MULTISCALE MODEL. SIMUL.
Vol. 15, No. 1, pp. 157–183

AN ASYMPTOTIC-PRESERVING STOCHASTIC GALERKIN METHOD FOR THE SEMICONDUCTOR BOLTZMANN EQUATION WITH RANDOM INPUTS AND DIFFUSIVE SCALINGS*

SHI JIN† AND LIU LIU‡

Abstract. In this paper, we develop a generalized polynomial chaos approach based stochastic Galerkin (gPC-SG) method for the linear semiconductor Boltzmann equation with random inputs and diffusive scalings. The random inputs are due to uncertainties in the collision kernel or initial data. We study the regularity (uniform in the Knudsen number) of the solution in the random space and prove the spectral accuracy of the gPC-SG method. We then use the asymptotic-preserving framework for the deterministic counterpart developed in [S. Jin and L. Pareschi, *J. Comput. Phys.*, 161 (2000), pp. 312–330] to come up with the stochastic asymptotic-preserving gPC-SG method for the problem under study which is efficient in the diffusive regime. Numerical experiments are conducted to validate the accuracy and asymptotic properties of the method.

Key words. semiconductor Boltzmann equation, uncertainty quantification, diffusion limit, asymptotic preserving, random inputs, generalized polynomial chaos, spectral accuracy

AMS subject classifications. 35Q20, 65M70

DOI. 10.1137/15M1053463

1. Introduction.

We consider the linear semiconductor Boltzmann equation with random inputs. Here the random inputs arise in the collision kernel or initial data due to modeling or measurement errors, which are typical for kinetic equations that are often derived via mean-field limits from particle systems [2, 17]. In recent years there has been significant interest in uncertainty quantification for physical models that contain uncertain coefficients, but few works have concentrated on kinetic equations which are of practical importance in mesoscopic modeling of physical, biological, and social sciences. Here we mention the stochastic Galerkin method employed for the neutron transport equation with random scattering coefficients [1] and, more recently, for the nonlinear Boltzmann equation [6].

Another challenge in numerical approximations of kinetic and transport equations arises from the varying magnitude of the Knudsen number, which is the dimensionless mean free path measuring the ratio between the particle mean free path and a typical length scale. When the Knudsen number is small the equation becomes numerically stiff and thus demands prohibitive, Knudsen number dependent mesh sizes and time steps. To overcome this difficulty, asymptotic-preserving (AP) schemes, which mimic

*Received by the editors December 18, 2015; accepted for publication (in revised form) November 2, 2016; published electronically January 17, 2017.
http://www.siam.org/journals/mms/15-1/M105346.html
Funding: This work was partially supported by NSF grants DMS-1522184 and DMS-1107291: RNMS KI-Net; by NSFC grant 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.

†Institute of Natural Sciences, Department of Mathematics, MOE-LSEC and SHL-MAC, Shanghai Jiao Tong University, Shanghai 200240, China, and Department of Mathematics, University of Wisconsin–Madison, Madison, WI 53706 (sjin@wisc.edu).
‡Department of Mathematics, University of Wisconsin–Madison, Madison, WI 53706 (lliu@math.wisc.edu).

the asymptotic transition from the kinetic equations to the hydrodynamic or diffusion limit, combined with efficient time integrators, have proved to be very efficient to handle small or multiple scales in kinetic or hyperbolic problems; see [7, 8]. For AP schemes for linear (deterministic) kinetic equations similar to the equation under study, see [14, 11, 10, 13]. For linear transport equations with diffusive scales and random inputs, stochastic asymptotic-preserving (s-AP) schemes were recently introduced in [12]. An s-AP scheme allows the use of mesh sizes, time steps, and the number of terms in the orthogonal polynomial expansions *independent of* the Knudsen number, yet can still capture the solution of the limiting, macroscopic equations.

In this paper, we aim to develop an s-AP scheme for the linear semiconductor Boltzmann equation with random inputs and diffusive scalings. Our method is based on the generalized polynomial chaos approach in the stochastic Galerkin (gPC-SG) framework [4, 15, 20]. The advantage of the gPC-SG method over the classical Monte Carlo method is that the former enjoys a spectral accuracy in the random space—provided sufficient regularity of the solution—while the latter converges with only half order accuracy.

It was realized in [12] that for random transport equations with diffusive scalings, upon adopting the gPC-SG formulation, the underlying system is a vector analogy of its deterministic counterpart, thus the deterministic AP machineries can be easily employed to give rise to s-AP schemes. It is also the case here. Once we use the standard gPC-SG approximation, the resulting system, which is deterministic, has the same form as the original Boltzmann equation except that it is vectorized. Indeed we will use the deterministic AP scheme developed in [10] for the gPC-SG system which induces an s-AP scheme, in the sense that its solution will approach, in the zero Knudsen number limit, that of the corresponding drift diffusion equation with random inputs. Here the physical equation is different from those studied in [10], arising in different areas. In addition to the development of the gPC-SG scheme, we also study the regularity (uniform in the Knudsen number) of the solution, proving that the regularity of the initial data in the random space is preserved at later time, *uniformly in the Knudsen number*. Further, we also establish the spectral convergence of the method in the random space.

Using the uniform regularity result, we will also explain why the stochastic collocation method (see for examples [5, 21]) is AP if one uses the deterministic AP method for each random sample.

The paper is organized as follows. In section 2, we introduce the semiconductor Boltzmann equation with random inputs and show its drift-diffusion limit. The regularity of the distribution function in the random space will also be studied. Section 3 describes the gPC-SG method for the Boltzmann equation. Section 4 focuses on the spectral convergence analysis in the random space, which also motivates the need for an AP method when the Knudsen number is small. In section 5, we adopt a fully discretized AP scheme under the diffusive scalings. A formal proof of the stochastic AP property is also given. In section 6, extensive numerical examples are presented to show the validity, spectral convergence, and asymptotic property of the proposed scheme.

2. The semiconductor Boltzmann equation with random inputs.

2.1. The model. We are interested in the linear Boltzmann equation for semiconductor devices [17] under the diffusive scaling with random inputs:

$$
(1)\qquad \begin{cases} \epsilon\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \dfrac{q}{m}\nabla_{\mathbf{x}}\phi(t,\mathbf{x},\mathbf{z})\cdot\nabla_{\mathbf{v}} f = \frac{1}{\epsilon}\mathcal{Q}(f)(t,\mathbf{x},\mathbf{v},\mathbf{z}) + \epsilon G(t,\mathbf{x},\mathbf{v},\mathbf{z}), \\ \qquad\qquad\qquad\qquad\qquad\qquad t>0, \mathbf{x}\in\Omega\subseteq\mathbb{R}^N, \mathbf{v}\in\mathbb{R}^d, \mathbf{z}\in I_{\mathbf{z}}, \\ f(0,\mathbf{x},\mathbf{v},\mathbf{z}) = f_I(\mathbf{x},\mathbf{v},\mathbf{z}), \\ f(t,\mathbf{x},\mathbf{v},\mathbf{z}) = g(t,\mathbf{x},\mathbf{v},\mathbf{z}), \quad \mathbf{x}\in\partial\Omega, \quad \mathbf{v}\cdot\mathbf{n}\le 0. \end{cases}
$$

Here $f(t,\mathbf{x},\mathbf{v},\mathbf{z})$ is the probability density distribution for particles at $\mathbf{x}\in\Omega$, with velocity $\mathbf{v}\in\mathbb{R}^d$. $\mathbf{n}$ is the unit outer normal vector to the boundary $\partial\Omega$ of the spatial domain, ϵ is the Knudsen number, $\phi(t,\mathbf{x},\mathbf{z})$ is the electric potential, and $\mathbf{E}(t,\mathbf{x},\mathbf{z}) = -\nabla_{\mathbf{x}}\phi(t,\mathbf{x},\mathbf{z})$ is the electric field. Here we consider the electric field given a priori for analytical study, though it is usually coupled with a Poisson equation [17] (the scheme can be easily extended to include the Poisson equation and will be tested in section 6). The constants q and m are respectively the elementary charge and the effective mass of the electron. In this paper, we set $q = m = 1$.

The anisotropic collision operator $\mathcal{Q}$ describes a linear approximation of the electron-phonon interaction. It is given by

$$
(2)\qquad \mathcal{Q}(f)(\mathbf{v},\mathbf{z}) = \int_{\mathbb{R}^d}\sigma(\mathbf{v},\mathbf{w},\mathbf{z})\left(M(\mathbf{v})f(\mathbf{w},\mathbf{z}) - M(\mathbf{w})f(\mathbf{v},\mathbf{z})\right)d\mathbf{w},
$$

where M is the normalized Maxwellian,

$$
M(\mathbf{v}) = \frac{1}{\pi^{d/2}}e^{-|\mathbf{v}|^2}.
$$

We assume the anisotropic scattering kernel σ to be symmetric and bounded,

$$
(3)\qquad \exists\ \sigma_0,\sigma_1>0, \quad \sigma_0\le\sigma(\mathbf{v},\mathbf{w},\mathbf{z}) = \sigma(\mathbf{w},\mathbf{v},\mathbf{z})\le\sigma_1,
$$

and the collision frequency

$$
(4)\qquad \lambda(\mathbf{v},\mathbf{z}) = \int_{\mathbb{R}^d}\sigma(\mathbf{v},\mathbf{w},\mathbf{z})M(\mathbf{w})d\mathbf{w}\le\lambda_0.
$$

The uncertainties may come from the collision kernel, the electric potential, initial data, or boundary data. The random variable $\mathbf{z}$ is an n-dimensional random vector with support $I_{\mathbf{z}}$ characterizing the random inputs of the system, so essentially all functions in (1) depend on $\mathbf{z}$. It has a prescribed probability density function $\pi(\mathbf{z})>0$.

We now introduce some spaces, inner product, and norms that will be used throughout this paper.

$$
(5)\quad H = L^2(I_{\mathbf{z}};\pi(\mathbf{z})d\mathbf{z}), \quad <f,g>_H = \int_{I_{\mathbf{z}}} fg\pi(\mathbf{z})d\mathbf{z}, \quad \|f\|_H = \left(\int_{I_{\mathbf{z}}} f^2\pi(\mathbf{z})d\mathbf{z}\right)^{1/2}.
$$

$$
(6)\qquad \|f(t,\cdot,\cdot,\cdot)\|^2_{\Gamma(t)} := \int_{\mathbb{R}^d}\int_{\Omega}||f(t,\mathbf{x},\mathbf{v},\mathbf{z})||^2_H\ e^{-2\phi(\mathbf{x},t)}/M(\mathbf{v})d\mathbf{x}d\mathbf{v}, \quad t\ge 0.
$$

An important property of the collision operator $\mathcal{Q}$ is the symmetry property [18]

$$
(7)\quad \int_{\mathbb{R}^d}\mathcal{Q}(f)(\mathbf{v})g(\mathbf{v})/M(\mathbf{v})d\mathbf{v} = -\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\sigma(\mathbf{v},\mathbf{w},\mathbf{z})M(\mathbf{v})M(\mathbf{w})
$$

$$
(8)\qquad \cdot\left(\frac{f(\mathbf{v})}{M(\mathbf{v})} - \frac{f(\mathbf{w})}{M(\mathbf{w})}\right)\left(\frac{g(\mathbf{v})}{M(\mathbf{v})} - \frac{g(\mathbf{w})}{M(\mathbf{w})}\right)d\mathbf{w}d\mathbf{v},
$$

from which one can deduce

$$\int_{\mathbb{R}^d} \mathcal{Q}(f)(\mathbf{v}) f(\mathbf{v})/M(\mathbf{v}) d\mathbf{v}$$

$$= -\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \sigma(\mathbf{v},\mathbf{w},\mathbf{z}) M(\mathbf{v}) M(\mathbf{w}) \left(\frac{f(\mathbf{v})}{M(\mathbf{v})} - \frac{f(\mathbf{w})}{M(\mathbf{w})}\right)^2 d\mathbf{w} d\mathbf{v} \le 0, \tag{9}$$

where the positivity $\sigma(\mathbf{v},\mathbf{w},\mathbf{z}) \ge \sigma_0(\mathbf{z}) > 0$ is used.

