^H(\gamma) = [u^H(z_1), \cdots, u^H(z_k)], \qquad U^H(\gamma) = \mathrm{span}\{u^H(\gamma)\}.$$

The bi-fidelity algorithm for approximating the high-fidelity solution consists of offline and online stages. In the offline stage, we employ the cheap low-fidelity model to explore the parameter space to find the most important parameter points. Within the online stage, we learn the approximation rule from the low-fidelity model for any given z, and apply it to construct the bi-fidelity approximation. The detailed algorithm is summarized in Algorithm.1.

Most of the steps in this algorithm are straightforward. It would be instructional to provide details for Step 3 (point selection) and Step 6 (bi-fidelity reconstruction).

Point selection. To select the subset γ_N, we shall search the parameter space by the greedy algorithm proposed in [53, 72]. Start with a trivial subspace $\gamma_0 = \varnothing$, and assume that the first $k-1$ important points $\gamma_{k-1} = \{z_{i_1}, \cdots, z_{i_{k-1}}\} \subset \Gamma$ have been selected. We shall choose the next point $z_{i_k} \in \Gamma$ as the point that maximizes the distance between its corresponding low-fidelity solution and the approximation space $U^L(\gamma_{k-1})$, spanned by the low-fidelity solutions on the existing point set γ_{k-1}, i.e.,

$$z_{i_k} = \arg \max_{z \in \Gamma} \mathrm{dist}(u^L(z), U^L(\gamma_{k-1})), \qquad \gamma_k = \gamma_{k-1} \cup z_{i_k}, \tag{3.1}$$

where $\mathrm{dist}(v, W)$ is the distance function between $v \in u^L(\Gamma)$ and subspace $W \subset U^L(\Gamma)$. The greedy procedure essentially serves the purpose of searching the linear independent basis set in the parameterized low-fidelity solution space. We remark that the whole algorithm allows an efficient implementation by standard linear algebra operations. See [53, 72] for more technical details.

Algorithm 1: bi-fidelity approximation

Offline:

1 Select a sample set $\Gamma = \{z_1, z_2, \ldots, z_M\} \subset I_z$.

2 Run the low-fidelity model $u_l(z_j)$ for each $z_j \in \Gamma$.

3 Select N "important" points from Γ and denote it by $\gamma_N = \{z_{i_1}, \cdots z_{i_N}\} \subset \Gamma$. Construct the low-fidelity approximation space $U^L(\gamma_N)$.

4 Run high-fidelity simulations at each sample point of the selected sample set γ_N. Construct the high-fidelity approximation space $U^H(\gamma_N)$.

Online:

5 For any given z, run the low-fidelity model to get the corresponding low-fidelity solution $u^L(z)$ and compute the low-fidelity coefficients by projection:

$$u^L(z) \approx \mathcal{P}_{U^L(\gamma_N)} u^L = \sum_{k=1}^{N} c_k(z) u^L(z_k).$$

6 Construct the bi-fidelity approximation by applying the sample approximation rule learned from the low-fidelity model:

$$u^B(z) = \sum_{k=1}^{N} c_k(z) u^H(z_k).$$

Bi-fidelity approximation. In the offline stage, we have constructed the low- and high-fidelity approximation space, $U^L(\gamma_N)$ (step 3) and $U^H(\gamma_N)$ (step 4), respectively. During the online stage, for any given sample point $z \in I_z$, we shall project the corresponding low-fidelity solution $u^L(z)$ onto the low-fidelity approximation space $U^L(\gamma_N)$:

$$u^L(z) \approx \mathcal{P}_{U^L(\gamma_N)}[u^L(z)] = \sum_{k=1}^{N} c_k(z) u^L(z_{i_k}),$$

where $\mathcal{P}_V$ is the projection operator onto a Hilbert space V and the corresponding projection coefficients $\{c_k\}$ are computed by the following projection:

$$\mathbf{G}^L \mathbf{c} = \mathbf{f}, \qquad \mathbf{f} = (f_k)_{1 \leq k \leq N}, \qquad f_k = \langle u^L(z), u^L(z_{i_k}) \rangle^L, \qquad (3.2)$$

where $\mathbf{G}^L$ is the Gramian matrix of $u^L(\gamma_N)$, defined by

$$(\mathbf{G}^L)_{ij} = \langle u^L(z_{i_k}), u^L(z_{j_k}) \rangle^L, \qquad 1 \leq k, i, j \leq N, \qquad (3.3)$$

with $\langle \cdot, \cdot \rangle^L$ the inner product associated with the approximation space $U^L(\gamma_N)$.

These low-fidelity coefficients $\{c_k\}$ serve as the surrogate of the corresponding high-fidelity coefficients of $u^H(z)$. Therefore, the sought bi-fidelity approximation of $u^H(z)$ can be constructed as follows:

$$u^B(z) = \sum_{k=1}^{N} c_k(z) u^H(z_{i_k}). \tag{3.4}$$

We emphasize that if the low-fidelity model can mimic the variations of the high-fidelity model in the parameter space, the low-fidelity coefficients can be a good approximation of the corresponding high-fidelity coefficients for a given sample z. We refer interested readers to [53, 72, 33] for details of the error analysis and justifications.

It is worth noting that since the number of low-fidelity basis is typically small ($\mathcal{O}(10)$ in our numerical tests), the cost of computing the low-fidelity projection coefficients by solving the linear system (3.2) is negligible. The dominant cost of the online step is one low-fidelity simulation run. If the low-fidelity solver is much cheaper than the high-fidelity solver, the speedup during the online stage can be significant.

3.2   The high- and low-fidelity models in our problem

Our purpose is to efficiently approximate high-fidelity solutions for the uncertain Boltzmann equation (2.9) for a fixed z, which is solved by a deterministic AP solver discussed in section 3.2.1. It is well known that existing solvers for deterministic kinetic equations are time-consuming and memory demanding due to its high-dimensional nature in the physical space. With the random parameter, it is more challenging to fully sweep the multi-dimensional parameter space by solving the Boltzmann equation repeatedly, especially given the complicated nonlinear collision operator in our model.

To mitigate this computational cost, we consider to choose the compressible Euler equations (2.12) as our low-fidelity model. It is a first-order approximation to the Boltzmann equation, which can mimic the variations of macroscopic quantities of the Boltzmann equation in the fluid regime up to a certain accuracy. Besides, it is worth noting that the macroscopic quantities do not depend on the velocity v in (2.10). Therefore, solving the deterministic Euler equation is much easier and more efficient in terms of memory and computational time compared to solving the deterministic Boltzmann equation (2.1). These facts motivate us to choose the Euler equation as the low-fidelity model in our numerical experiments. A comparison of the computation cost (CPU time) for the two models are given in Section 5.

3.2.1   A high-fidelity solver

To solve the high-fidelity model Boltzmann equation, we shall resort to a high-fidelity asymptotic-preserving (AP) solver. There have been many works in developing robust numerical schemes for kinetic equations in the framework of the asymptotic-preserving scheme, see for example [5, 44, 43, 40, 16, 28]. As pointed in [26], AP scheme for the kinetic equation has two major merits: 1) as the Knudsen number ε go to zero, it automatically becomes a consistent and stable scheme for

the limiting fluid equation, with the stability condition independent of ε (i.e., Δt independent of ε); 2) the implicit collision terms can be implemented explicitly, free of Newton-type nonlinear algebraic solvers. Compared with multi-physics domain decomposition methods [18], AP schemes avoid the coupling of physical equations of different scales where coupling conditions and interface locations are difficult to determine. In contrast with many existing multiscale solvers, the AP schemes only require solving one equation – the kinetic equation and it becomes a robust macroscopic solver automatically when $\varepsilon \to 0$.

For our problem, we shall employ an AP scheme developed in [26] for the deterministic rescaled Boltzmann equation (2.1) as our high-fidelity solver. The main idea of [26] is to penalize the collision term $\mathcal{Q}(f, f)$ by the BGK operator $P(f) = M - f$ which can be inverted easily, thus the scheme can be solved explicitly. Let the initial distribution function be f_{in} and consider periodic boundary conditions. The basic scheme consists of the following two major steps:

1. We first discretize (2.1) in time by the following first-order semi-discrete scheme:
$$\frac{f^{n+1} - f^n}{\Delta t} + v \cdot \nabla_x f^n = \frac{\mathcal{Q}(f^n, f^n) - \beta^n(M^n - f^n)}{\varepsilon} + \frac{\beta^n(M^{n+1} - f^{n+1})}{\varepsilon}. \quad (3.5)$$

   f^{n+1} can be rewritten as follows:
$$f^{n+1} = \frac{1}{1 + \frac{\Delta t}{\varepsilon}\beta^n} \left[f^n - \Delta t \, v \cdot \nabla_x f^n + \frac{\Delta t}{\varepsilon} \left(\mathcal{Q}(f^n, f^n) - \beta^n(M^n - f^n) + \beta^n M^{n+1} \right) \right],$$
   $$(3.6)$$

   where β is some constant that depends on the spectral radius of the linearized collision operator of $\mathcal{Q}$ around the Maxwellian M. We refer to [26, Section 2] on the intuition and justification of choosing β. For example, one can set
$$\beta^n = \sup \left| \frac{\mathcal{Q}(f^n, f^n)}{f^n - M^n} \right|,$$

   and other choices β are also available [26, 28]. We numerically evaluate the collision term $\mathcal{Q}(f^n, f^n)$ in (3.5) by applying the fast spectral method developed in [52].

2. Though the above equation appears to be implicit due to M^{n+1}, it can be solved explicitly, thanks to the conservation property of $\mathcal{Q}(f, f)$ and the BGK operator $P(f)$. By multiplying the equation (3.5) with the vector $m(v)$ in (2.5), we can get the following equation:
$$\frac{W^{n+1} - W^n}{\Delta t} + \nabla_x \cdot \langle vmf^n \rangle = 0. \quad (3.7)$$

   where $W := (\rho, \rho u, E)$ that consists of the macroscopic quantities (mass, momentum and energy). With W^{n+1}, one computes M^{n+1} from the Maxwellian (2.7). Finally, we can update f^{n+1} explicitly from (3.6).

For the spatial discretization in (3.6), we employ a second order upwind MUSCL scheme as in [26], and a second order minmod slope limiter is used to suppress possible spurious oscillations near discontinuities or sharp gradients [46]. In addition to (3.7), a second order TVD scheme with a minmod slope limiter is also applied, see [5, 28] for details of implementation.

9

3.3 A low-fidelity solver

For the low-fidelity model, instead of solving the Euler system (2.12) directly, we shall semi-discretize its equivalent form (3.7) with f replaced by the Maxwellian $M(v)_{\rho,u,T}$, i.e.,

$$\frac{W^{n+1} - W^n}{\Delta t} + \nabla_x \cdot \langle vmM^n \rangle = 0, \tag{3.8}$$

where the relation between $W := (\rho, \rho u, E)$ and M is given in (2.11) and (2.7). The initial data of ρ, u and E are obtained from the initial distribution f_{in} for the Boltzmann equation, by using (2.8) and (2.11). That is, the initial data for the low- and high-fidelity models are consistent. The scheme (3.8) is numerically solved in the same way as equation (3.7) in the AP solver for the Boltzmann equation. Since Euler system is marching the macroscopic quantities, instead of marching the distribution solution f to the Boltzmann equation, the scheme (3.8) can be solved with a much reduced computational cost and memory consumption.