For each $\mathbf{z}$, (1) is a deterministic equation. As $\epsilon \to 0$, $\mathcal{Q}(f) = 0$, then $f(\mathbf{x},\mathbf{v},t) = \rho(\mathbf{x},t)M(\mathbf{v})$, where ρ satisfies the drift-diffusion equation [18, 17]

$$\partial_t \rho = \nabla_{\mathbf{x}}(D(\nabla_{\mathbf{x}}\rho + 2\rho \mathbf{E}(\mathbf{x}))), \tag{10}$$

where the diffusion matrix D is defined by

$$D = \int_{\mathbb{R}^d} \frac{\mathbf{v}\otimes\mathbf{v} M(\mathbf{v})}{\lambda} d\mathbf{v}.$$

The limit $\epsilon \to 0$ is known as the drift-diffusion limit.

2.2. A stability result. In this section, we assume z is a one-dimensional random variable. We will study the stability of f in the random variable z in this section. To this aim, in this and the next subsection, we assume σ depends on z *linearly*. This is mostly the case when one approximates a random field, for example, by the Karhunen–Loeve expansion [16],

$$\sigma(\mathbf{v},\mathbf{w},\mathbf{z}) \approx \bar{\sigma}_0(\mathbf{v},\mathbf{w}) + \sum_{i=1}^{n} \tilde{\sigma}_i(\mathbf{v},\mathbf{w}) z_i, \tag{11}$$

with $z_1, z_2, \ldots, z_n$ independent random variables with probability density function $\pi(\mathbf{z})$.

We assume I_z to be a bounded domain. z can follow several distributions, such as the uniform or Beta distribution. Then the boundedness of σ in z, as a linear function of z, can be assumed. For simplicity one also assumes $\phi = \phi(\mathbf{x},t)$ independent of z.

We begin with a simple lemma on the estimate of $||f||_{\Gamma(t)}$.

LEMMA 2.1.

$$||f||_{\Gamma(t)} \le e^{\hat{\phi} t} ||f_I||_{\Gamma(0)},$$

where $\hat{\phi} = ||\frac{\partial \phi}{\partial t}||_{L^\infty(\Omega\times(0,\infty))}$.

Proof. Upon multiplying $f e^{-2\phi}\pi(z)/M(\mathbf{v})$ to both sides of (1) and integrating on $\Omega \times \mathbb{R}^d \times I_z$,

$$\frac{\epsilon^2}{2}\partial_t ||f||^2_{\Gamma(t)} + \epsilon^2 \int_\Omega \int_{\mathbb{R}^d} \partial_t \phi ||f||^2_H e^{-2\phi}/M(\mathbf{v}) d\mathbf{v} d\mathbf{x}$$

$$= \int_{\mathbb{R}^d}\int_\Omega \langle \mathcal{Q}(f), f\rangle_H e^{-2\phi}/M(\mathbf{v}) d\mathbf{v} d\mathbf{x}.$$

Using (9), RHS ≤ 0, then

$$\frac{1}{2}\partial_t ||f||^2_{\Gamma(t)} + \int_\Omega \int_{\mathbb{R}^d} \partial_t \phi ||f||^2_H e^{-2\phi}/M(\mathbf{v}) d\mathbf{v} d\mathbf{x} \le 0.$$

By Gronwall's inequality, one has

$$||f||_{\Gamma(t)} \le e^{\hat{\phi} t} ||f_I||_{\Gamma(0)}\,. \tag{12}$$

∎

2.3. Uniform regularity. In this section, we prove that the solution f will preserve the regularity of the initial data in the random space. We will obtain an energy bound for higher-order derivatives in the random space, in suitable weighted Sobolev norm, *uniformly in* ϵ.

THEOREM 2.2. *Assume that*

$$\min_z |\sigma| \geq \gamma, \qquad \max_z |\partial_z \sigma| \leq \gamma_1,$$

and $||f_I||_{W(0)} \leq \beta$. *Define the* W *norm as*

$$||f||^2_{W^m(t)} = \sum_{l=1}^{m} \frac{1}{4} \left(\frac{\gamma_1}{\gamma}\right)^{2l} \left(\frac{m\,!}{(m-l)\,!}\right)^2 ||\partial_z^{m-l} f||^2_{\Gamma(t)} + \frac{1}{2}||\partial_z^m f||^2_{\Gamma(t)};$$

then we have

$$||f||_{W(t)} \leq \beta e^{\hat{\phi} t/2},$$

where $\hat{\phi} = ||\frac{\partial \phi}{\partial t}||_{L^\infty(\Omega \times (0,\infty))}$.

Proof. Take the lth-order derivative of z on (1),

$$\epsilon^2 \partial_t(\partial_z^l f) + \epsilon \mathbf{v} \cdot \nabla_{\mathbf{x}}(\partial_z^l f) + \epsilon \nabla_{\mathbf{x}} \phi \cdot \nabla_{\mathbf{v}}(\partial_z^l f) = \partial_z^l Q(f), \tag{13}$$

for all $l = 0, 1, \ldots, m$, where

$$\begin{aligned} \partial_z^l Q(f) = &\int_{\mathbb{R}^d} \sigma(\mathbf{v}, \mathbf{w}, z) \left(M(\mathbf{v}) \partial_z^l f(\mathbf{w}, z) - M(\mathbf{w}) \partial_z^l f(\mathbf{v}, z)\right) d\mathbf{w} \\ &+ l \int_{\mathbb{R}^d} \partial_z \sigma(\mathbf{v}, \mathbf{w}, z) \left(M(\mathbf{v}) \partial_z^l f(\mathbf{w}, z) - M(\mathbf{w}) \partial_z^l f(\mathbf{v}, z)\right) d\mathbf{w}. \end{aligned} \tag{14}$$

We multiply by $\frac{1}{2}(\frac{\gamma_1}{\gamma})^{2l}(\frac{m\,!}{(m-l)\,!})^2(\partial_z^{m-l} f) e^{-2\phi(x,t)} \pi(z)/M(v)$ (for $l = 1, \ldots, m$) and $(\partial_z^m f) e^{-2\phi(x,t)} \pi(z)/M(v)$ (for $l = 0$) to both sides of (14) and integrate on $\Omega \times \mathbb{R}^d \times I_z$. Adding up these equations for $l = 0, \ldots, m$, summation of the right-hand side becomes

$$\begin{aligned} &\int_{I_z} \int_\Omega \int_{\mathbb{R}^d} \partial_z^m Q(f) \partial_z^m f(\mathbf{v})/M(\mathbf{v}) d\mathbf{v} e^{-2\phi} d\mathbf{x} \pi(z) dz \\ &\quad + \sum_{l=1}^{m} \frac{1}{2} \left(\frac{\gamma_1}{\gamma}\right)^{2l} \left(\frac{m\,!}{(m-l)\,!}\right)^2 \int_{I_z} \int_\Omega \int_{\mathbb{R}^d} \partial_z^{m-l} Q(f) \partial_z^{m-l} f(\mathbf{v})/M(\mathbf{v}) d\mathbf{v} e^{-2\phi} d\mathbf{x} \pi(z) dz \\ &= -\frac{1}{2} \int_{I_z} \int_\Omega \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \sigma(\mathbf{v}, \mathbf{w}, z) M(\mathbf{v}) M(\mathbf{w}) \left(\frac{\partial_z^m f(\mathbf{v})}{M(\mathbf{v})} - \frac{\partial_z^m f(\mathbf{w})}{M(\mathbf{w})}\right)^2 d\mathbf{w} d\mathbf{v} e^{-2\phi} d\mathbf{x} \pi(z) dz \\ &\quad - \frac{m}{2} \int_{I_z} \int_\Omega \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \partial_z \sigma(\mathbf{v}, \mathbf{w}, z) M(\mathbf{v}) M(\mathbf{w}) \left(\frac{\partial_z^{m-1} f(\mathbf{v})}{M(\mathbf{v})} - \frac{\partial_z^{m-1} f(\mathbf{w})}{M(\mathbf{w})}\right) \\ &\qquad \times \left(\frac{\partial_z^m f(\mathbf{v})}{M(\mathbf{v})} - \frac{\partial_z^m f(\mathbf{w})}{M(\mathbf{w})}\right) d\mathbf{w} d\mathbf{v} e^{-2\phi} d\mathbf{x} \pi(z) dz \\ &\quad + \sum_{l=1}^{m-1} \frac{1}{2} \left(\frac{\gamma_1}{\gamma}\right)^{2l} \left(\frac{m\,!}{(m-l)\,!}\right)^2 \Big[-\frac{1}{2} \int_{I_z} \int_\Omega \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \sigma(\mathbf{v}, \mathbf{w}, z) M(\mathbf{v}) M(\mathbf{w}) \\ &\qquad \times \left(\frac{\partial_z^{m-l} f(\mathbf{v})}{M(\mathbf{v})} - \frac{\partial_z^{m-l} f(\mathbf{w})}{M(\mathbf{w})}\right)^2 d\mathbf{w} d\mathbf{v} \end{aligned}$$

$$
\begin{aligned}
&\times e^{-2\phi} d\mathbf{x}\pi(z)dz - \frac{(m-l)}{2}\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \partial_z\sigma(\mathbf{v},\mathbf{w},z)M(\mathbf{v})M(\mathbf{w})\\
&\times\left(\frac{\partial_z^{m-l-1}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-l-1}f(\mathbf{w})}{M(\mathbf{w})}\right)\left(\frac{\partial_z^{m-l}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-l}f(\mathbf{w})}{M(\mathbf{w})}\right)d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\Big]\\
&+\frac{1}{2}\left(\frac{\gamma_1}{\gamma}\right)^{2m}(m!)^2\left(-\frac{1}{2}\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\sigma(\mathbf{v},\mathbf{w},z)M(\mathbf{v})M(\mathbf{w})\right.\\
&\left.\times\left(\frac{f(\mathbf{v})}{M(\mathbf{v})}-\frac{f(\mathbf{w})}{M(\mathbf{w})}\right)^2 d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\right)\\
\le&-\frac{\gamma}{2}\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\left(\frac{\partial_z^m f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^m f(\mathbf{w})}{M(\mathbf{w})}\right)^2 d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\\
&+\frac{m\gamma_1}{2}\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\left(\frac{\partial_z^{m-1}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-1}f(\mathbf{w})}{M(\mathbf{w})}\right)\\
&\times\left(\frac{\partial_z^m f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^m f(\mathbf{w})}{M(\mathbf{w})}\right)d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\\
&+\sum_{l=1}^{m-1}\frac{1}{2}\left(\frac{\gamma_1}{\gamma}\right)^{2l}\left(\frac{m!}{(m-l)!}\right)^2\Big[-\frac{\gamma}{2}\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\\
&\times\left(\frac{\partial_z^{m-l}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-l}f(\mathbf{w})}{M(\mathbf{w})}\right)^2 d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\\
&+\frac{\gamma_1}{2}(m-l)\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\left(\frac{\partial_z^{m-l-1}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-l-1}f(\mathbf{w})}{M(\mathbf{w})}\right)\\
&\times\left(\frac{\partial_z^{m-l}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-l}f(\mathbf{w})}{M(\mathbf{w})}\right)d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\Big]\\
&-\frac{\gamma}{4}\left(\frac{\gamma_1}{\gamma}\right)^{2m}(m!)^2\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\left(\frac{f(\mathbf{v})}{M(\mathbf{v})}-\frac{f(\mathbf{w})}{M(\mathbf{w})}\right)^2 d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\\
=&-\frac{\gamma}{2}\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\left[\left(\frac{\partial_z^m f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^m f(\mathbf{w})}{M(\mathbf{w})}\right)\right.\\
&\left.-\frac{m}{2}\frac{\gamma_1}{\gamma}\left(\frac{\partial_z^{m-1}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-1}f(\mathbf{w})}{M(\mathbf{w})}\right)\right]^2 d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\\
&-\sum_{l=1}^{m-1}\frac{\gamma}{8}\left(\frac{\gamma_1}{\gamma}\right)^{2l}\left(\frac{m!}{(m-l)!}\right)^2\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\Big[\left(\frac{\partial_z^{m-l}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-l}f(\mathbf{w})}{M(\mathbf{w})}\right)\\
&-(m-l)\frac{\gamma_1}{\gamma}\left(\frac{\partial_z^{m-l-1}f(\mathbf{v})}{M(\mathbf{v})}-\frac{\partial_z^{m-l-1}f(\mathbf{w})}{M(\mathbf{w})}\right)\Big]^2 d\mathbf{w}d\mathbf{v}e^{-2\phi}d\mathbf{x}\pi(z)dz\\
&-\frac{\gamma}{8}\left(\frac{\gamma_1}{\gamma}\right)^{2m}(m!)^2\int_{I_z}\int_{\Omega}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})
\end{aligned}
$$

$$\times \left(\frac{f(\mathbf{v})}{M(\mathbf{v})} - \frac{f(\mathbf{w})}{M(\mathbf{w})} \right)^2 d\mathbf{w} d\mathbf{v} e^{-2\phi} d\mathbf{x} \pi(z) dz$$

$$\leq 0.$$

Combining summation of the right-hand side, we then have

$$\partial_t ||f||^2_{W(t)} \leq \hat{\phi} ||f||^2_{W(t)}.$$

Applying Gronwall's inequality,

$$||f||_{W(t)} \leq e^{\hat{\phi} t/2} ||f_I||_{W(0)} \leq \beta e^{\hat{\phi} t/2}. \qquad \square$$

3. A gPC-sG method. By the stochastic Galerkin method, one seeks an orthogonal polynomial expansion for the solution of problem (1). That is, for random variable $\mathbf{z} \in I_{\mathbf{z}}$,

$$f(t, \mathbf{x}, \mathbf{v}, \mathbf{z}) \approx f_K(t, \mathbf{x}, \mathbf{v}, \mathbf{z}) = \sum_{|\mathbf{k}|=1}^{K} \alpha_{\mathbf{k}}(t, \mathbf{x}, \mathbf{v}) \psi_{\mathbf{k}}(\mathbf{z}) = \boldsymbol{\alpha} \cdot \boldsymbol{\psi}, \quad K = \binom{n+P}{n}. \tag{15}$$

Here $\mathbf{k} = (k_1, \ldots, k_n)$ is a multi-index with $|\mathbf{k}| = k_1 + \cdots + k_n$, and the coefficient vectors are given by

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K), \qquad \boldsymbol{\psi} = (\psi_1, \ldots, \psi_K).$$

$\{\Psi_{\mathbf{k}}(\mathbf{z})\}$ are the orthonormal basis functions that form $\mathbb{P}_P^n$ (the set of n-variate orthonormal polynomials of degree up to $P \geq 1$) and satisfy

$$\int_{I_{\mathbf{z}}} \psi_{\mathbf{k}}(\mathbf{z}) \psi_{\mathbf{l}}(\mathbf{z}) \pi(\mathbf{z}) d\mathbf{z} = \delta_{\mathbf{kl}}, \qquad 1 \leq |\mathbf{k}|, |\mathbf{l}| \leq K = \dim(\mathbb{P}_P^n),$$

where $\delta_{\mathbf{kl}}$ is the Kronecker Delta function. The commonly used pairs of $\{\psi_{\mathbf{k}}(\mathbf{z})\}$ and $\pi(\mathbf{z})$ include Hermite–Gaussian, Legendre-uniform, Laguerre-Gamma, etc. If the random dimension $n > 1$, one can reorder the multidimensional polynomials $\{\psi_{\mathbf{k}}(\mathbf{z})\}$ of $\mathbf{z}$ into a single index k. One can refer to section 5.2 of [20], where the graded lexicographic ordering is introduced.

The initial values of each component of $\boldsymbol{\alpha}$ are the gPC coefficients of the initial datum for f which is denoted by $f_I(\mathbf{x}, \mathbf{v}, \mathbf{z})$,

$$\alpha_{\mathbf{k}}(0, \mathbf{x}, \mathbf{v}) = \langle f_I(\mathbf{x}, \mathbf{v}, \cdot), \psi_{\mathbf{k}} \rangle_H, \quad |\mathbf{k}| = 1, \ldots, K. \tag{16}$$

We apply f_K into (1) and take an inner product with $\psi_{\mathbf{k}}(\mathbf{z})$ to get

$$\begin{aligned} &\epsilon \frac{\partial \boldsymbol{\alpha}}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} \boldsymbol{\alpha} + \nabla_{\mathbf{x}} \phi \cdot \nabla_{\mathbf{v}} \boldsymbol{\alpha} \\ &\quad = \langle \mathcal{Q}(f_K), \boldsymbol{\psi} \rangle_H \\ &\quad = \frac{1}{\epsilon} \int_{\mathbb{R}^d} B(\mathbf{v}, \mathbf{w}) \left[M(\mathbf{v}) \boldsymbol{\alpha}(\mathbf{w}) - M(\mathbf{w}) \boldsymbol{\alpha}(\mathbf{v}) \right] d\mathbf{w} + \epsilon \tilde{G}(\mathbf{v}) \\ &\quad = \frac{1}{\epsilon} \mathbf{Q}(\boldsymbol{\alpha}) + \epsilon \tilde{G}, \end{aligned} \tag{17}$$

where we denote

$$\mathbf{Q}(\boldsymbol{\alpha}) = \int_{\mathbb{R}^d} B(\mathbf{v},\mathbf{w})\left[M(\mathbf{v})\boldsymbol{\alpha}(\mathbf{w}) - M(\mathbf{w})\boldsymbol{\alpha}(\mathbf{v})\right] d\mathbf{w}, \tag{18}$$

and matrices $B(\mathbf{v},\mathbf{w}) = (B_{\mathbf{ij}})_{K\times K}$, and the vector $\tilde{G} = (\tilde{G}_{\mathbf{k}})_{K\times 1}$ are given by

$$B_{\mathbf{ij}}(\mathbf{v},\mathbf{w}) = \int_{I_{\mathbf{z}}} \sigma(\mathbf{v},\mathbf{w},\mathbf{z})\psi_{\mathbf{i}}(\mathbf{z})\psi_{\mathbf{j}}(\mathbf{z})\pi(\mathbf{z})d\mathbf{z},$$

$$\tilde{G}_{\mathbf{k}}(t,\mathbf{x},\mathbf{v}) = \int_{I_{\mathbf{z}}} G(t,\mathbf{x},\mathbf{v},\mathbf{z})\psi_{\mathbf{k}}(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}. \tag{19}$$

Since $\sigma, \lambda > 0$, the matrix B is symmetric and positive definite [22].

Similar to the property that $\ker \mathcal{Q} = \operatorname{span}\{M(\mathbf{v})\}$, we have the following result which will be used when we prove the stochastic AP property of our scheme. The proof is similar to that in [18] for the collision operator $\mathcal{Q}$ defined in (2).

LEMMA 3.1. *The unique solution of the system of equations* $\mathbf{Q}(\mathbf{r}) = 0$ *is* $\mathbf{r}(\mathbf{v}) = cM(\mathbf{v})$, *where* $c = \int_{\mathbb{R}^d} \mathbf{r}(\mathbf{w})d\mathbf{w}$ *is independent of* $\mathbf{v}$.

Proof. Using the symmetry property $\mathbf{Q}$, one has

$$\begin{aligned}
\int_{\mathbb{R}^d} \frac{\mathbf{Q}(\mathbf{r})^T\mathbf{r}(\mathbf{v})}{M(\mathbf{v})}d\mathbf{v} &= \frac{1}{2}\Bigg(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\mathbf{r}(\mathbf{w})^T}{M(\mathbf{w})} B(\mathbf{v},\mathbf{w})M(\mathbf{v})M(\mathbf{w})\frac{\mathbf{r}(\mathbf{v})}{M(\mathbf{v})}d\mathbf{w}d\mathbf{v} \\
&\quad + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\mathbf{r}(\mathbf{v})^T}{M(\mathbf{v})} B(\mathbf{w},\mathbf{v})M(\mathbf{w})M(\mathbf{v})\frac{\mathbf{r}(\mathbf{w})}{M(\mathbf{w})}d\mathbf{v}d\mathbf{w} \\
&\quad - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\mathbf{r}(\mathbf{v})^T}{M(\mathbf{v})} B(\mathbf{v},\mathbf{w})M(\mathbf{w})M(\mathbf{v})\frac{\mathbf{r}(\mathbf{v})}{M(\mathbf{v})}d\mathbf{w}d\mathbf{v} \\
&\quad - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\mathbf{r}(\mathbf{w})^T}{M(\mathbf{w})} B(\mathbf{w},\mathbf{v})M(\mathbf{v})M(\mathbf{w})\frac{\mathbf{r}(\mathbf{w})}{M(\mathbf{w})}d\mathbf{v}d\mathbf{w}\Bigg) \\
&= -\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left(\frac{\mathbf{r}(\mathbf{w})}{M(\mathbf{w})} - \frac{\mathbf{r}(\mathbf{v})}{M(\mathbf{v})}\right)^T B(\mathbf{v},\mathbf{w})M(\mathbf{v})M(\mathbf{w}) \left(\frac{\mathbf{r}(\mathbf{w})}{M(\mathbf{w})} - \frac{\mathbf{r}(\mathbf{v})}{M(\mathbf{v})}\right) d\mathbf{v}d\mathbf{w} \\
&\le 0,
\end{aligned}$$

since B is positive definite. $B(\mathbf{v},\mathbf{w}) = B(\mathbf{w},\mathbf{v})$ is used in the second equality. Clearly if $\mathbf{Q}(\boldsymbol{\alpha}) = 0$, the above inequality becomes equality. Since B is positive definite, the term in the last line implies $\frac{\mathbf{r}(\mathbf{w})}{M(\mathbf{w})} = \frac{\mathbf{r}(\mathbf{v})}{M(\mathbf{v})}$ for any $\mathbf{v}$ and $\mathbf{w}$. This implies $r(\mathbf{v}) = cM(\mathbf{v})$ for c independent of $\mathbf{v}$. Integrating over $\mathbf{v}$ gives $c = \int_{\mathbb{R}^d} r(\mathbf{v})d\mathbf{v}$. ∎

4. Spectral convergence analysis. We now prove that the gPC-SG method is convergent and stable under a suitably defined energy norm given in (6). An error estimate and spectral convergence rate in the probability space will also be presented. The proofs are similar to the convergence analysis of the moment method for linear kinetic model [23].