Remark 3.1. *We remark that instead of taking the solution f to the Boltzmann equation via the scheme (3.6) as the high-fidelity solutions, we consider its corresponding macroscopic quantities of interest $\mathcal{U}^H(z) = [\rho^H(z), u^H(z), T^H(z)]^\top$ as the high-fidelity snapshot solutions in order to connect the macroscopic quantities computed from the low-fidelity models. The low-fidelity solutions $\mathcal{U}^L(z) = [\rho^L(z), u^L(z), T^L(z)]^\top$ we considered are computed from the Euler system by using (3.8). During the point selection step to construct γ_N in Algorithm 1, we shall select the important parameter points based on the concatenated macroscopic quantity snapshot, namely $\mathcal{U}^L(z) = [\rho^L(z), u^L(z), T^L(z)]^\top$.*

Remark 3.2. *We acknowledge that there could be other choices of low-fidelity models that lead to more accurate bi-fidelity approximation, e.g., the compressible Navier-Stokes equations. To estimate if the low-fidelity model would be useful for constructing a reasonable accurate bi-fidelity approximation, one can explore an a priori estimate developed in the recent work [33].*

4 Accuracy and convergence analysis

To establish the accuracy and convergence results, we first give a summary of the hypocoercivity framework and notations used in [50], then introduce the relation between the solutions to the Boltzmann and compressible Euler system in suitable norms. To study the difference between the high- and bi-fidelity solutions, one can split it into two parts: the projection error and the remainder. In section 4.1, we show the estimate for the projection error in Theorem 4.1. In section 4.2, we give the regularity of high-fidelity solution, then prove the accuracy and convergence results of our bi-fidelity method adapted to the Boltzmann equation in Theorem 4.2.

4.1 The projection error

The subject of hydrodynamic limits and rigorous derivations of macroscopic models such as the fundamental PDEs of fluid mechanics from the kinetic theory of gases is a challenging task and has been studied for decades, see for example [24, 51, 10, 4, 3, 37, 32]. We shall show that for each fixed $z \in I_z$, the error between solutions to the Euler system and the macroscopic quantities (2.8) obtained from the Boltzmann equation (with consistent initial data) is small and of order ε, which will be described in (4.3).

Hypocoercivity framework. First, we review the hypocoercivity framework and notations for the norms used in [50]. Let f be the solution to the Boltzmann equation (2.1). Consider a linearization around the global equilibrium and perturbation of f:

$$f = \mathcal{M} + \varepsilon \sqrt{\mathcal{M}}\, h,$$

with

$$\mathcal{M}(v) = \frac{1}{(2\pi)^{\frac{d_v}{2}}} e^{-\frac{|v|^2}{2}},$$

then h satisfies the perturbed equation

$$\partial_t h + v \cdot \nabla_x h = \frac{1}{\varepsilon}\mathcal{L}(h) + \mathcal{F}(h, h), \tag{4.1}$$

where the linearized operator $\mathcal{L}$ and the nonlinear operator $\mathcal{F}$ are given by

$$\mathcal{L}(h) = \left(\sqrt{\mathcal{M}}\right)^{-1} \left[\mathcal{Q}(\sqrt{\mathcal{M}}h, \mathcal{M}) + \mathcal{Q}(\mathcal{M}, \sqrt{\mathcal{M}}h)\right],$$

$$\mathcal{F}(h, h) = 2\left(\sqrt{\mathcal{M}}\right)^{-1} \mathcal{Q}(\sqrt{\mathcal{M}}h, \sqrt{\mathcal{M}}h).$$

Denote $\partial_l^j := \partial/\partial v_j\, \partial/\partial x_l$ for multi-indices j and l. Introduce the following Sobolev norms:

$$||h||^2_{H^s_{x,v}} = \sum_{|j|+|l|\leq s} ||\partial_l^j h||_{L^2_{x,v}}, \qquad ||h||_{H^{s,r}_{x,v}} = \sum_{|\nu|\leq r} ||\partial^\nu h||^2_{H^s_{x,v}},$$

$$||h||_{H^s_{x,v}H^r_z} = \int_{I_z} ||h||_{H^{s,r}_{x,v}}\, \pi(z)dz, \qquad ||h||_{H^{s,r}_{x,v}L^\infty_z} = \sup_{z\in I_z} ||h||_{H^{s,r}_{x,v}}. \tag{4.2}$$

Refer to [9, Theorem 2.5], we extend its analysis to our case of the Boltzmann equation in the acoustic regime. Let h_ε be the perturbed solution to the linearized equation (4.1). Suppose the initial data for (4.1) and (2.12) are consistent for each z. If the initial distribution $h_{in} \in \text{Null}(\mathcal{L})$ and $h_{in} \in H^s_{x,v}$, then for each z, $(h_\varepsilon)_{\varepsilon>0}$ converges strongly to

$$h(t, x, v, z) = \left[\rho(t,x,z) + v\cdot u(t,x,z) + \frac{1}{2}(|v|^2 - d_v)T(t,x,z)\right]\mathcal{M}(v)$$

in $L^2_{[0,T]}H^s_x L^2_v$ as the Knudsen number $\varepsilon \to 0$, where ρ, u, T (with E obtained by (2.11)) satisfy the Euler system (2.12). We adapt our acoustic scaling to [9, Theorem 2.5] and get the follows: For all $z \in I_z$, if h_{in} belongs to $H^s_x L^2_v$, then

$$\sup_{t\in[0,\infty]} ||h(t,z) - h_\varepsilon(t,z)||_{L^2_{x,v}} \leq \sup_{t\in[0,\infty]} ||h(t,z) - h_\varepsilon(t,z)||_{H^s_x L^2_v} \leq C\max\{\varepsilon,\, \varepsilon V_T(\varepsilon)\}, \tag{4.3}$$

11

where $\forall T > 0$, $V_T(\varepsilon)$ is defined as

$$V_T(\varepsilon) = \sup_{t \in [0,T]} ||h(t,z) - h_\varepsilon(t,z)||_{L_x^\infty L_v^2} \to 0, \qquad \text{as } \varepsilon \to 0.$$

Error splitting. Let $\langle \cdot \rangle^H$ be an inner product space corresponding to the high-fidelity solution and $||\cdot||^H$ be the corresponding induced norm, see [53]. For each z, to study the total error $||u^H(z) - u^B(z)||^H$, one can split it into two parts:

$$\left|\left|u^H(z) - u^B(z)\right|\right|^H \leq \left|\left|u^H(z) - P_{U^H(\gamma_N^L)}u^H(z)\right|\right|^H + \left|\left|P_{U^H(\gamma_N^L)}u^H(z) - u^B(z)\right|\right|^H. \tag{4.4}$$

[53, Lemma 4.3] shows the estimate for the second term:

$$\left|\left|P_{U^H(\gamma_N^L)}u^H(z) - u^B(z)\right|\right|^H \leq c\left|\left|P_{U^H(\gamma_N^L)}u^H(z)\right|\right|^H + \left|\left|\sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\mathbf{Q}\,\mathbf{f}^L\right|\right|, \tag{4.5}$$

where we refer c to $\varepsilon_1 + \varepsilon_2 + \varepsilon_1\varepsilon_2$ in [53, Lemma 4.3], which is small, based on the reasonable assumptions made there. The last term above is related to the non-invertibility of high-fidelity Gramian matrix and usually negligible. Here $\mathbf{G}^L$ (or $\mathbf{G}^H$) is the Gramian matrix of $u^L(\gamma)$ (or $u^H(\gamma)$) given by (3.3), the vector $\mathbf{f}^L$ has entries

$$f_n^L = \left\langle u^L(z_n), u^L(z)\right\rangle^L,$$

and $\mathbf{Q} := \mathbf{I} - \mathbf{P}$ is the orthogonal projection onto its kernel (with $\mathbf{P}$ the orthogonal projection matrix onto its range), see [53]. In addition, since $\left|\left|P_{U^H(\gamma_N^L)}u^H(z)\right|\right|^H \leq ||u^H(z)||^H$, thus

$$\left|\left|P_{U^H(\gamma_N^L)}u^H(z) - u^B(z)\right|\right|^H \leq c\,||u^H(z)||^H + \left|\left|\sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\mathbf{Q}\,\mathbf{f}^L\right|\right|. \tag{4.6}$$

Projection error. The rest of this section will study the estimate for the projection error (the first term on the right-hand-side of (4.4)) and conclude it in Theorem 4.1. We now adapt the analysis in [53, subsection 4.1] and incorporate our high-fidelity (Boltzmann) and low-fidelity (Euler) models, by utilizing the knowledge of (4.3). Denote

$$z_n^H = \arg\max_{z \in I_z} d^H(u^H(z), U^H(\gamma_{n-1}^H)), \qquad \gamma_n^H = \gamma_{n-1}^H \cup \{z_n^H\}, \tag{4.7a}$$

$$z_n^L = \arg\max_{z \in I_z} d^L(u^L(z), U^L(\gamma_{n-1}^L)), \qquad \gamma_n^L = \gamma_{n-1}^L \cup \{z_n^L\}. \tag{4.7b}$$

The "best" achievable distance for approximation from a general N-dimensional subspace is the Kolmogorov N-width, defined by

$$d_N^H(u^H(I_z)) = \inf_{\dim(V_N)=N} \sup_{v \in u^H(I_z)} d^H(v, V_N). \tag{4.8}$$

The following is similar to [53, Lemma 4.1], except that we use the inequality (4.3). Since it is lengthy, we present as Lemma 1.1 with its proof in the Appendix.

Theorem 4.1. *If all the assumptions in Lemma 1.1 are satisfied for each $n = 0, 1, \cdots, N-1$, then*

$$\sup_{z \in I_z} ||u^H(z) - P_{U^H(\gamma_N^L)}u^H(z)||^H \leq C\sqrt{d_{N/2}(u^H(I_z))}, \tag{4.9}$$

12

where $d_N(u^H(I_z))$ is the N-width of the functional manifold $u^H(I_z)$, and the constant $C = \sqrt{2}(\delta_1 - \delta_2\,\varepsilon)^{-1}$ with $0 < \delta_1 - \delta_2\,\varepsilon < 1$.

Proof. Lemma 1.1 shows that it is a weak greedy procedure to use the nodal choices of z_n^L: $\exists\, 0 < \delta_1 - \delta_2\,\varepsilon < 1$ such that for all $n = 0, 1, \cdots, N-1$,

$$d^H(u^H(z_n^L), U^H(\gamma_{n-1}^L)) \geq (\delta_1 - \delta_2\,\varepsilon) \sup_{z \in I_z} d^H(u^H(z), U^H(\gamma_{n-1}^L)),$$

and [19, Corollary 3.3] indicates that (4.9) holds with $C = \sqrt{2}(\delta_1 - \delta_2\,\varepsilon)^{-1}$. $\qquad\square$

4.2 Smoothness of the solution and convergence analysis

In this section, we first study the smoothness of the high-fidelity solution u^H : $I_z \mapsto V^H$, where z is a multivariate random parameter and V^H is a Hilbert space with an inner product $\langle \cdot, \cdot \rangle^H$. Then we shall establish an estimate bound for the Kolmogorov width in Theorem 4.1, finally combine all the arguments and show the convergence result for our bi-fidelity procedure in Theorem 4.2.

We introduce the standard multivariate notation. Denote the countable set of "finitely supported" sequences of nonnegative integers by

$$\mathcal{F} := \{\nu = (\nu_1, \nu_2, \cdots) : \nu_j \in \mathbb{N}, \text{ and } \nu_j \neq 0 \text{ for only a finite number of } j\},$$

with $|\nu| := \sum_{j \geq 1} |\nu_j|$. For $\nu \in \mathcal{F}$ supported in $\{1, \cdots, J\}$, one defines the partial derivative in z

$$\partial^\nu u = \frac{\partial^{|\nu|} u}{\partial^{\nu_1} z_1 \cdots \partial^{\nu_J} z_J}, \tag{4.10}$$

and the multi-factorial $\nu! := \prod_{j \geq 1} \nu_j!$, where $0! := 0$.