Let f be the solution to the Boltzmann equation (1). We define the projection operator

$$P_K f = \sum_{|\mathbf{k}|=1}^{K} \langle f, \psi_{\mathbf{k}}\rangle_H \psi_{\mathbf{k}}.$$

The error arisen from the gPC-SG can be split into two parts, R_K and e_K,

$$f - f_K = f - P_K f + P_K f - f_K := R_K + e_K, \tag{20}$$

where $R_K = f - P_K f$ is the projection error, and

$$e_K = P_K f - f_K = \sum_{|\mathbf{k}|=1}^{K} \big(\langle f, \psi_{\mathbf{k}}\rangle_H - f_{\mathbf{k}}\big)\psi_{\mathbf{k}} = \hat{\mathbf{e}} \cdot \boldsymbol{\psi},$$

where $\hat{\mathbf{e}} = \big(\langle f, \psi_1\rangle_H - f_1, \ldots, \langle f, \psi_K\rangle_H - f_K\big)$ is the numerical error, and $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_K)$.

We first give the projection error. By the standard error estimate for orthogonal polynomial approximations and Theorem 2.2,

$$||R_K||_{\Gamma(t)} \le C_0 K^{-m} ||\partial_z^m f||_{\Gamma(t)} \le \frac{C_1}{K^m} e^{\hat{\phi} t/2}, \tag{21}$$

where C_0, C_1 are constants independent of ϵ.

It remains to estimate e_K. Define the operator

$$\mathcal{L} = \epsilon^2 \partial_t + \epsilon \mathbf{v} \cdot \nabla_{\mathbf{x}} + \epsilon \nabla_{\mathbf{x}} \phi \cdot \nabla_{\mathbf{v}} - \mathcal{Q}. \tag{22}$$

We first prove two results.

LEMMA 4.1. $\langle \mathcal{L}(R_K), \boldsymbol{\psi}\rangle_H = -\langle \mathcal{Q}(R_K), \boldsymbol{\psi}\rangle_H$.

Proof. Since $R_K = f - P_K f = f - \sum_{|\mathbf{k}|=1}^{K} \langle f, \psi_{\mathbf{k}}\rangle_H \psi_{\mathbf{k}}$,

$$\begin{aligned}
\langle \partial_t R_K, \boldsymbol{\psi}\rangle_H &= \langle \partial_t f, \boldsymbol{\psi}\rangle_H - \left\langle \sum_{|\mathbf{k}|=1}^{K} \partial_t \langle f, \psi_{\mathbf{k}}\rangle_H \psi_{\mathbf{k}}, \boldsymbol{\psi} \right\rangle_H \\
&= \langle \partial_t f, \boldsymbol{\psi}\rangle_H - \sum_{|\mathbf{k}|=1}^{K} \langle \partial_t f, \psi_{\mathbf{k}}\rangle_H \langle \psi_{\mathbf{k}}, \boldsymbol{\psi}\rangle_H \\
&= \langle \partial_t f, \boldsymbol{\psi}\rangle_H - \langle \partial_t f, \boldsymbol{\psi}\rangle_H = 0.
\end{aligned} \tag{23}$$

The second and third terms of $\langle \mathcal{L}(R_K), \boldsymbol{\psi}\rangle_H$ are also zeros by similar arguments. Therefore, $\langle \mathcal{L}(R_K), \boldsymbol{\psi}\rangle_H = -\langle \mathcal{Q}(R_K), \boldsymbol{\psi}\rangle_H$. ∎

LEMMA 4.2.

$$||\mathcal{Q}(R_K)||_{\Gamma(t)} \le \frac{C_1}{K^m} e^{\hat{\phi} t/2}.$$

Proof. Denote

$$\begin{aligned}
Q(R_K)(\mathbf{v}, z) &= M(\mathbf{v}) \int_{\mathbb{R}^d} \sigma(\mathbf{v}, \mathbf{w}, z) R_K(\mathbf{w}, z) d\mathbf{w} - \int_{\mathbb{R}^d} \sigma(\mathbf{v}, \mathbf{w}, z) M(\mathbf{w}) d\mathbf{w} R_K(\mathbf{v}, z) \\
&:= I + II,
\end{aligned}$$

then $||Q(R_K)||^2_{\Gamma(t)} \le 2(||I||^2_{\Gamma(t)} + ||II||^2_{\Gamma(t)})$.

It is easy to see that $||II||^2_{\Gamma(t)} \le \sigma_1^2 ||R_K||^2_{\Gamma(t)}$. Now we estimate I.

$$\begin{aligned}
\|I\|^2_{\Gamma(t)} &= \left|\left|M(\mathbf{v})\int_{\mathbb{R}^d}\sigma(\mathbf{v},\mathbf{w},z)R_K(\mathbf{w},z)d\mathbf{w}\right|\right|^2_{\Gamma(t)} \\
&= \int_{I_{\mathbf{z}}}\int_{\mathbb{R}^d} M(\mathbf{v})\int_\Omega\left(\int_{\mathbb{R}^d}\sigma(\mathbf{v},\mathbf{w},z)R_K(\mathbf{w},z)d\mathbf{w}\right)^2 e^{-2\phi}d\mathbf{x}d\mathbf{v}\pi(z)dz \\
&\le \int_{I_{\mathbf{z}}}\int_{\mathbb{R}^d} M(\mathbf{v})\int_\Omega\left(\int_{\mathbb{R}^d}\sigma^2(\mathbf{v},\mathbf{w},z)M(\mathbf{w})d\mathbf{w}\right)\left(\int_{\mathbb{R}^d}R_K^2(\mathbf{w},z)/M(\mathbf{w})d\mathbf{w}\right) \\
&\quad\cdot e^{-2\phi}\ d\mathbf{x}d\mathbf{v}\pi(z)dz \le \sigma_1^2||R_K||^2_{\Gamma(t)},
\end{aligned}$$

where we used

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\sigma^2(\mathbf{v},\mathbf{w},z)M(\mathbf{v})M(\mathbf{w})d\mathbf{w}d\mathbf{v} \le \sigma_1^2\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(\mathbf{v})M(\mathbf{w})\ d\mathbf{v}d\mathbf{w} = \sigma_1^2.$$

Therefore, $||Q(R_K)||^2_{\Gamma(t)} \le 2(\sigma_1^2||R_K||^2_{\Gamma(t)} + \sigma_1^2||R_K||^2_{\Gamma(t)})$; thus

$$||\mathcal{Q}(R_K)||_{\Gamma(t)} \lesssim ||R_K||_{\Gamma(t)} \le \frac{C_1}{K^m}e^{\hat{\phi}t/2}. \qquad \square$$

Since $\mathcal{L}(f) = 0$ and $P_K\mathcal{L}(f_K) = 0$, from (20), $\langle\mathcal{L}(e_K),\psi_{\mathbf{k}}\rangle_H = -\langle\mathcal{L}(R_K),\psi_{\mathbf{k}}\rangle_H$, for $|\mathbf{k}| = 1,\dots K$, that is,

$$\langle\mathcal{L}(e_K),\boldsymbol{\psi}\rangle_H = -\langle\mathcal{L}(R_K),\boldsymbol{\psi}\rangle_H. \tag{24}$$

Taking the scalar product of (24) with $e^{-2\phi(\mathbf{x},t)}\cdot\hat{\mathbf{e}}/M(\mathbf{v})$ and integrating on $\Omega\times\mathbb{R}^d$ gives

$$\begin{aligned}
&\epsilon^2\left(\frac{1}{2}\partial_t\int_{I_z}\int_{\mathbb{R}^d}\int_\Omega e^{-2\phi}e_K^2/M(\mathbf{v})d\mathbf{x}d\mathbf{v}\pi(z)dz\right. \\
&\qquad\left.+\int_{I_z}\int_{\mathbb{R}^d}\int_\Omega(\partial_t\phi)e^{-2\phi}e_K^2/M(\mathbf{v})d\mathbf{x}d\mathbf{v}\pi(z)dz\right) \\
&\qquad-\int_{\mathbb{R}^d}\int_\Omega\langle\mathcal{Q}(e_K),\boldsymbol{\psi}\rangle_H\cdot\hat{\mathbf{e}}\ \ e^{-2\phi}/M(\mathbf{v})d\mathbf{x}d\mathbf{v} \\
&= -\int_{\mathbb{R}^d}\int_\Omega\langle\mathcal{L}(R_K),\boldsymbol{\psi}\rangle_H\cdot\hat{\mathbf{e}}\ \ e^{-2\phi}/M(\mathbf{v})d\mathbf{x}d\mathbf{v} \\
&= \int_{\mathbb{R}^d}\int_\Omega\langle\mathcal{Q}(R_K),\boldsymbol{\psi}\rangle_H\cdot\hat{\mathbf{e}}\ \ e^{-2\phi}/M(\mathbf{v})d\mathbf{x}d\mathbf{v} \\
&= \int_{\mathbb{R}^d}\int_\Omega\langle\mathcal{Q}(R_K),e_K\rangle_H\, e^{-2\phi}/M(\mathbf{v})d\mathbf{x}d\mathbf{v} \\
&\le \int_{\mathbb{R}^d}\int_\Omega||\mathcal{Q}(R_K)||_H\cdot||e_K||_H\, e^{-2\phi}/M(\mathbf{v})d\mathbf{x}d\mathbf{v} \\
&\le \left(\int_{\mathbb{R}^d}\int_\Omega||\mathcal{Q}(R_K)||^2_H\ \ e^{-2\phi}/M(\mathbf{v})d\mathbf{x}d\mathbf{v}\right)^{\frac{1}{2}}\left(\int_{\mathbb{R}^d}\int_\Omega||e_K||^2_H\ \ e^{-2\phi}/M(\mathbf{v})d\mathbf{x}d\mathbf{v}\right)^{\frac{1}{2}} \\
&= ||\mathcal{Q}(R_K)||_{\Gamma(t)}\cdot||e_K||_{\Gamma(t)},
\end{aligned}$$

where the second equality uses Lemma 4.1 and the inequality is by the Cauchy–Schwartz inequality.

From (9), one has

$$\frac{\epsilon^2}{2}\partial_t\left(||e_K||^2_{\Gamma(t)}\right) \le \epsilon^2\hat{\phi}||e_K||^2_{\Gamma(t)} + ||\mathcal{Q}(R_K)||_{\Gamma(t)} \cdot ||e_K||_{\Gamma(t)}.$$

Canceling an $\|e_K\|_{\Gamma(t)}$ term on both sides from the above inequality and using Lemma 4.2, one gets

$$\partial_t||e_K||_{\Gamma(t)} \le \hat{\phi}||e_K||_{\Gamma(t)} + \frac{1}{\epsilon^2}\frac{C_1}{K^m}e^{\hat{\phi}t/2}\,.$$

Apply Gronwall's inequality,

$$||e_K||_{\Gamma(t)} \le e^{\hat{\phi}t}||e_K(0)||_{\Gamma(0)} + \frac{1}{\epsilon^2}\frac{C_1}{K^m}e^{\hat{\phi}t/2}\int_0^t e^{\hat{\phi}(t-\tau)}d\tau\,. \tag{25}$$

We are now ready to summarize the result in the following theorem.

THEOREM 4.3 (spectral convergence). *Assuming that σ depends on z linearly, $\hat{\phi}$ is bounded, and*

$$\max_z|\sigma| \le \gamma_0, \qquad \max_z|\partial_z\sigma| \le \gamma_1,$$

we have

$$||f - f_K||_{\Gamma(t)} \le \frac{C_2}{\epsilon^2}\frac{1}{K^m}e^{\hat{\phi}t/2}\,, \tag{26}$$

where C_2 is independent of ϵ and m.

Proof. From (20), one has

$$\|f - f_K\|_{\Gamma(t)} \le \|R_K\|_{\Gamma(t)} + \|e_K\|_{\Gamma(t)}\,.$$

Note $e_K(0) = P_kf - f_k|_{t=0} = 0$. Now combining (21) and (25) gives the desired inequality (26). ∎

Remark 4.4. The $1/\epsilon^2$ factor on the right-hand side of (26) is due to the truncation error from the gPC approximation of the collision operator Q and thus is intrinsic, like the error in deterministic problems [8]. Equation (26) shows the necessity of using an ϵ-dependent K (actually it shows larger K is needed for smaller ϵ), motivating the development of the s-AP schemes.

In [9] for the linear transport equation with the special case of isotropic collision kernel, a sharp (ϵ-uniform) spectral convergence was established. Whether one can obtain such a sharp convergence of the anisotropic collision studied in this paper remains to be investigated. It is worthwhile to point out that even if one obtains such a sharp convergence of the gPC-SG approximation, the time and spatial discretizations still need to be AP as is the case of the deterministic cases. This is what the subsequent sections will be aimed at.

5. A fully discrete stochastic AP scheme. For clarity, we consider $v \in \mathbb{R}$. The space, time, and velocity discretization here follow that of the deterministic case in [11], using the even- and odd-parities formulation in velocity, the Wild sum approximation to handle the stiffness in time, center and upwind approximations in space, and the spectral approximation in velocity.

First, (1) can be split into two equations, one for $v > 0$ and one for $-v < 0$,

$$\begin{cases} \epsilon\partial_t f + v\partial_x f - E\partial_v f = \dfrac{1}{\epsilon}\mathcal{Q}(f) + \epsilon G, \\ \epsilon\partial_t f - v\partial_x f + E\partial_v f = \dfrac{1}{\epsilon}\mathcal{Q}(f) + \epsilon G. \end{cases} \tag{27}$$

Introduce the even and odd parities r and j for $v > 0$,

$$\begin{aligned} r(t,x,v,\mathbf{z}) &= \frac{1}{2}\left(f(t,x,v,\mathbf{z}) + f(t,x,-v,\mathbf{z})\right), \\ j(t,x,v,\mathbf{z}) &= \frac{1}{2\epsilon}(f(t,x,v,\mathbf{z}) - f(t,x,-v,\mathbf{z})); \end{aligned} \tag{28}$$

adding and subtracting the two equations in (27), one has

$$\begin{cases} \partial_t r + v\partial_x j - E\partial_v j = \dfrac{1}{\epsilon^2}\mathcal{Q}(r) + G, \\ \partial_t j + \frac{1}{\epsilon^2}(v\partial_x r - E\partial_v r) = -\dfrac{1}{\epsilon^2}\lambda j. \end{cases} \tag{29}$$

From now on we will only consider the case of $v > 0$.

By the gPC-SG approach, one inserts the approximate solutions r_K, j_K

$$r_K(t,x,v,\mathbf{z}) = \sum_{|\mathbf{k}|=1}^{K} \hat{r}_{\mathbf{k}}(t,x,v)\psi_{\mathbf{k}}(\mathbf{z}), \quad j_K(t,x,v,\mathbf{z}) = \sum_{|\mathbf{k}|=1}^{K} \hat{j}_{\mathbf{k}}(t,x,v)\psi_{\mathbf{k}}(\mathbf{z})$$

into (29) and enforces the residual to be orthogonal to the polynomial space spanned by $\boldsymbol{\psi}(\mathbf{z}) = (\psi_1(\mathbf{z}), \dots, \psi_K(\mathbf{z}))^T$. Denote

$$\hat{\mathbf{r}}(t,x,v) = (\hat{r}_1(t,x,v), \dots, \hat{r}_K(t,x,v))^T, \qquad \hat{\mathbf{j}}(t,x,v) = (\hat{j}_1(t,x,v), \dots, \hat{j}_K(t,x,v))^T$$

and the matrix

$$F(\mathbf{v}) = \int_{\mathbb{R}^d} B(\mathbf{v},\mathbf{w})M(\mathbf{w})d\mathbf{w}; \tag{30}$$

then F is symmetric and positive definite.

A system of equations for vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{j}}$ can then be obtained as

$$\begin{cases} \partial_t\hat{\mathbf{r}} + v\partial_x\hat{\mathbf{j}} - E\partial_v\hat{\mathbf{j}} = \dfrac{1}{\epsilon^2}\mathbf{Q}(\hat{\mathbf{r}}) + \int_{I_{\mathbf{z}}} G(t,x,v,\mathbf{z})\boldsymbol{\psi}(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}, \\ \partial_t\hat{\mathbf{j}} + \dfrac{1}{\epsilon^2}(v\partial_x\hat{\mathbf{r}} - E\partial_v\hat{\mathbf{r}}) = -\dfrac{1}{\epsilon^2}F(v)\hat{\mathbf{j}}, \end{cases} \tag{31}$$

where matrices B, F are defined in (19). As was done in [10], we rewrite (31) into a diffusive relaxation system,

$$\begin{cases} \partial_t\hat{\mathbf{r}} + v\partial_x\hat{\mathbf{j}} - E\partial_v\hat{\mathbf{j}} = \dfrac{1}{\epsilon^2}\mathbf{Q}(\hat{\mathbf{r}}) + \int_{I_{\mathbf{z}}} G(t,x,v,\mathbf{z})\boldsymbol{\psi}(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}, \\ \partial_t\hat{\mathbf{j}} + \phi\left(v\partial_x\hat{\mathbf{r}} - E\partial_v\hat{\mathbf{r}}\right) = -\dfrac{1}{\epsilon^2}\left[F(v)\hat{\mathbf{j}} + (1-\epsilon^2\phi)(v\partial_x\hat{\mathbf{r}} - E\partial_v\hat{\mathbf{r}})\right], \end{cases} \tag{32}$$

where $\phi = \phi(\epsilon)$ is a control parameter satisfying $0 \le \phi \le 1/\epsilon^2$. A simple choice of ϕ is

$$\phi(\epsilon) = \min\left\{1, \frac{1}{\epsilon^2}\right\}.$$

5.1. An s-AP time-splitting. Here the main difficulty is to handle the stiff collision operator in an efficient way. We employ the conventional time-splitting procedure [11, 10] for the diffusive relaxation system (32), which is composed of a relaxation step,

$$\begin{cases} \partial_t \hat{\mathbf{r}} = \dfrac{1}{\epsilon^2}\mathbf{Q}(\hat{\mathbf{r}}) + \int_{I_{\mathbf{z}}} G(t,x,v,\mathbf{z})\boldsymbol{\psi}(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}, \\ \partial_t \hat{\mathbf{j}} = -\dfrac{1}{\epsilon^2}\left(F(v)\hat{\mathbf{j}} + (1-\epsilon^2\phi)\mathbf{d}(\hat{\mathbf{r}})\right), \end{cases} \tag{33}$$

where the vector $\mathbf{d}(\hat{\mathbf{r}}) = v\partial_x\hat{\mathbf{r}} - E\partial_v\hat{\mathbf{r}}$, followed by a convection step,

$$\begin{cases} \partial_t \hat{\mathbf{r}} + v\partial_x\hat{\mathbf{j}} - E\partial_v\hat{\mathbf{j}} = 0, \\ \partial_t \hat{\mathbf{j}} + \phi(v\partial_x\hat{\mathbf{r}} - E\partial_v\hat{\mathbf{r}}) = 0. \end{cases} \tag{34}$$

The AP property of the splitting can be shown as follows. Under the diffusive scaling, as $\epsilon \to 0$, (33) becomes $\mathbf{Q}(\hat{\mathbf{r}}) = 0$. By Lemma 3.1, the solution is given by

$$\hat{r}_{\mathbf{k}}(v) = \hat{\rho}_{\mathbf{k}} M(v), \tag{35}$$

where $\hat{\rho}_{\mathbf{k}} = \int_{\mathbb{R}} \hat{\mathbf{r}}_{\mathbf{k}}(w)dw$. The second equation of (33) gives

$$\hat{j}_{\mathbf{k}} = -\sum_{\mathbf{l}} (F^{-1})_{\mathbf{kl}}(v\partial_x\hat{r}_{\mathbf{l}} - E\partial_v\hat{r}_{\mathbf{l}}). \tag{36}$$

Applying (35) and (36) to the first equation of (34) and integrating over v, one gets

$$\partial_t\hat{\rho}_{\mathbf{k}} = \partial_x\left(T\sum_{\mathbf{l}}(F^{-1})_{\mathbf{kl}}(\partial_x\hat{\rho}_{\mathbf{l}} + 2\hat{\rho}_{\mathbf{l}}E)\right), \tag{37}$$

where

$$T = \int_{\mathbb{R}} v^2 M(v)dv.$$

On the other hand, applying the ansatz $\rho(t,x,\mathbf{z}) = \sum_{|\mathbf{k}|=1}^{K} \hat{\rho}_{\mathbf{k}}(t,x)\psi_{\mathbf{k}}(\mathbf{z})$ and conducting the Galerkin projection for the limiting drift-diffusion equation (10), we obtain

$$\partial_t\hat{\rho}_{\mathbf{k}} = \partial_x\left(T\sum_{\mathbf{l}} S_{\mathbf{kl}}(\partial_x\hat{\rho}_{\mathbf{l}} + 2\hat{\rho}_{\mathbf{l}}E)\right), \tag{38}$$

where the matrix $S = (S_{\mathbf{kl}})_{K\times K}$ is given by

$$S_{\mathbf{kl}} = \int_{I_{\mathbf{z}}} \frac{1}{\lambda(\mathbf{z})}\psi_{\mathbf{k}}(\mathbf{z})\psi_{\mathbf{l}}(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}.$$

We will demonstrate in the numerical test that (37) is a good approximation of (38), namely, $S = (S_{\mathbf{kl}})_{K\times K} \sim F^{-1} = (F_{\mathbf{kl}})^{-1}_{K\times K}$ with spectral accuracy.

5.2. The fully discretized systems. The relaxation step (33) is stiff; thus we adopt the Wild sum approximation as in [3].