Regularity of u^H. It is not hard to extend the sensitivity analysis in [50] to the multi-dimensional random variable case, see Remark 2.8 in [65] on the extension. [50] tells us that 1) the uncertainties from the initial data and collision kernel (under suitable assumptions) will eventually diminish and the solution will exponentially decay in time to the deterministic global equilibrium $\mathcal{M}$; 2) the regularity of the initial data in the random space is preserved at later time.

Let h_ε be a perturbed solution to the equation (4.1). If its random initial data satisfies for all r,

$$||h_\varepsilon^{in}(z)||_{H_{x;v}^{s,r} L_z^\infty} \leq C_I, \tag{4.11}$$

then at time $t > 0$,

$$||h_\varepsilon(t, z)||_{H_{x,v}^{s,r} L_z^\infty} \leq C_r\, e^{-\varepsilon\tau t},$$

Moreover, h is *analytic* in the random space, meaning that

$$||\partial^\nu h_\varepsilon(t, z)||_{H_{x,v}^s L_z^\infty} \leq C\, e^{-\varepsilon\tau t}(|\nu|!), \qquad \text{for any } \nu \geq 0. \tag{4.12}$$

See Appendix B for a discussion on the linearized Boltzmann case.

We now use a weaker version by letting $s = 1$ in (4.11) and (4.12). If the initial data $||h_\varepsilon^{in}(z)||_{H_{x;v}^{1,r} L_z^\infty} \leq C_I$, then

$$||\partial^\nu h_\varepsilon(t, z)||_{L_{x,v}^2 L_z^\infty} \leq ||\partial^\nu h_\varepsilon(t, z)||_{H_{x,v}^1 L_z^\infty} \leq C\, e^{-\varepsilon\tilde{\tau} t}.$$

By the definition of u containing perturbed macroscopic quantities (see [9, section 2.2.4]) by the Cauchy-Schwarz inequality, one easily gets

$$||\partial^\nu u(t,z)||_{L^2_x L^\infty_z} \lesssim ||\partial^\nu h_\varepsilon(t,z)||_{L^2_{x,v} L^\infty_z} \leq C\, e^{-\varepsilon \tau t}, \tag{4.13}$$

for $|\nu| \leq r$, where C and $\tau > 0$ are all generic constants independent of ε. We assume that the high-fidelity solution u^H follows a similar behavior as the analytic solution u (computed from h_ε) to the Boltzmann equation, with an error that depends on the numerical scheme used in the high-fidelity model. If the initial distribution of the high-fidelity model satisfies

$$||h_\varepsilon^{in}(z)||_{H^{1,r}_{x,v} L^\infty_z} \leq C_I, \tag{4.14}$$

then for a fixed time $t > 0$ and $|\nu| \leq r$,

$$\sup_{z \in I_z} ||\partial^\nu u^H(t,z)||^H \leq C'\, e^{-\varepsilon \tau t} + \xi, \tag{4.15}$$

where ξ depends on the order and discretization parameters Δt, Δx, Δv used in the high-fidelity solver. Thus for all $z \in I_z$, one gets

$$||\partial^\nu u^H(t,z)||^H \leq C'\, e^{-\varepsilon \tau t} + \xi.$$

Note that C_I, C, C' and τ in the inequalities (4.11)–(4.15) are all positive generic constants independent of ε.

We make the following assumption on the random collision kernel:

Assumption 1. *Assume the collision kernel take the form*

$$B(|v - v^*|, \cos\theta, z) = \Phi(|v - v^*|)b(\cos\theta, z),\ \Phi(|v - v^*|) = C|v - v^*|^m,\ m \in [0,1], \tag{4.16}$$

and $(\psi_j)_{j \geq 1}$ be an affine representer (see definition in [15]) of the cross section b, that is,

$$b(\eta, z) = \bar{b}(\eta) + \sum_{j \geq 1} z_j \psi_j(\eta), \qquad z := (z_j)_{j \geq 1}, \quad \eta = \cos\theta, \tag{4.17}$$

where the sequence $\left(||\psi_j||_{L^\infty(\eta)}\right)_{j \geq 1} \in \ell^p$ for $0 < p < 1$ (see [15]). One also assumes that

$$|b(\eta, z)| \leq C_0, \quad |\partial_\eta b(\eta, z)| \leq C_1, \quad |\partial^\nu b(\eta, z)| \leq C_2, \tag{4.18}$$

for all $\eta \in [-1, 1]$, $|\nu| \leq r$. Here C, C_0, C_1, C_2 are all positive constants.

The above (4.16)–(4.18) extend the conditions in [50, Theorem 4.4 (ii)] from one-dimensional random space to the multi-dimensional random space.

An upper bound for N-width. We can now utilize the result in [15, Corollary 3.11], which works for general parametric PDEs. Define the norm

$$||u - v||_{L^\infty(I_z, V)} := \sup_{z \in I_z} ||u(z) - v(z)||_V,$$

where V is the physical space considered, in our case $V = L_x^2$. Under **Assumption 1**, by the analyticity of u^H given in (4.15), for $z \in I_z$, one obtains

$$\left\|u^H - \sum_{\nu \in \Lambda_N} w_\nu P_\nu \right\|_{L^\infty(I_z, V)} \leq C(N+1)^{-q}, \qquad q = \frac{1}{p} - 1, \qquad (4.19)$$

where $(P_k)_{k \geq 0}$ is the sequence of renormalized Legendre polynomials on $[-1, 1]$, Λ_N is the set of indices that corresponds to the N largest $\|w_\nu\|_V$, and the constant $C := \left\|(\|w_\nu\|_V)_{\nu \in \mathcal{F}}\right\|_{\ell^p} < \infty$. By (4.15), one formally gets $C = c\, e^{-\varepsilon \tau t} + \xi$.

Recall (4.8) and the definition $d^H(v, V_N) := \min_{w \in V_N} \|v - w\|_V$ (refer to [15, equation (8.3)]), the best achievable error in $L^\infty(I_z, V)$ is described by the N-width of the solution manifold $\mathcal{M} = u^H(I_z)$:

$$d_N(\mathcal{M})_V = \inf_{\dim(V_N) = N} \sup_{v \in \mathcal{M}} \min_{w \in V_N} \|v - w\|_V. \qquad (4.20)$$

One can use the polynomial approximation bound in $L^\infty(I_z, V)$ to estimate an upper bound for $d_N(\mathcal{M})_V$, as studied in [15, Section 4]. Let the N-dimensional subspace

$$V_N := \operatorname{span}\{c_\nu : \nu \in \Lambda_N\}.$$

For $N \geq 1$, one observes that

$$d_N(\mathcal{M})_V \leq \sup_{v \in \mathcal{M}} \min_{w \in V_N} \|v - w\|_V = \sup_{z \in I_z} \min_{w \in V_N} \|u^H(z) - w\|_V$$

$$\leq \|u^H - \sum_{\nu \in \Lambda_N} w_\nu P_\nu \|_{L^\infty(I_z, V)} \leq C(N+1)^{-q}, \qquad q = \frac{1}{p} - 1. \qquad (4.21)$$

where $\sum_{\nu \in \Lambda_N} w_\nu P_\nu$ is the truncated Legendre expansion and (4.19) is used in the last inequality.

We now conclude with our main result on convergence analysis:

Theorem 4.2. *If the assumptions for the random initial data, random collision kernel, namely (4.14) and **Assumption 1** are satisfied, for fixed time $t > 0$ and fixed numerical discretization parameters Δt, Δx and Δv, then for all $z \in I_z$,*

$$\left\|u^H(t, z) - u^B(t, z)\right\|^H \leq C_1 \frac{e^{-\frac{\varepsilon \tau t}{2}}}{(N/2 + 1)^{q/2}} + C_2\, e^{-\varepsilon \tau t} + \chi + \left\|\sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\mathbf{Q}\,\mathbf{f}^L(z)\right\|, \qquad (4.22)$$

*where N is the size of the subspace γ_N in Algorithm 1, and q is given in (4.19) with p depending on the ℓ^p-summability assumption of $(\psi_j)_{j \geq 1}$, $C_1 = \mathcal{O}\left(\frac{1}{\delta_1 - \delta_2 \varepsilon}\right)$, C_2 and τ are constants that depend on the initial data u^{in} and **Assumption 1** on the collision kernel. δ_1, δ_2 are all sufficiently small with $0 < \delta_1 - \delta_2 \varepsilon < 1$. Definitions of $\mathbf{G}^L$, $\mathbf{G}^H$, $\mathbf{Q}$ and f^L are given below (4.5). χ is associated to the order and discretization mesh in the high-fidelity solver. ε is the Knudsen number in the Boltzmann equation (2.1).*

Proof. According to the inequalities (4.4), (4.6), (4.21), (4.15) and Theorem 4.1, one gets for all $z \in I_z$,

$$\left\|u^H(t, z) - u^B(t, z)\right\|^H \leq C\sqrt{d_{N/2}(u^H(I_z))} + c\,\|u^H(z)\|^H + \left\|\sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\mathbf{Q}\,\mathbf{f}^L(z)\right\|$$

$$\leq C_1 \frac{\sqrt{e^{-\varepsilon \tau t} + \xi}}{(N/2 + 1)^{q/2}} + C_2 \, e^{-\varepsilon \tau t} + \xi + \left\| \sqrt{\mathbf{G}^H} (\mathbf{G}^L)^{-1} \mathbf{Q} \, \mathbf{f}^L(z) \right\|$$

$$\leq C_1 \frac{e^{-\frac{\varepsilon \tau t}{2}}}{(N/2 + 1)^{q/2}} + C_2 \, e^{-\varepsilon \tau t} + \chi + \left\| \sqrt{\mathbf{G}^H} (\mathbf{G}^L)^{-1} \mathbf{Q} \, \mathbf{f}^L(z) \right\|,$$

here C, C_1, C_2, τ are all generic positive constants independent of ε, $C_2 = c \, C'$ which is small, and $\chi = C_1 \sqrt{\xi} + \xi$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Theorem 4.2 indicates that the error between the bi- and high-fidelity solutions decays algebraically with respect to the number of high-fidelity runs N. The convergence rate $q/2$ is independent of the dimension of the random space and the regularity of the initial data; it only relates to the ℓ^p summability of the affine representer $(\psi_j)_{j \geq 1}$ in **Assumption 1**.

Remark 4.3. *We make the following remarks:*

1. *The estimate in Theorem 4.2 may not be sharp. Deriving a sharper estimate requires a better understanding on the role of the Knudsen number ε in the accuracy analysis.*

2. *It is not our goal of the current work to establish stability and error analysis for the deterministic AP method for the multiscale Boltzmann equation in [26], which is difficult due to the penalization used in the scheme thus has not been studied to our best knowledge. Thus deriving (4.15) from (4.13) rigorously remains challenging.*

We hope to report more results regarding the above two issues in future researches.

5 Numerical Tests

To examine the performance of the method, we shall compute numerical errors in the following way: we choose a fixed set of points $\{\hat{z}_i\}_{i=1}^n \subset I_z$ that is independent of the point sets Γ, and evaluate the following error between the bi-fidelity and high fidelity solutions at a final time t:

$$\|u^H(t) - u^B(t)\|_{L^2(D \times I_z)} \approx \frac{1}{n} \sum_{i=1}^n \|u^H(\hat{z}_i, t) - u^B(\hat{z}_i, t)\|_{L^2(D)}, \qquad (5.1)$$

where $\|\cdot\|_{L^2(D)}$ is the L^2 norm in the physical domain $D = \Omega \times \mathbb{R}^2$. The error can be considered as an approximation to the average L^2 error in the whole space of $D \times I_z$.