Denote $\mathbf{P}(\hat{\mathbf{r}}) = \mathbf{Q}(\hat{\mathbf{r}}) + \lambda_0 \hat{\mathbf{r}}$ and rewrite (33) as

$$\partial_t \hat{\mathbf{r}} = \frac{1}{\epsilon^2}(\mathbf{P}(\hat{\mathbf{r}}) - \lambda_0 \hat{\mathbf{r}}), \tag{39}$$

$$\partial_t \hat{\mathbf{j}} = -\frac{1}{\epsilon^2}\left(F(v)\hat{\mathbf{j}} + (1-\epsilon^2\phi)\mathbf{d}(\hat{\mathbf{r}})\right). \tag{40}$$

We discretize time using a uniform time step $\Delta t = t^{n+1} - t^n$, where n is the temporal index. Denote by the index $(*)$ the intermediate numerical values obtained after one relaxation step from t^n. Introducing the new variables [10]

$$\tau = 1 - e^{-\lambda_0 t/\epsilon^2}, \quad \hat{\mathbf{R}} = \hat{\mathbf{r}} e^{\lambda_0 t/\epsilon^2},$$

(39) becomes

$$\partial_\tau \hat{\mathbf{R}} = \frac{1}{\lambda_0(1-\tau)}\mathbf{P}(\hat{\mathbf{R}}).$$

By Taylor expansion,

$$\hat{\mathbf{R}}(v,\tau) = \sum_{k=0}^{\infty} \tau^k \hat{\mathbf{r}}^{(k)}(v)$$

with

$$\hat{\mathbf{r}}^{(k+1)}(v) = \frac{1}{k+1}\sum_{n=0}^{k}\frac{1}{\lambda_0}P(\hat{\mathbf{r}}^{(n)}), \quad k \geq 0. \tag{41}$$

We revert to the old notation and truncate the expansion

$$\hat{\mathbf{r}}(v,t) = (1-\tau)\sum_{k=0}^{\infty}\tau^k \hat{\mathbf{r}}^{(k)}(v)$$

by replacing the higher-order terms with the corresponding local equilibrium state $\hat{\mathbf{r}}^{(\infty)} = \hat{\boldsymbol{\rho}}^* M = \hat{\boldsymbol{\rho}}^n M$, which leads to

$$\hat{\mathbf{r}}^* = (1-\tau)\sum_{k=0}^{m}\tau^k \hat{\mathbf{r}}^{(k)} + \tau^{m+1}\hat{\mathbf{r}}^{(\infty)}. \tag{42}$$

The first order time-relaxed scheme (corresponding to $m = 1$) is obtained by using (41) in (42) to give

$$\hat{\mathbf{r}}^* = (1-\tau)\hat{\mathbf{r}}^n + \tau(1-\tau)\frac{\mathbf{P}(\hat{\mathbf{r}}^n)}{\lambda_0} + \tau^2 \hat{\boldsymbol{\rho}}^n M$$

with $\hat{\boldsymbol{\rho}}^n = \int_{\mathbb{R}} \hat{\mathbf{r}}^n(v)dv$.

To update the values for $\hat{\mathbf{j}}$, we solve (40) by the backward Euler method to get $\hat{\mathbf{j}}^\star$,

$$\hat{\mathbf{j}}^\star = \left(I + F(v)\frac{\Delta t}{\epsilon^2}\right)^{-1}\left[\hat{\mathbf{j}}^n - \frac{\Delta t(1-\epsilon^2\phi)}{\epsilon^2}\mathbf{d}(\hat{\mathbf{r}}^\star)\right],$$

where the matrix $I + F(v)\frac{\Delta t}{\epsilon^2}$ is invertible, thanks to the positive definite F.

For notational clarity, we describe the spatial discretization in one space dimension. Consider the spatial domain $\Omega = [x_L, x_R]$ which is partitioned into N grid cells with a uniform mesh size $\Delta x = 1/N$. Define the left boundary x_L as $x_{1/2}$, the right boundary x_R as $x_{N+1/2}$, and choose the spatial grid points $x_{i-1/2} = x_{1/2} + (i-1)\Delta x$, for $i = 1, \ldots, N+1$. The ith interior cell is $[x_{i-1/2}, x_{i+1/2}]$, for $i = 1, \ldots, N$, with the cell average at time level t^n given by

$$\hat{\mathbf{U}}_i^n = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \hat{\mathbf{U}}(t^n, x, v, \mathbf{z}) dx.$$

In the relaxation step, the second-order center difference is used for the spatial derivative. In the transport step, since it is not stiff, we solve it by using an explicit second-order upwind scheme with slope limiters, which need to be implemented on the diagonal form on the Riemann invariants:

$$\hat{\mathbf{U}} = (\hat{\mathbf{r}} + \phi^{-\frac{1}{2}} \hat{\mathbf{j}})/2, \qquad \hat{\mathbf{V}} = (\hat{\mathbf{r}} - \phi^{-\frac{1}{2}} \hat{\mathbf{j}})/2, \tag{43}$$

which solve

$$\begin{cases} \partial_t \hat{\mathbf{U}} + \sqrt{\phi}(v \partial_x \hat{\mathbf{U}} - E \partial_v \hat{\mathbf{U}}) = 0, \\ \partial_t \hat{\mathbf{V}} - \sqrt{\phi}(v \partial_x \hat{\mathbf{V}} - E \partial_v \hat{\mathbf{V}}) = 0. \end{cases}$$

The scheme is given by, for $j = 1, \ldots, N$, if $\lambda_1 = \sqrt{\phi} v \frac{\Delta t}{\Delta x} > 0$,

$$\begin{cases} \hat{\mathbf{U}}_i^{n+1} = \hat{\mathbf{U}}_i^* - \lambda_1 \left[\hat{\mathbf{U}}_i^* - \hat{\mathbf{U}}_{i-1}^* + \frac{\Delta x}{4} (\boldsymbol{\sigma}_i^+ - \boldsymbol{\sigma}_{i-1}^+) \right] + \sqrt{\phi} (\partial_v \hat{\mathbf{U}}_i^*) E_i \Delta t, \\ \hat{\mathbf{V}}_i^{n+1} = \hat{\mathbf{V}}_i^* + \lambda_1 \left[\hat{\mathbf{V}}_{i+1}^* - \hat{\mathbf{V}}_i^* - \frac{\Delta x}{4} (\boldsymbol{\sigma}_{i+1}^- - \boldsymbol{\sigma}_i^-) \right] - \sqrt{\phi} (\partial_v \hat{\mathbf{V}}_i^*) E_i \Delta t. \end{cases}$$

If $\lambda_1 < 0$, the upwind stencil has to be changed. Here $\boldsymbol{\sigma}_i^{\pm}$ is the slope of $\hat{\mathbf{r}} \pm \phi^{-\frac{1}{2}} \hat{\mathbf{j}}$ on the ith cell at the $(*)$th time step, with mth component denoted by $(\sigma_i^{\pm})^{(m)}$ and given by

$$(\sigma_i^{\pm})^{(m)} = \frac{1}{\Delta x} \left[\pm r_{i\pm1}^{(m)} + \phi^{-\frac{1}{2}} j_{i\pm1}^{(m)} \mp r_i^{(m)} - \phi^{-\frac{1}{2}} j_i^{(m)} \right] \psi((\theta_i^{\pm})^{(m)}), \qquad (1 \le m \le K)$$

where $r^{(m)}$, $j^{(m)}$ are the mth components of $\hat{\mathbf{r}}$ and $\hat{\mathbf{j}}$, respectively. $(\theta_i^{\pm})^{(m)}$ is defined by

$$(\theta_i^{\pm})^{(m)} = \left(\frac{r_i^{(m)} \pm \phi^{-\frac{1}{2}} j_i^{(m)} - r_{i-1}^{(m)} \mp \phi^{-\frac{1}{2}} j_{i-1}^{(m)}}{r_{i+1}^{(m)} \pm \phi^{-\frac{1}{2}} j_{i+1}^{(m)} - r_i^{(m)} \mp \phi^{-\frac{1}{2}} j_i^{(m)}} \right)^{\pm}.$$

A simple minmod slope limiter function is chosen,

$$\psi(\theta) = \max\{0, \min\{1, \theta\}\}.$$

One can then update the values for $\hat{\mathbf{r}}_i^{n+1}$ and $\hat{\mathbf{j}}_i^{n+1}$ by using (43).

Now combine the AP property of the time splitting shown in section 5.1 with the AP property of the deterministic scheme given in [10], easily implying that the fully discrete time and space approximations are s-AP, in the sense that its $\epsilon \to 0$ limit, with $\Delta t, \Delta x$ fixed, approaches the fully discrete gPC-SG approximation of the drift-diffusion equation (10). We omit the details.

The velocity discretization is performed using spectral approximation based on the Hermite polynomials, which is equivalent to the moment method. We refer the reader to [19, 10].

5.3. Boundary conditions. Consider the inflow boundary conditions

$$f(t,x,v)|_{x_L} = F_L(v,\mathbf{z}), \qquad f(t,x,-v)|_{x_R} = F_R(v,\mathbf{z}), \quad \text{for } v>0.$$

Then

$$f_k(t,x,v)|_{x_L} = (\vec{F}_1(v))_k, \qquad f_k(t,x,-v)|_{x_R} = (\vec{F}_2(v))_k, \quad \text{for } 1\le k\le K,$$

where

$$(\vec{F}_1(v))_k = \int_{I_{\mathbf{z}}} F_L(v,\mathbf{z})\psi_k(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}, \qquad (\vec{F}_2(v))_k = \int_{I_{\mathbf{z}}} F_R(v,\mathbf{z})\psi_k(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}.$$

By the gPC Galerkin approach (for $v>0$ only),

$$\hat{r}_k + \epsilon\hat{j}_k|_{x_L} = (\vec{F}_1(v))_k, \qquad \hat{r}_k - \epsilon\hat{j}_k|_{x_R} = (\vec{F}_2(v))_k, \tag{44}$$

where r_k, j_k are the kth components of $\hat{\mathbf{r}}$ and $\hat{\mathbf{j}}$, respectively.

As $\epsilon\to 0$, (33) leads to

$$\hat{\mathbf{j}} = -F^{-1}(v)\mathbf{d}(\hat{\mathbf{r}}), \tag{45}$$

which is plugged into (44) to form the numerical boundary conditions. To obtain a second-order numerical boundary conditions, we use two ghost cells outside x_L, namely, $[x_{-3/2}, x_{-1/2}]$ and $[x_{-1/2}, x_{1/2}]$, with average cell values $\hat{\mathbf{U}}_{-1}$, $\hat{\mathbf{U}}_0$ defined respectively. Two ghost cells outside x_R are $[x_{N+1/2}, x_{N+3/2}]$ and $[x_{N+3/2}, x_{N+5/2}]$ with $\hat{\mathbf{U}}_{N+1}$, $\hat{\mathbf{U}}_{N+2}$ are similarly defined. Using the second-order central difference for spatial derivative of $\hat{\mathbf{r}}$ in (44) and (45), we define the ghost cells values $\hat{\mathbf{r}}_{-1}$, $\hat{\mathbf{r}}_0$ by the following equations:

$$\begin{aligned}
&\frac{\hat{\mathbf{r}}_{-1}+\hat{\mathbf{r}}_2}{2} - \epsilon\left[F^{-1}(v)\left(\frac{v}{3\Delta x}(\hat{\mathbf{r}}_2-\hat{\mathbf{r}}_{-1}) - E(x_{1/2})\nabla_v\hat{\mathbf{r}}(x_{1/2})\right)\right] = \vec{F}_1,\\
&\frac{\hat{\mathbf{r}}_0+\hat{\mathbf{r}}_1}{2} - \epsilon\left[F^{-1}(v)\left(\frac{v}{\Delta x}(\hat{\mathbf{r}}_1-\hat{\mathbf{r}}_0) - E(x_{1/2})\nabla_v\hat{\mathbf{r}}(x_{1/2})\right)\right] = \vec{F}_1,
\end{aligned} \tag{46}$$

where $\nabla_v\hat{\mathbf{r}}$ is approximated by

$$\nabla_v\hat{\mathbf{r}}(x_{1/2}) \approx \int_{I_{\mathbf{z}}} (\nabla_v F_L)\boldsymbol{\psi}(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}.$$

Notice that the matrix $\frac{1}{2}I + \frac{\epsilon v}{\Delta x}F^{-1}$ is positive definite, so the above equations are solvable. For $\epsilon \ll 1$, (46) is an $O(\epsilon^2)$ approximation to (44) and (45). Boundary values for $\hat{\mathbf{j}}$ can be obtained by using approximation to (45),

$$\begin{aligned}
\hat{\mathbf{j}}_{-1} &= -2F^{-1}(v)\left(\frac{v}{3\Delta x}(\hat{\mathbf{r}}_2-\hat{\mathbf{r}}_{-1}) - E(x_{1/2})\nabla_v\hat{\mathbf{r}}(x_{1/2})\right) - \hat{\mathbf{j}}_2,\\
\hat{\mathbf{j}}_0 &= -2F^{-1}(v)\left(\frac{v}{\Delta x}(\hat{\mathbf{r}}_1-\hat{\mathbf{r}}_0) - E(x_{1/2})\nabla_v\hat{\mathbf{r}}(x_{1/2})\right) - \hat{\mathbf{j}}_1.
\end{aligned}$$

The boundary condition at x_R is treated similarly.

5.4. The stochastic collocation method. The stochastic collocation (SC) method will be employed for numerical comparison. Let $\{\mathbf{z}^{(j)}\}_{j=1}^{N_s} \subset I_{\mathbf{z}}$ be the set of collocation nodes and N_s the number of samples. For each fixed individual sample $\mathbf{z}^{(j)}$, $j = 1, \ldots, N_s$, one applies the AP scheme to the deterministic equations as in [10], obtains the solution ensemble $u_j(t, x, v) = u(t, x, v, \mathbf{z}^{(j)})$, then adopts an interpolation approach to construct a gPC approximation, such as

$$u(t, x, v, \mathbf{z}) = \sum_{j=1}^{N_s} u_j(t, x, v) l_j(\mathbf{z}),$$

where $l_j(\mathbf{z})$ depends on the construction method. The Lagrange interpolation method is used here by choosing $l_j(\mathbf{z}^{(i)}) = \delta_{ij}$. This is straightforward and brings no coding difficulty. An overview of such an SC method can be found in [20].

For kinetic equations with uncertainy and diffusive scaling, so far there has not been any analysis conducted for the s-AP property of the SC method. The behavior of the SC method for such problems can be understood, at least formally, as follows. If one uses a deterministic AP scheme, such as the one in [10], for each sample $z^{(j)}$, according to the typical behavior of an AP scheme [8], one could expect a uniform convergence in ϵ, with a convergence rate, say, $O(\delta^{\alpha})$, where δ is the numerical parameter (time step, mesh size in space or velocity, etc.). If one uses a βth order interpolation or quadrature rule in the SC method, there could be an *additional* interpolation error of $O(N_s^{-\beta})$, from the regularity estimate of Theorem 2.2. So the overall error is $(\delta^{\alpha} + N_s^{-\beta})$, which is independent of ϵ. This shows that the SC method is s-AP.

In the next section, we will use the SC results to compare with the results of the s-AP gPC-SG method.

6. Numerical examples. In this section, several numerical tests are shown to illustrate the validity and effectiveness of our stochastic AP scheme. Randomness will arise from collision kernel, initial data, or boundary data. For simplicity, we will always assume the random variable $\mathbf{z}$ obeys a uniform distribution, defined on $[-1, 1]^n$ with n up to 2, so the Legendre polynomial gPC basis is used.

Often one is only interested in the solution statistics, such as the mean and standard deviation of the macroscopic physical quantities. The macroscopic quantities ρ, $\mathbf{u}$ that stand for density and bulk velocity are defined by

$$\rho = \int_{\mathbb{R}} f(v) dv, \qquad u = \frac{1}{\rho} \int_{\mathbb{R}} f(v) v dv.$$

Throughout this paper, we deal with the integral in v by using the Hermite quadrature rule for N_v positive quadrature points $(v_1, \ldots, v_{N_v})$ with the corresponding weights $(\omega_1, \ldots, \omega_{N_v})$. In all of our computations, we take $N_v = 8$.

Given the gPC coefficients $f_{\mathbf{k}}$ of f, the statistical mean, variance, and standard deviation are

$$[f] \approx f_0, \qquad \mathrm{Var}[f] \approx \sum_{|\mathbf{k}|=1}^{M} f_{\mathbf{k}}^2, \qquad \mathcal{S}[f] = \sqrt{\sum_{|\mathbf{k}|=1}^{M} f_{\mathbf{k}}^2}.$$

To measure the difference in the solutions, we use the difference in mean and in standard deviation, with L^2 norm in x,

$$\mathcal{E}_{\mathrm{mean}}(t) = \left|\left| E[u^h] - E[u] \right|\right|_{L^2},$$

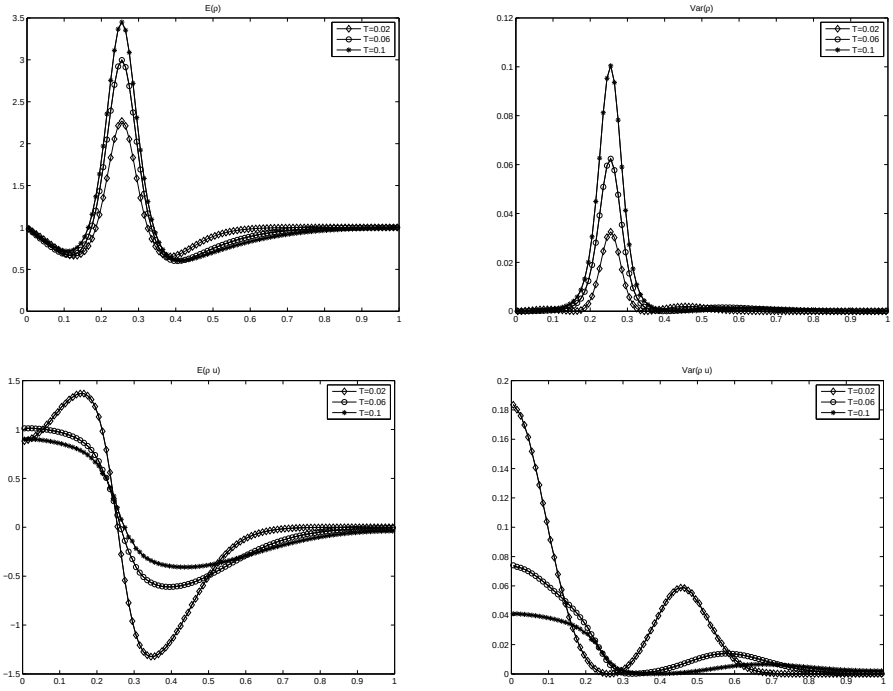

FIG. 1. *Test* I *with* $\sigma(z) = 2 + 2$. *Mean and variance of density* ρ *and momentum density* ρu *at different time* T. $\Delta x = 0.01$, $\Delta t = 5 \times 10^{-5}$, $\epsilon = 0.001$. *Star: gPC-SG with* $K = 4$. *Solid line: the reference solutions by the collocation method using* $N_z = 16$.

$$\mathcal{E}_{\text{std}}(t) = \left|\left|\sigma[u^h] - \sigma[u]\right|\right|_{L^2},$$

where u^h and u are respectively numerical solutions of gPC-SG and the reference solutions obtained by the high-order collocation method.

For the collocation method, with samples $\{\mathbf{z}^{(j)}\}$ and corresponding weights $\{w^{(j)}\}$ chosen from the quadrature rule, the integrals are approximated by

$$\int_{I_{\mathbf{z}}} f(t,x,v,\mathbf{z})\pi(\mathbf{z})d\mathbf{z} \approx \sum_{j=1}^{N_s} f(t,x,v,\mathbf{z}^{(j)})w^{(j)}.$$

For all the numerical examples below, we choose $G = 0$.

6.1. Test I: A random collision kernel. We first consider a problem with a random collision kernel, with

$$x \in [0,1], \quad F_L(v) = M(v), \quad F_R(-v) = M(v),$$
$$\phi = \exp\left(-50\exp(1)(1/4 - x)^2\right), \quad \epsilon = 0.001.$$

The initial distribution is $f(x,v,t=0) = M(v)$. $\sigma(z) = 2 + z \geq 0$ and $\sigma(\mathbf{z}) = 2 + 0.3z_1 + 0.7z_2$ for the one- and two-dimensional tests, respectively.

Lacking the analytic solutions, we use the high-order SC method with 16 Legendre–Gauss quadrature points to obtain the reference solutions. A satisfactory agreement between gPC-SG solutions and the reference solutions are observed at different time T, for both one- and two-dimensional random variables (see Figures 1 and 2 respectively).

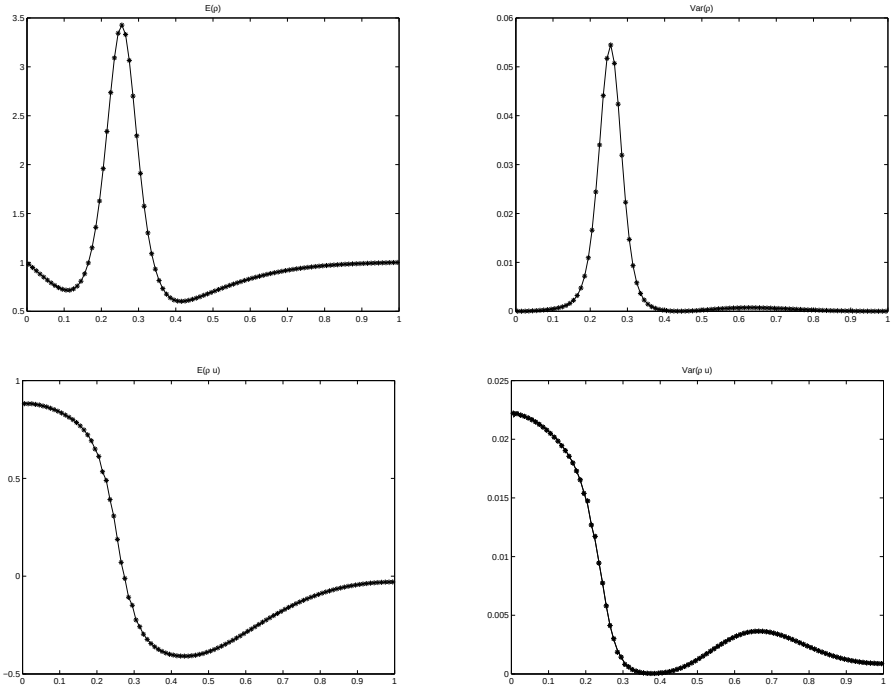

FIG. 2. *Test* I *with* $\sigma(\mathbf{z}) = 2 + 0.3z_1 + 0.7z_2$. *Mean and variance of* ρ *and* ρu. $T = 0.1$, $\Delta x = 0.01$, $\Delta t = 10^{-4}$, $\epsilon = 0.001$. *Star: gPC-SG with* $K = 4$. *Solid line: the reference solutions.*

In Figure 3, we show the differences in the $L^2(x, v)$ norm between the fourth-order gPC-SG solutions f and ρM for Test I. The errors are defined by

$$
\begin{aligned}
||f - \rho M||^2_{L^2(x,v)} &= \int_{I_z} \int_{\mathbb{R}^d} \int_{\Omega} |f - \rho M|^2 \, dx dv \pi(z) dz \\
&\approx \sum_{k=0}^{K} \sum_{i=1}^{N_x} \sum_{j=1}^{2N_v} \left(r_k(x_i, v_j) + \epsilon j_k(x_i, v_j) - \rho_k(x_i) M(v_j)\right)^2 \omega_j \Delta x. \qquad (47)
\end{aligned}
$$

We only compute $r_k(v_j)$, $j_k(v_j)$ for $j = 1, \ldots, N_v$ in our algorithm. One can get $r_k(-v_j)$, $j_k(-v_j)$ by using the even and odd property of r_k and j_k, and $\rho_k(x_i) = \sum_{j=1}^{N_v} r_k(x_i, v_j)\omega_j$. The difference is $\mathcal{O}(\epsilon)$ before it saturates and the numerical errors due to space, time, and velocity discretizations start to dominate.