Since our goal is to numerically solve the multiscale Boltzmann equation with random inputs, we solve the Boltzmann equation (2.1) as the high-fidelity AP solver discussed in Section 3.2. We assume the random collision kernel in the form of

$$B(v, v_*, \sigma, z) = b(z)|v - v_*|^\lambda, \qquad (5.2)$$

and consider Maxwell molecules, i.e., $\lambda = 0$ in (5.2). The low-fidelity model is chosen as the Euler equation solved by the forward Euler in time and second-order MUSCL scheme in space, by using the same spatial and temporal resolutions as the

high-fidelity model, but with a different number of quadrature points N_v^l in velocity discretization.

In all the examples, the spatial domain is chosen to be $[0, 1]$ with N_x grid points, and periodic boundary condition is assumed except for the shock tube tests. The velocity domain is chosen as $[-L_v, L_v]^2$ with $L_v = 8.4$ and N_v grid points in each dimension. Without loss of generality, the d-dimensional random variable $\mathbf{z}$ is assumed to follow the uniform distribution on $[-1, 1]^d$. The training set Γ is chosen to be $M = 1000$ random samples of $\mathbf{z}$. We examine the error of bi-fidelity approximation with respect to the number of high-fidelity runs by computing the norm defined in (5.1) (evaluated over an independent set of $n = 1000$ Monte Carlo samples). All numerical experiments are conducted by MATLAB R2018b on macOS Mojave system with 2.4 GHz Intel Core i5 processor and 8GB DDR3 memory),

5.1 A double-peak initial data test

We first consider the following initial data to mimic the Karhunen-Loeve expansion of the random field:

$$
\begin{cases}
\rho_0(x, \mathbf{z}^\rho) = \dfrac{1}{3}\left(2 + \sin(2\pi x) + 0.2\sum_{k=1}^{d_1}\sin[2\pi(k+1)x]\,\dfrac{z_k^\rho}{2k}\right), \\[2mm]
\mathbf{u}_0 = (0.2, 0), \\[2mm]
T_0(x, \mathbf{z}^T) = \dfrac{1}{4}\left(3 + \cos(2\pi x) + 0.2\sum_{k=1}^{d_1}\cos[2\pi(k+1)x]\,\dfrac{z_k^T}{2k}\right), \\[2mm]
f_0(x, \mathbf{v}, \mathbf{z}) = \dfrac{\rho_0}{4\pi T_0}\left(\exp(-\dfrac{|\mathbf{v} - \mathbf{u_0}|^2}{2T_0}) + \exp(-\dfrac{|\mathbf{v} + \mathbf{u_0}|^2}{2T_0})\right).
\end{cases}
\tag{5.3}
$$

The uncertain collisional cross section is given by

$$
b(z) = 1 + 0.5 z_1^b,
\tag{5.4}
$$

Here $\mathbf{z}^\rho = \left(z_1^\rho, \cdots, z_{d_1}^\rho\right)$, $\mathbf{z}^T = \left(z_1^T, \cdots, z_{d_1}^T\right)$, and $\mathbf{z}^b = z_1^b$ represent the random variables in the collision kernel, initial density and temperature. Let the initial distribution f_0 follow a double-peak non-equilibrium initial data [26]. Set $d_1 = 7$, thus this is a $d = 15$ dimensional problem in the random space. We use the Boltzmann equation as the high-fidelity model and the Euler system as the low-fidelity model, set $\Delta x = 0.01$, $\Delta t = 8 \times 10^{-4}$ (in both the high- and low-fidelity models), $N_v^h = 16$, and the final time $t = 0.1$.

In Figure 1, we consider the fluid regime with $\varepsilon = 10^{-4}$. This figure shows the mean L^2 errors of ρ, u_1, T between the high- and bi-fidelity solutions with different quadrature points in velocity space. Here $\mathbf{u_1}$ in the figures below stands for the first component of the two-dimensional bulk velocity $\mathbf{u}$. It is clear that the error decays fast with the number of high-fidelity runs. In addition, when N_v^l increases, the error between the high- and bi-fidelity solutions decreases. This is expected because the Euler equation solved by more quadrature points in velocity space can capture more information about the high-fidelity model.

In Figure 2, fluid regime is considered and we vary ε from $\varepsilon = 10^{-2}$ to $\varepsilon = 10^{-4}$. The Euler equation is chosen as the low-fidelity model, solved by the same forward Euler in time and second-order MUSCL scheme in space, and the same spatial and temporal meshes as the Boltzmann equation in the high-fidelity model, and with $N_v^l = 8$ velocity quadrature points. One observes that the smaller ε is, the lower level the errors saturate. This is expected, because when the Knudsen number ε approaches to zero, the Euler equation as the low-fidelity model commits less modeling error and can capture more information of the high-fidelity model.

In Figure 3, we investigate the performance of the bi-fidelity approximation for the kinetic regime with $\varepsilon = 1$. Fast convergence of the mean L^2 errors with respect to the number of high-fidelity runs is observed. Even though ε is relatively large compared to the previous two tests, a satisfactory accuracy in characterizing the behaviors of the solution in the random space is achieved in both cases: $N_v^l = 8$ and $N_v^l = 4$; and the errors with $N_v^l = 8$ is smaller than that of $N_v^l = 4$. On the right column of Figure 3, we plot the high-, low- and the corresponding bi-fidelity solutions (with $r = 20$, $N_v^l = 8$) for a particular sample point z. One observes that the high- and bi-fidelity solutions match quite well, whereas the low-fidelity solutions appear inaccurate at some spatial points. This example seems to indicate that although in the kinetic regime, the fluid description breaks down in the *physical space*, the bi-fidelity solution can still capture important variations of the high-fidelity model (Boltzman equation) in the *random space*.
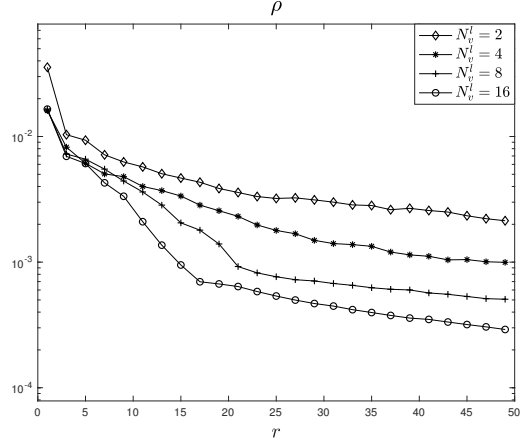
(a)
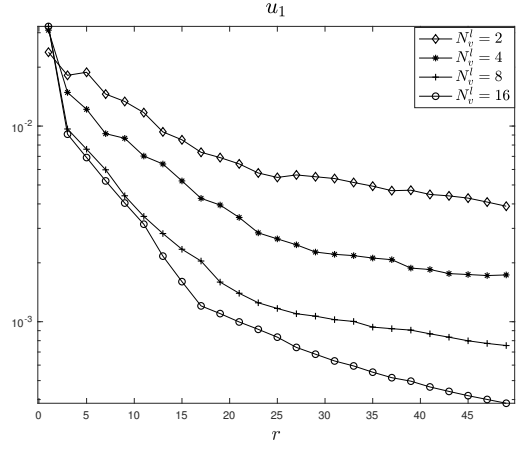
(b)
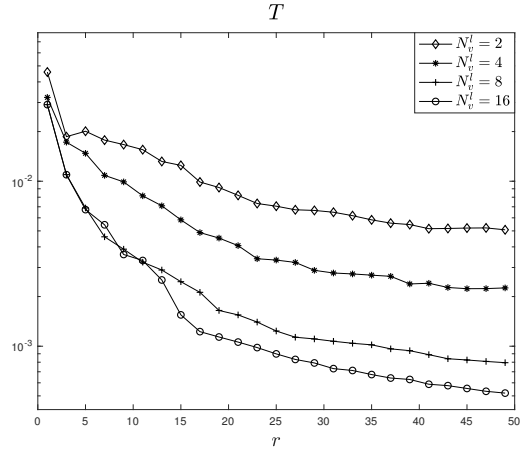
(c)

Figure 1: The mean L^2 error of the bi-fidelity approximation of ρ, u_1, T with respect to the number of high-fidelity simulation runs, based on the low-fidelity model with different N_v^l.
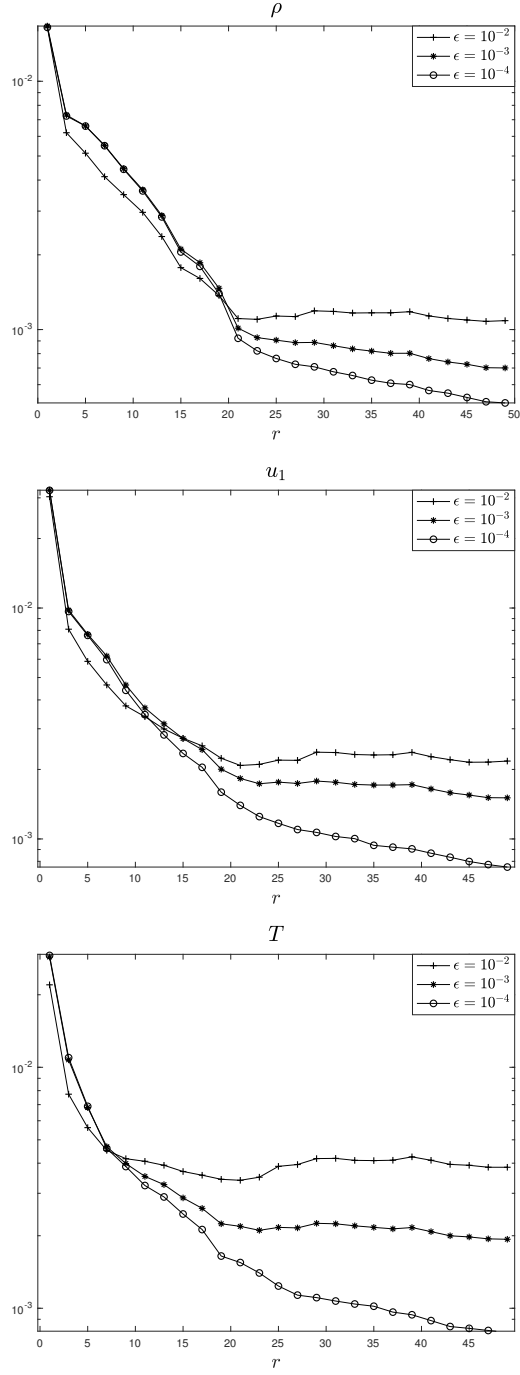
Figure 2: The mean L^2 error of the bi-fidelity approximation of ρ, u_1, T with respect to the number of high-fidelity simulation runs, based on the low-fidelity model with $N_v^l = 8$ for different ε.
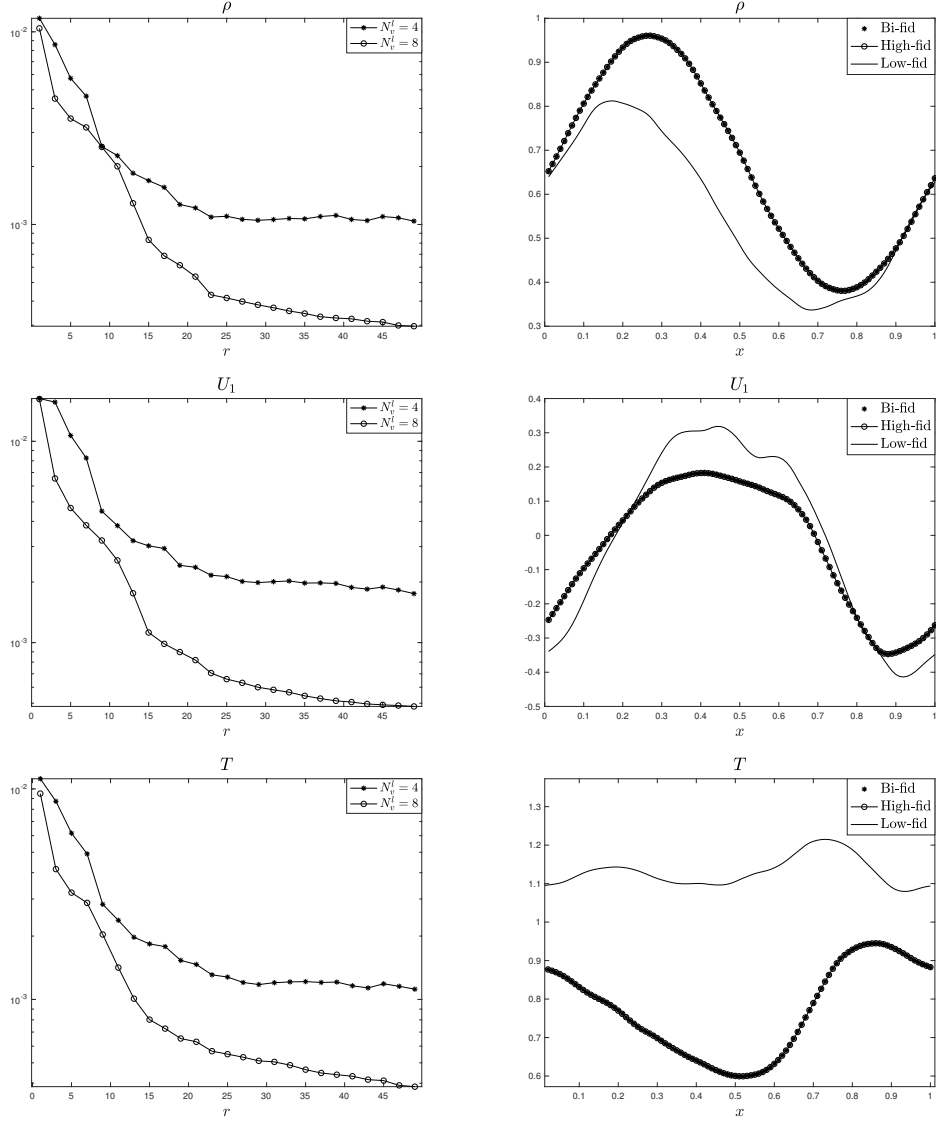
Figure 3: (Left) The mean L^2 error of the bi-fidelity approximation of ρ, u_1, T with respect to the number of high-fidelity simulation runs, based on the low-fidelity model with different N_v^l. (Right) Comparison of the low-fidelity solution ($N_v^l = 8$), high-fidelity solutions ($N_v^l = 16$), and the corresponding bi-fidelity approximations with $r = 20$ for a fixed z.

5.2 Sod shock tube test

We next consider a more challenging problem where the initial data is discontinuous. Assume the random collision kernel in the form

$$b(\mathbf{z}^b) = 1 + 0.5 \sum_{k=1}^{d_1+1} \frac{z_k^b}{2k},$$

and the random initial distribution

$$f^0(x, \mathbf{v}, \mathbf{z}) = \frac{\rho^0}{2\pi T^0} e^{-\frac{|\mathbf{v}-u^0|^2}{2T^0}},$$

where the initial data for ρ^0, u^0 and T^0 is given by

$$\begin{cases} \rho_l = 1, \quad u_l = (0,0), \quad T_l(\mathbf{z}^T) = 1 + 0.4 \sum_{k=1}^{d_1} \frac{z_k^T}{2k}, \quad x \le 0.5, \\ \\ \rho_r = \frac{1}{8}, \quad u_r = (0,0), \quad T_r(\mathbf{z}^T) = \frac{1}{8}(1 + 0.4 \sum_{k=1}^{d_1} \frac{z_k^T}{2k}), \quad x > 0.5. \end{cases}$$

Here $\mathbf{z}^b = \left(z_1^b, \cdots, z_{d_1+1}^b\right)$ and $\mathbf{z}^T = \left(z_1^T, \cdots, z_{d_1}^T\right)$ represent the random variables in the collision kernel and initial temperature. Set $d_1 = 7$, then the total dimension d of the random space is 15. We use the Boltzmann equation as the high-fidelity model, and solve it by $\Delta x = 0.01$, $\Delta t = 8 \times 10^{-4}$, and $N_v^h = 24$, until the final time $t = 0.15$. We shall employ the Euler equation as the low-fidelity model, and solve it with the same spacial and temporal resolution with the high-fidelity model but with $N_v^l = 12$. We consider the fluid regime with $\varepsilon = 10^{-4}$ in this test.

From the left column of Figure 4, we see a fast convergence of L^2 errors between the high- and bi-fidelity solutions. With only 10 high-fidelity runs, the bi-fidelity approximation can reach an accuracy level $\mathcal{O}(10^{-3})$ for a 15-dimensional problem in random space, while the low-fidelity approximation is quite poor with an accuracy level $\mathcal{O}(10^{-1})$. To further illustrate the performance of our bi-fidelity method, we compared the high-, low- and the corresponding bi-fidelity solutions (with $r = 10$) for a particular sample point z. One observes that the high- and bi-fidelity solutions match really well, whereas the low-fidelity solutions seem to be quite inaccurate at some points in the spatial domain. Even in this case, the bi-fidelity solutions can approximate the high-fidelity solutions very well.

Figure 5 shows clearly that the mean and standard deviation of the bi-fidelity approximation of ρ, u_1 and T agree well with the high-fidelity solutions by using only 10 high-fidelity runs. The result is a bit surprising yet reasonable, suggesting that even though the Euler model may be inaccurate in the physical space, it still can capture the behaviors and characteristics of the solution to the Boltzmann equation in the random space. Moreover, since the high-fidelity model (Boltzmann) with $N_v^h = 24$ costs approximately 43 times of the low-fidelity solver (Euler) with $N_v^l = 12$ (the former takes 30.6 seconds, the latter takes 0.7 seconds for one single run; a significant speedup is quite noticeable in this case.
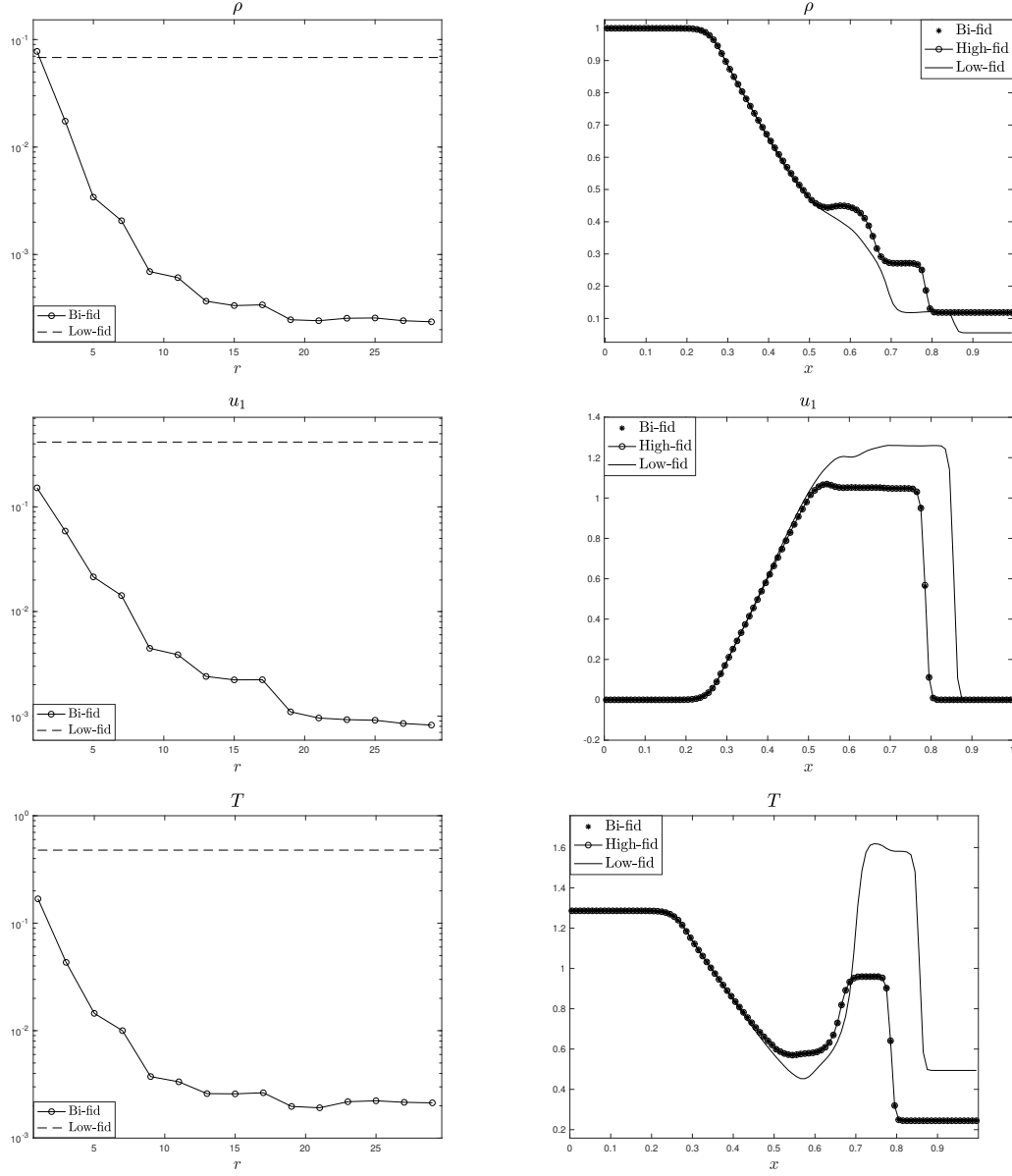
22

Figure 4: (Left) The mean L^2 errors between high-fidelity and low- or bi-fidelity solutions with respect to the number of high-fidelity runs; (Right) Comparison of the low-fidelity solution ($N_v^l = 12$), high-fidelity solutions ($N_v^l = 24$), and the corresponding bi-fidelity approximations $r = 10$ for a fixed z.
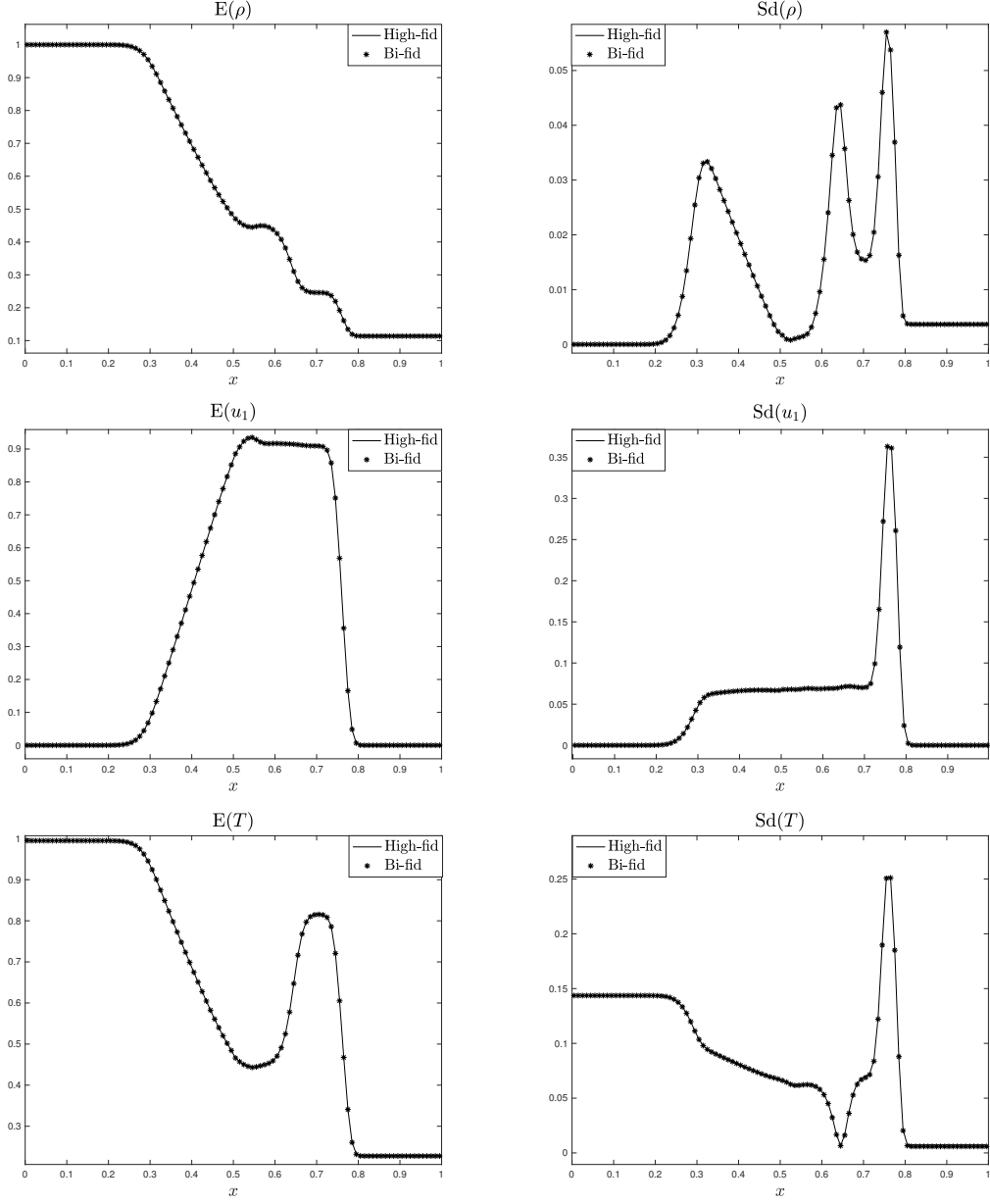
Figure 5: Mean and standard deviation of ρ, u_1, T of high-fidelity solutions and bi-fidelity solutions with $r = 10$.

5.3 Mixed regime test

The next test we shall consider is more challenging than the previous two tests. Because various scales are involved, good accuracy of the AP scheme for the Boltzmann equation is required for all ranges of ε. We consider a mixed regime with the

Knudsen number ε varying in space show in Figure. 6 and given by

$$\varepsilon(x) = 10^{-3} + \frac{1}{2}\left[\tanh\left(1 - \frac{11}{2}(x - 0.5)\right) + \tanh\left(1 + \frac{11}{2}(x - 0.5)\right)\right]. \qquad (5.5)$$

The random initial data and collision kernel are given by (5.3) and (5.4). The total dimension of the random space is $d = 15$.
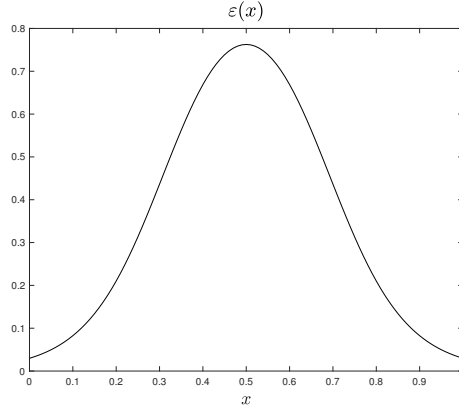

Figure 6: The distribution of $\varepsilon(x)$ in (5.5).

All the numerical parameters used in temporal and spatial discretizations are the same as that in Section 5.1. We solve the Boltzmann equation (2.1) for the high-fidelity solution with $N_v^h = 24$, and the Euler system for the low-fidelity solution with $N_v^l = 8$.

From the left column of Figure 7, we observe a fast convergence of L^2 errors between the high- and bi-fidelity solutions, where they saturate quickly when r reaches about 25. It is worth noting that the dotted lines that represent the errors between the high- and low-fidelity solutions are much larger $\mathcal{O}(10^{-1})$ compared to that between the high- and bi-fidelity solutions. This indicates that even though the low-fidelity solutions alone are relatively not accurate in the spatial domain, it might be still able to behave similarly in the random space, therefore the resulted bi-fidelity approximation based on a small number of high-fidelity runs (say $r = 25$) can reach a reasonable accuracy level up to $\mathcal{O}(10^{-3})$.

The right column of Figure 7 shows the high-, low- and bi-fidelity solutions at a randomly chosen sample point z. One can see that the high- and bi-fidelity solutions match really well, whereas the low-fidelity solutions are not accurate. In addition, with $N = 1000$ low-fidelity runs of the Euler model, together with only 25 runs of the AP solver to the Boltzmann model, one can get the bi-fidelity solutions which are able to capture behavior of the solutions to the Boltzmann equation in the random space, up to an accuracy of 10^{-3}; on the other hand, using the low-fidelity model (Euler equation) alone can not achieve this result, especially under the multiple scalings where ε ranges from 10^{-3} to 1 (since the errors between macroscopic quantities calculated from the Boltzmann and Euler equation deteriorate when ε becomes large).

25

This observation certainly highlights the merits of our bi-fidelity method. Figure 8 presents the mean and standard deviation of ρ, u_1, T by using 25 high-fidelity runs. One can see that the high- and bi-fidelity solutions match well, indicating that the bi-fidelity solutions have captured well the characteristics of the macroscopic quantities in the random space.

Once we construct the bi-fidelity model, the online computational cost can be significantly reduced. In this example, the high-fidelity model (Boltzmann) with $N_v^h = 16$ takes about 50 times computation time of the low-fidelity model (the former takes 12.3 seconds, while the latter takes 0.23 seconds for a single run). Since the dominant cost of the bi-fidelity reconstruction for a high-fidelity solution lies in the corresponding low-fidelity run, a significant amount of computational cost is saved in our method.
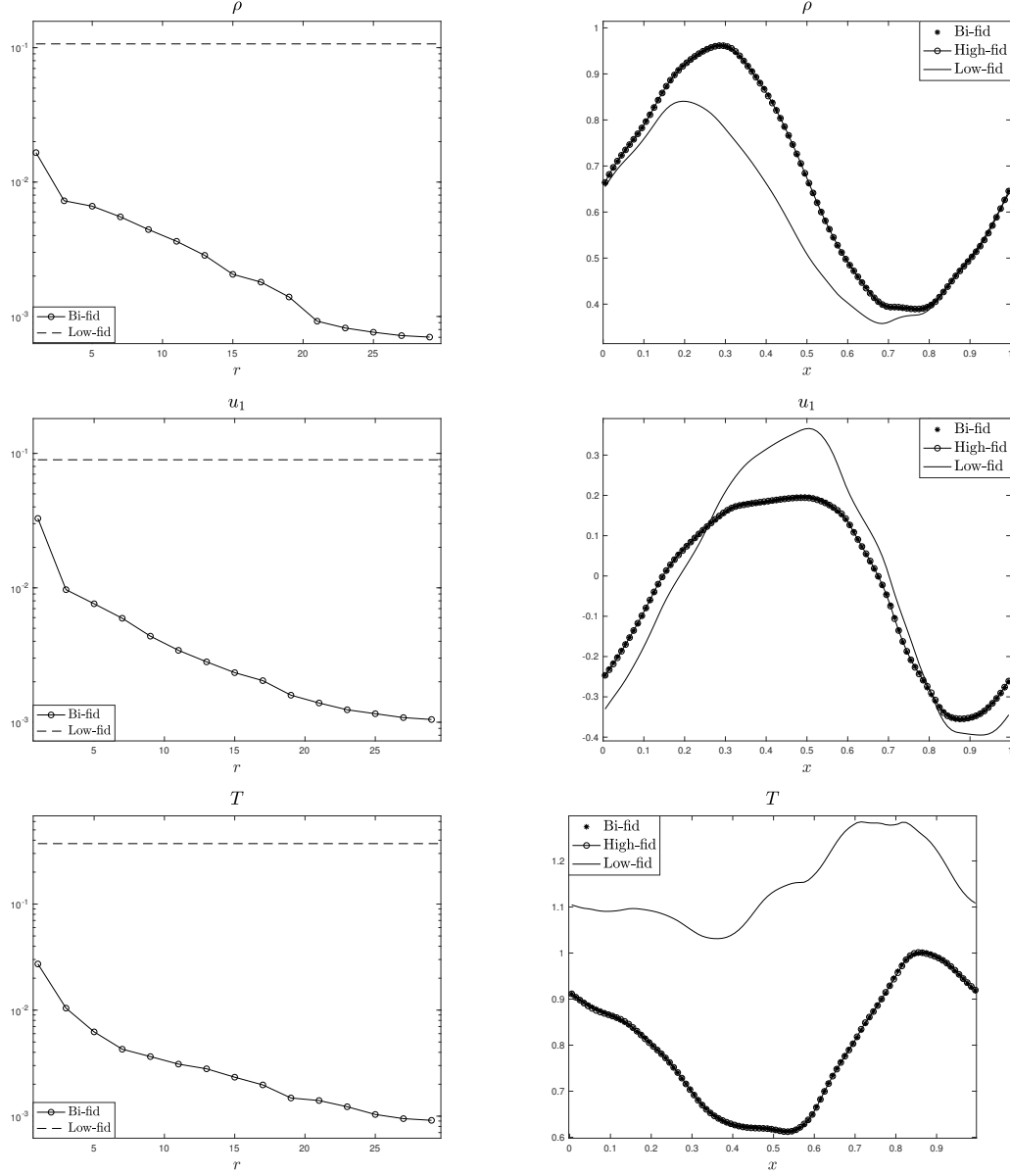
Figure 7: (Left) The mean L^2 errors between high-fidelity and low- or bi-fidelity solutions with respect to the number of high-fidelity run; (Right) Comparison of the low-fidelity solution ($N_v^l = 8$), high-fidelity solutions ($N_v^l = 16$), and the corresponding bi-fidelity approximations $r = 25$ for a fixed z.
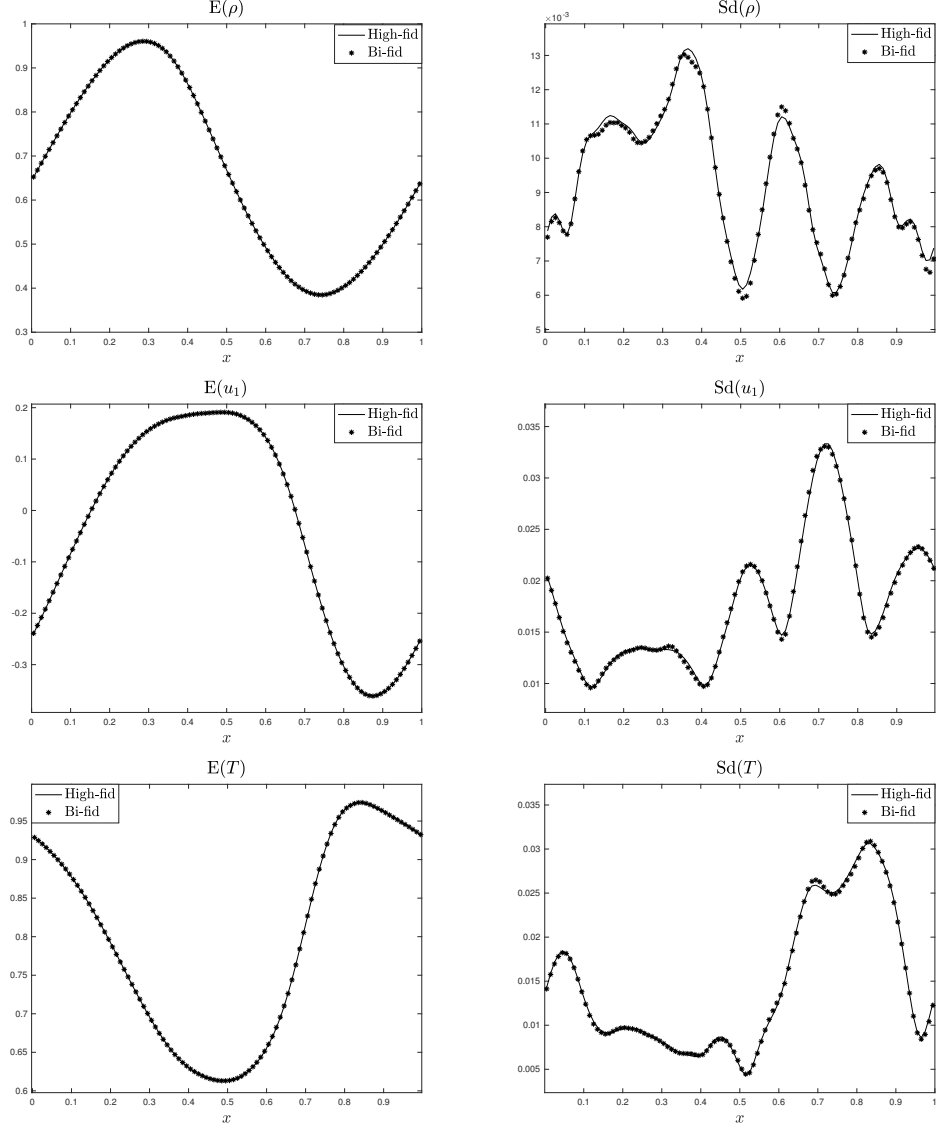
Figure 8: The mean and standard deviation of ρ, u_1, T of high-fidelity solutions and bi-fidelity solutions with $r = 25$.

6   Conclusion

In this work, we study the multiscale Boltzmann equation with multi-dimensional random parameters by a bi-fidelity collocation method [53, 72]. By choosing the low-fidelity solution as the solution of the corresponding first order macro-model – the compressible Euler equations with a consistent initial data, our bi-fidelity approximation can capture well the variations of macroscopic quantities computed from the high-fidelity AP solver of the Boltzmann equation with multiple scales, at a much

reduced computational cost and memory footprint. An error analysis developed in [53] has been extended by incorporating the knowledge of the regularity of our high-fidelity and low-fidelity solutions. The computational accuracy and efficiency are demonstrated in various numerical examples and holds promise to accelerate the computation for more complex problems in multiscale kinetic equations with uncertainty.

Appendices

A   The Lemma proof

Lemma 1.1. *Let z_*^H, z_*^L be the maximizers of (4.7a) and (4.7b), respectively. One assumes that*

1. $\exists$ positive constants K^H, K^L, C_1, C_2 such that

$$\left| ||u^L(z_*^H)||^L - K^H ||u^H(z_*^H)||^H \right| \leq C_1\,\varepsilon,$$

$$\left| ||u^H(z_*^L)||^H - K^L ||u^L(z_*^L)||^L \right| \leq C_2\,\varepsilon,$$

2. $\exists \eta^L \leq 1$ and $\eta^H \leq 1$ such that

$$||P_{U_n^L} u^L(z_*^H)||^L \leq \eta^H ||P_{U_n^H} u^H(z_*^H)||^H,$$

$$||P_{U_n^H} u^H(z_*^L)||^H \leq \eta^L ||P_{U_n^L} u^L(z_*^L)||^L.$$

3. $\exists$ constants A^H and A^L satisfying

$$1 \leq A^H < \frac{\eta^H}{\max\{0, \eta^H - K^H\}}, \qquad 1 \leq A^L < \frac{\eta^L}{\max\{0, \eta^L - K^L\}}$$

such that

$$||u^H(z_*^H)||^H \leq A^H d^H(u^H(z_*^H), U_n^H),$$

$$||u^L(z_*^L)||^L \leq A^L d^L(u^L(z_*^L), U_n^L).$$

Then

$$d^H(u^H(z_*^L), U_n^H) \geq (\delta_1 - \delta_2\,\varepsilon)\, d^H(u^H(z_*^H), U_n^H),$$

with $0 < \delta_1 - \delta_2\,\varepsilon < 1$.

Proof. Using the above assumptions, one gets

$$
\begin{aligned}
d^L(u^L(z_*^H), U_n^L) &= ||u^L(z_*^H)||^L - ||P_{U_n^L} u^L(z_*^H)||^L \\
&\geq K^H ||u^H(z_*^H)||^H - C_1\,\varepsilon - ||P_{U_n^L} u^L(z_*^H)||^L \\
&\geq K^H ||u^H(z_*^H)||^H - C_1\,\varepsilon - \eta^H ||P_{U_n^H} u^H(z_*^H)||^H \\
&= \eta^H \left[||u^H(z_*^H)||^H - ||P_{U_n^H} u^H(z_*^H)||^H \right] - C_1\,\varepsilon + (K^H - \eta^H) ||u^H(z_*^H)||^H \\
&\geq \eta^H d^H(u^H(z_*^H), U_n^H) + C_1\,\varepsilon - \max\{0, \eta^H - K^H\} A^H d^H(u^H(z_*^H), U_n^H)
\end{aligned}
$$

29

$$= \left[\eta^H - \max\{0, \eta^H - K^H\}A^H\right] d^H(u^H(z_*^H), U_n^H) - C_1\,\varepsilon,$$

Similarly, we have

$$d^H(u^H(z_*^L), U_n^H) \geq \left[\eta^L - \max\{0, \eta^L - K^L\}A^L\right] d^L(u^L(z_*^L), U_n^L) - C_2\,\varepsilon.$$

Therefore,

$$
\begin{aligned}
d^H(u^H(z_*^L), U_n^H) &\geq \left[\eta^L - \max\{0, \eta^L - K^L\}A^L\right] d^L(u^L(z_*^L), U_n^L) - C_2\,\varepsilon \\
&\geq \left[\eta^L - \max\{0, \eta^L - K^L\}A^L\right] d^L(u^L(z_*^H), U_n^L) - C_2\,\varepsilon \\
&\geq \left[\eta^L - \max\{0, \eta^L - K^L\}A^L\right] \cdot \left[\eta^H - \max\{0, \eta^H - K^H\}A^H\right] d^H(u^H(z_*^H), U_n^H) \\
&\quad - \left\{\left[\eta^L - \max\{0, \eta^L - K^L\}A^L\right] C_1 + C_2\right\}\varepsilon \\
&\geq (\delta_1 - \delta_2\,\varepsilon)\, d^H(u^H(z_*^H), U_n^H),
\end{aligned}
$$

where we assume that

$$0 < \delta_1 = \left[\eta^L - \max\{0, \eta^L - K^L\}A^L\right] \cdot \left[\eta^H - \max\{0, \eta^H - K^H\}A^H\right] < 1,$$

and $\exists\,\delta_2$ such that $0 < \delta_1 - \delta_2\,\varepsilon < 1$ with

$$\left[\eta^L - \max\{0, \eta^L - K^L\}A^L\right] C_1 + C_2 \leq \delta_2\, d^H(u^H(z_*^H), U_n^H).$$

$$\square$$

B Proof of analyticity for the linearized equation

For simplicity, consider the linearized Boltzmann equation under the acoustic scaling. The perturbative solution satisfies

$$\partial_t h + v \cdot \nabla_x h = \frac{1}{\varepsilon}\mathcal{L}(h). \tag{B.1}$$

For simplicity, we write out the proof for one-dimensional random variable z. Assume the collision cross-section has the form $b(\cos\theta) = b_0(\cos\theta) + b_1(\cos\theta)\,z$, and we require $|b_1| \leq C_b$. Define the operator $\mathcal{T}_\varepsilon := \frac{1}{\varepsilon}\mathcal{L} - v \cdot \nabla_x$, and the notation $h_l := \partial_z^l h$ for all $l \geq 0$. Take ∂_z^l of the equation (B.1), then

$$\partial_t h_l = \mathcal{T}_\varepsilon(h_l) + \frac{1}{\varepsilon}\mathcal{L}_{b_1}(h_{l-1}),$$

where $\mathcal{L}_{b_1}$ is defined by substituting $b(\cos\theta)$ by $b_1(\cos\theta)$ in the linearized collision operator $\mathcal{L}$. Thus

$$\frac{d}{dt}\|h_l\|_{H^k}^2 = 2\langle\mathcal{T}_\varepsilon(h_l), h_l\rangle_{H^k} + \frac{2}{\varepsilon}l\,\langle\mathcal{L}_{b_1}(h_{l-1}), h_l\rangle_{H^k}.$$

By the hypocoercivity theory and $\mathcal{L}$ being a bounded operator,

$$\frac{d}{dt}\|h_l\|_{H^k}^2 \leq -\varepsilon C_0\,\|h_l\|_{H_\Lambda^k}^2 + \frac{C_1}{\varepsilon}l\,\|h_{l-1}\|_{H^k}\|h_l\|_{H^k},$$

where $\|h\|^2_{H^k_\Lambda} = \sum_{|j|+|l|=k} \left\| \frac{\partial^2}{\partial v_j \partial x_l} h \right\|^2_\Lambda$ and $\|h\|_\Lambda = \left\| h(1+|v|)^{\gamma/2} \right\|_{L^2_{x,v}}$, with γ shown in (2.4). Since $\|h_l\|_{H^k_\Lambda}$ controls $\|h_l\|_{H^k}$, then by setting $g_l = \|h_l\|_{H^k}$,

$$\frac{d}{dt} g_l \leq -\varepsilon C_0\, g_l + \frac{C_1}{\varepsilon} l\, g_{l-1}.$$

Using [47, Lemma 6] in a similar way, one gets

$$g_l(t) \leq e^{-\varepsilon C_0 t} \sum_{m=0}^{l} \frac{l!}{(l-m)!\, m!} \left(\frac{C_1}{\varepsilon} t \right)^m h_{l-m}(0).$$

Following [47, Theorem 5], if the initial data satisfies $\|\partial^l_z h_0(z)\|_{H^k} \leq R^l$, $\forall l \geq 0$, then at time t,

$$\|\partial^l_z h(z)\|_{H^k} \leq C e^{-\varepsilon C_0 t} \left(R + \frac{C_1}{\varepsilon} t \right)^l,$$

where C, C_0, C_1 are independent of l and ε. The convergence radius for h at any point z_0 is defined by

$$r(z_0) = \frac{1}{\limsup_{l\to\infty} \left(g_l(z_0) / l! \right)^{1/l}} = \infty,$$

which is independent of z_0, thus h (or f) is *analytic* in z.

References

[1] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Rev., 52 (2010), pp. 317–355.

[2] C. BARDOS, F. GOLSE, AND C. D. LEVERMORE, *Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation*, Comm. Pure Appl. Math., 46 (1993), pp. 667–753.

[3] ——, *Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation*, Comm. Pure Appl. Math., 46 (1993), pp. 667–753.

[4] C. BARDOS, F. GOLSE, AND D. LEVERMORE, *Fluid dynamic limits of kinetic equations. I. Formal derivations*, J. Statist. Phys., 63 (1991), pp. 323–344.

[5] M. BENNOUNE, M. LEMOU, AND L. MIEUSSENS, *Uniformly stable numerical schemes for the Boltzmann equation preserving the compressible Navier-Stokes asymptotics*, J. Comput. Phys., 227 (2008), pp. 3781–3803.

[6] M. BIERI AND C. SCHWAB, *Sparse high order FEM for elliptic sPDEs*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 1149–1170.

[7] G. A. BIRD, *Molecular gas dynamics and the direct simulation of gas flows*, vol. 42 of Oxford Engineering Science Series, The Clarendon Press, Oxford University Press, New York, 1995. Corrected reprint of the 1994 original, With 1 IBM-PC floppy disk (3.5 inch; DD), Oxford Science Publications.

[8] F. Bouchut, F. Golse, and M. Pulvirenti, *Kinetic equations and asymptotic theory*, Elsevier, 2000.

[9] M. Briant, *From the Boltzmann equation to the incompressible Navier-Stokes equations on the torus: a quantitative error estimate*, J. Differential Equations, 259 (2015), pp. 6072–6141.

[10] R. E. Caflisch, *The fluid dynamic limit of the nonlinear Boltzmann equation*, Comm. Pure Appl. Math., 33 (1980), pp. 651–666.

[11] C. Cercignani, *The Boltzmann equation in the whole space*, in The Boltzmann Equation and Its Applications, Springer, 1988, pp. 40–103.

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

[13] S. Chapman and T. G. Cowling, *The mathematical theory of non-uniform gases. An account of the kinetic theory of viscosity, thermal conduction and diffusion in gases*, Third edition, prepared in co-operation with D. Burnett, Cambridge University Press, London, 1970.

[14] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Computing and Visualization in Science, 14 (2011), p. 3.

[15] A. Cohen and R. DeVore, *Approximation of high-dimensional parametric PDEs*, Acta Numer., 24 (2015), pp. 1–159.

[16] F. Coron and B. Perthame, *Numerical passage from kinetic to fluid equations*, SIAM J. Numer. Anal., 28 (1991), pp. 26–42.

[17] E. S. Daus, S. Jin, and L. Liu, *Spectral convergence of the stochastic galerkin approximation to the boltzmann equation with multiple scales and large random perturbation in the collision kernel*, Kinetic and Related Models, 12 (2019).

[18] P. Degond, S. Jin, and L. Mieussens, *A smooth transition model between kinetic and hydrodynamic equations*, J. Comput. Phys., 209 (2005), pp. 665–694.

[19] R. DeVore, G. Petrova, and P. Wojtaszczyk, *Greedy algorithms for reduced bases in Banach spaces*, Constr. Approx., 37 (2013), pp. 455–466.

[20] G. Dimarco and L. Pareschi, *Exponential Runge-Kutta methods for stiff kinetic equations*, SIAM J. Numer. Anal., 49 (2011), pp. 2057–2077.

[21] ——, *Numerical methods for kinetic equations*, Acta Numerica, 23 (2014), pp. 369–520.

[22] ——, *Multi-scale control variate methods for uncertainty quantification in kinetic equations*, J. Comput. Phys., 388 (2019), pp. 63–89.

[23] G. Dimarco, L. Pareschi, and M. Zanella, *Uncertainty quantification for kinetic models in socio-economic and life sciences*, in Uncertainty quantification for hyperbolic and kinetic equations, vol. 14 of SEMA SIMAI Springer Ser., Springer, Cham, 2017, pp. 151–191.

[24] R. J. DiPerna and P.-L. Lions, *On the Cauchy problem for Boltzmann equations: global existence and weak stability*, Ann. of Math. (2), 130 (1989), pp. 321–366.

[25] M. S. Eldred, L. W. Ng, M. F. Barone, and S. P. Domino, *Multifidelity uncertainty quantification using spectral stochastic discrepancy models*, Handbook of Uncertainty Quantification, (2017), pp. 991–1036.

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

[27] I. M. Gamba, J. R. Haack, C. D. 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.

[28] I. M. Gamba, S. Jin, and L. Liu, *Micro-macro decomposition based asymptotic-preserving numerical schemes and numerical moments conservation for collisional nonlinear kinetic equations*, J. Comput. Phys., 382 (2019), pp. 264–290.

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

[30] M. B. Giles, *Multilevel Monte Carlo path simulation*, Operations Research, 56 (2008), pp. 607–617.

[31] M. D. Gunzburger, C. G. Webster, and G. Zhang, *Stochastic finite element methods for partial differential equations with random input data*, Acta Numer., 23 (2014), pp. 521–650.

[32] Y. Guo, J. Jang, and N. Jiang, *Local Hilbert expansion for the Boltzmann equation*, Kinet. Relat. Models, 2 (2009), pp. 205–214.

[33] J. Hampton, H. R. Fairbanks, A. Narayan, and A. Doostan, *Practical error bounds for a non-intrusive bi-fidelity approach to parametric/stochastic model reduction*, Journal of Computational Physics, 368 (2018), pp. 315–332.

[34] J. Hu and S. Jin, *Uncertainty quantification for kinetic equations*, in Uncertainty quantification for hyperbolic and kinetic equations, vol. 14 of SEMA SIMAI Springer Ser., Springer, Cham, 2017, pp. 193–229.

[35] J. Hu, S. Jin, and S. Ruiwen, *On stochastic galerkin approximation of the nonlinear boltzmann equation with uncertainty in the fluid regime*, preprint, (2019).

[36] J. D. Jakeman and S. G. Roberts, *Stochastic Galerkin and collocation methods for quantifying uncertainty in differential equations: a review*, ANZIAM J., 50 (2008), pp. C815–C830.

[37] N. Jiang, C. D. Levermore, and N. Masmoudi, *Remarks on the acoustic limit for the Boltzmann equation*, Comm. Partial Differential Equations, 35 (2010), pp. 1590–1609.

[38] S. Jin, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput., 21 (1999), pp. 441–454.

[39] ———, *Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review*, Riv. Math. Univ. Parma (N.S.), 3 (2012), pp. 177–216.

[40] S. Jin, L. Pareschi, and G. Toscani, *Uniformly accurate diffusive relaxation schemes for multiscale transport equations*, SIAM J. Numer. Anal., 38 (2000), pp. 913–936.

[41] S. Jin, D. Xiu, and X. Zhu, *Asymptotic-preserving methods for hyperbolic and transport equations with random inputs and diffusive scalings*, Journal of Computational Physics, 289 (2015), pp. 35–52.

[42] M. C. Kennedy and A. O'Hagan, *Predicting the output from a complex computer code when fast approximations are available*, Biometrika, 87 (2000), pp. 1–13.

[43] A. Klar, *An asymptotic preserving numerical scheme for kinetic equations in the low Mach number limit*, SIAM J. Numer. Anal., 36 (1999), pp. 1507–1527.

[44] M. Lemou and L. Mieussens, *A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit*, SIAM J. Sci. Comput., 31 (2008), pp. 334–368.

[45] M. E. Leo Wai-Tsun Ng, *Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation*, 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA, (2012), p. 1852.

[46] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser, 1992.

[47] Q. Li and L. Wang, *Uniform regularity for linear kinetic equations with random input based on hypocoercivity*, SIAM/ASA J. Uncertain. Quantif., 5 (2017), pp. 1193–1219.

[48] L. Liu, *Uniform spectral convergence of the stochastic Galerkin method for the linear semiconductor Boltzmann equation with random inputs and diffusive scaling*, Kinet. Relat. Models, 11 (2018), pp. 1139–1156.

[49] ———, *A stochastic asymptotic-preserving scheme for the bipolar semiconductor Boltzmann-Poisson system with random inputs and diffusive scalings*, J. Comput. Phys., 376 (2019), pp. 634–659.

[50] L. Liu and S. Jin, *Hypocoercivity based sensitivity analysis and spectral convergence of the stochastic Galerkin approximation to collisional kinetic equations with multiple scales and random inputs*, Multiscale Model. Simul., 16 (2018), pp. 1085–1114.

[51] T. P. Liu, *Solutions in the large for the equations of nonisentropic gas dynamics*, Indiana Univ. Math. J., 26 (1977), pp. 147–177.

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

[53] A. Narayan, C. Gittelson, and D. Xiu, *A stochastic collocation algorithm with multifidelity models*, SIAM J. Sci. Comput., 36 (2014), pp. A495–A521.

[54] A. Narayan and D. Xiu, *Stochastic collocation methods on unstructured grids in high dimensions via interpolation*, SIAM J. Sci. Comput., 34 (2012), pp. A1729–A1752.

[55] L. W.-T. Ng and M. Eldred, *Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation*, in 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA, 2012, p. 1852.

[56] F. Nobile, R. Tempone, and C. G. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2309–2345.

[57] P. S. Palar, T. Tsuchiya, and G. T. Parks, *Corrigendum to "Multi-fidelity non-intrusive polynomial chaos based on regression" [Comput. Methods Appl. Mech. Engrg. 305 (2016) 579–606] [MR3491948]*, Comput. Methods Appl. Mech. Engrg., 307 (2016), pp. 489–490.

[58] 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.

[59] B. Peherstorfer, T. Cui, Y. Marzouk, and K. Willcox, *Multifidelity importance sampling*, Computer Methods in Applied Mechanics and Engineering, 300 (2016), pp. 490–509.

[60] B. Peherstorfer, K. Willcox, and M. Gunzburger, *Optimal model management for multifidelity monte carlo estimation*, SIAM Journal on Scientic Computing, 38 (2016), pp. A3163–A3194.

[61] P. Perdikaris, M. Raissi, A. Damianou, N. Lawrence, and G. E. Karniadakis, *Nonlinear information fusion algorithms for data-efficient multifidelity modelling*, Proc. R. Soc. A, 473 (2017), p. 20160751.

[62] A. D. M. A. S. Ryan King, Jennifer Annoni, *Enabling predictive reduced order modeling of high-fidelity wind plant simulations with in-situ modal decomposition and basis interpolation*, Bulletin of the American Physical Society, 63.

[63] E. L. P. J. A. E. Ryan W. Skinner, Alireza Doostan and K. E. Jansen, *Reduced-basis multifidelity approach for efficient parametric study of naca airfoils*, AIAA Journal, 57, pp. 1481–1491.

[64] R. Shu, J. Hu, and S. Jin, *A stochastic galerkin method for the boltzmann equation with multi-dimensional random inputs using sparse wavelet bases*, Numerical Mathematics: Theory, Methods and Applications, 10 (2017), pp. 465–488.

[65] R. Shu and S. Jin, *Uniform regularity in the random space and spectral accuracy of the stochastic Galerkin method for a kinetic-fluid two-phase flow model with random initial inputs in the light particle regime*, ESAIM Math. Model. Numer. Anal., 52 (2018), pp. 1651–1678.

[66] R. Tuo, C. J. Wu, and D. Yu, *Surrogate modeling of computer experiments with different mesh densities*, Technometrics, 56 (2014), pp. 372–380.

[67] C. Villani, *A review of mathematical topics in collisional kinetic theory*, in Handbook of mathematical fluid dynamics, Vol. I, North-Holland, Amsterdam, 2002, pp. 71–305.

[68] D. Xiu, *Efficient collocational approach for parametric uncertainty analysis*, Commun. Comput. Phys., 2 (2007), pp. 293–309.

[69] L. Yan and T. Zhou, *An adaptive multi-fidelity pc-based ensemble kalman inversion for inverse problems*, Int. J. Uncertaitny Quantif., 9 (2019), pp. 205–220.

[70] L. Yan and T. Zhou, *Adaptive multi-fidelity polynomial chaos approach to Bayesian inference in inverse problems*, J. Comput. Phys., 381 (2019), pp. 110–128.

[71] X. Yang, G. Tartakovsky, and A. Tartakovsky, *Physics-informed kriging: A physics-informed Gaussian process regression method for data-model convergence*, arXiv preprint arXiv:1809.03461, (2018).

[72] X. Zhu, A. Narayan, and D. Xiu, *Computational aspects of stochastic collocation with multifidelity models*, SIAM/ASA J. Uncertain. Quantif., 2 (2014), pp. 444–463.

[73] X. Zhu and D. Xiu, *A multi-fidelity collocation method for time-dependent parameterized problems*, AIAA SciTech Forum, 19th AIAA Non-Deterministic Approaches Conference, (2017).