In Figure 4, we compare mean and variance of ρ between the third-order gPC-SG solutions and that of the diffusion system using (38). One observes a satisfactory agreement between the two solutions.

In Figure 5, we show a log-log plot of the L^2 error in space for mean and variance of ρ between fourth-order gPC-SG solutions and that obtained by solving the diffusion system via (38). One observes that the difference is of $\mathcal{O}(\epsilon)$ and it saturates as the numerical errors due to space, time, and velocity discretizations dominate. This result indicates that K can be chosen independent of ϵ, verifying the stochastic AP property.

In Figure 6, we compare the limiting system for gPC-SG scheme (37) and the Galerkin system for the diffusion equation given by (38). The analytic solutions of two equations are unknown, so the forward Euler in time and the central difference

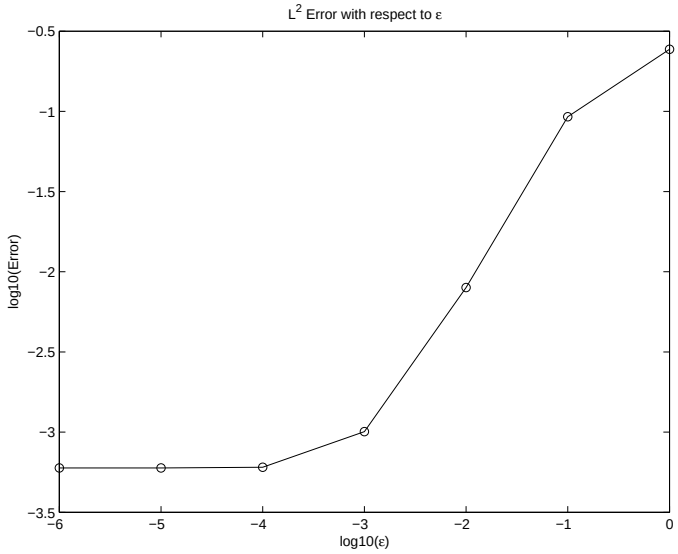

FIG. 3. *Test* I. L^2 *norm of* $f-\rho M$ *defined in* (47)*, with respect to different* ϵ*.* $K=4$*,* $T=0.05$*,* $\Delta x=0.01$*,* $\Delta t=5\times 10^{-5}$*.*

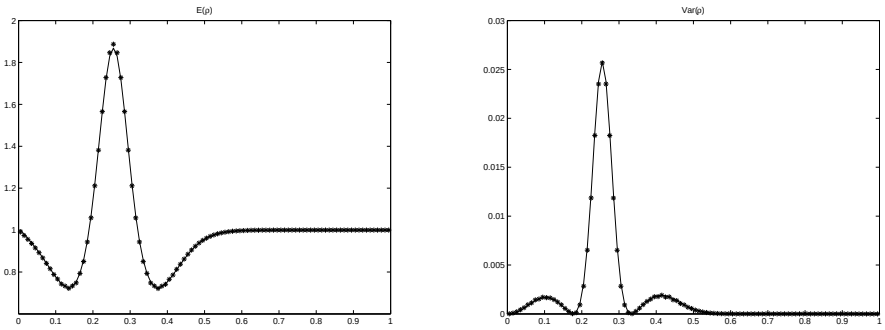

FIG. 4. *Test* I. *Mean and variance of* ρ *of third-order gPC-SG solutions (star) and reference ones (solid line) obtained by using* (38) *for the limiting diffusion equations.* $T=0.01$*,* $\Delta x=0.01$*,* $\Delta t=10^{-5}$*,* $\epsilon=10^{-4}$*.*

scheme in space are used to compute the approximated solutions. One can observe that two sets of solutions are in good agreement.

Figure 7 shows a fast exponential decay of L^2 error in space for mean and variance of ρ by solving two diffusion systems with respect to different gPC orders K. These two figures give an indication of $S\sim F^{-1}$ with spectral accuracy, as discussed in section 5.1.

Figure 8 compares the errors between the gPC-SG solutions and the reference solutions at different time T. This result shows that the errors of physical quantities remain the same order of magnitude in time, which indicates the error estimate in Theorem 4.3 may not be sharp.

Figure 9 gives a semi-log plot of the errors of mean and variance of physical

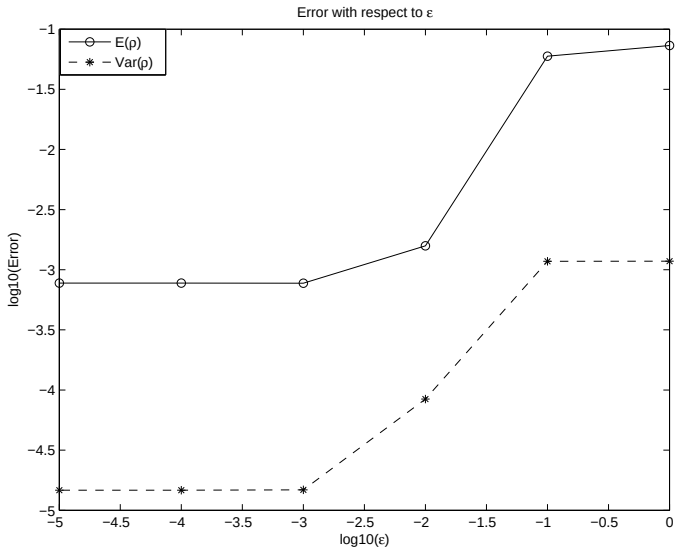

FIG. 5. *Test* I. L^2 *error in space for mean and variance of* ρ*, between fourth-order gPC-SG solutions using* $\Delta x = 0.01$*,* $\Delta t = 5 \times 10^{-5}$ *and that obtained by solving the diffusion system via* (38)*, using* $\Delta x = 0.005$*,* $\Delta t = 10^{-6}$*. We compute the error with respect to* ϵ *at* $T = 0.05$*.*

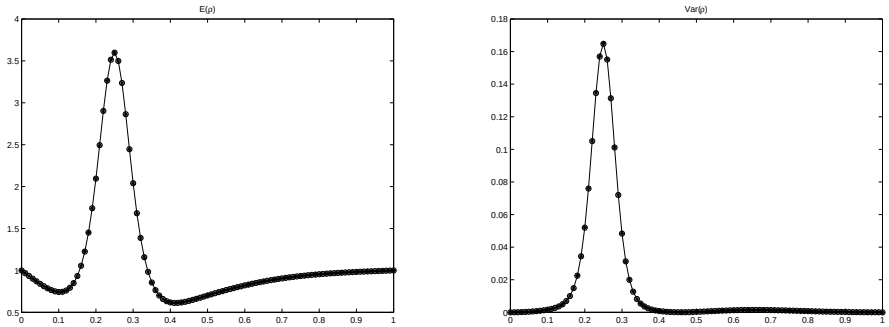

FIG. 6. *Test* I. *Mean and variance of* ρ*, obtained from two limiting diffusion systems* (37) *(circle) and* (38) *(star).* $T = 0.1$*,* $\Delta x = 0.01$*,* $\Delta t = 5 \times 10^{-5}$ *with gPC order* $K = 4$*.*

quantities at $\epsilon = 10^{-5}$, $T = 0.05$ and $T = 0.1$, $\Delta t = 5 \times 10^{-5}$, $\Delta x = 0.01$, with respect to different gPC orders, even for moderate to very small ϵ. We demonstrate a fast exponential convergence with respect to increasing K. The errors quickly saturate at modest gPC orders when $K = 4$, which indicates that the errors arising from the temporal and spatial discretization contribute more than those from the gPC expansion. This result verifies again the stochastic AP property that one can choose K independent of ϵ.

6.2. Test II: Random initial data. Now we consider the case involving random initial data, with

$$x \in [0,1], \quad \phi = \exp\left(-50\exp(1)(1/4 - x)^2\right), \quad \sigma(v,w) = 1,$$

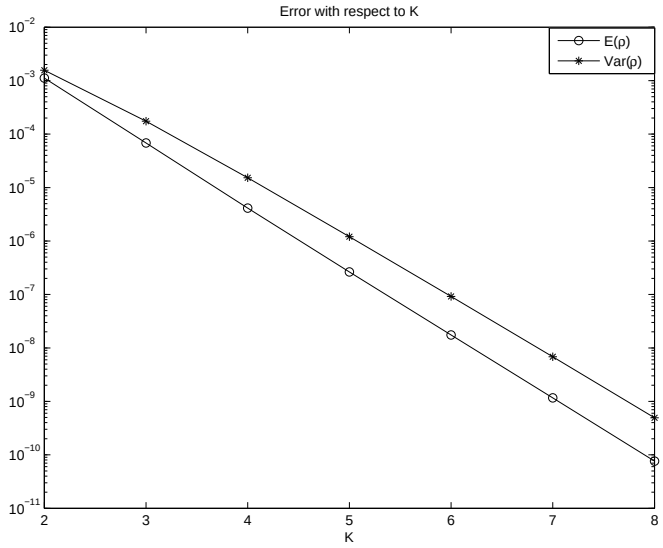

FIG. 7. *Test* I. L^2 *error in space for mean and variance of* ρ *obtained by solving two limiting diffusion systems* (37) *and* (38). $T = 0.05$, $\Delta x = 0.01$, $\Delta t = 5 \times 10^{-5}$.

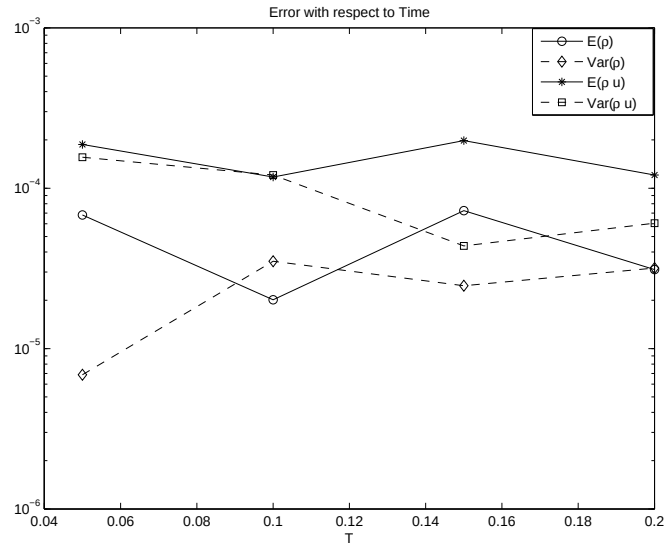

FIG. 8. *Test* I. L^2 *error in space between gPC-SG and the reference solutions for mean and variance of* ρ *and* ρu *at different time* $T = 0.05, 0.1, 0.15, 0.2$. $\Delta x = 0.01$, $\Delta t = 5 \times 10^{-5}$, $\epsilon = 0.001$.

$$f^0(x, v, z) = \frac{\rho^0(z)}{\sqrt{\pi}} e^{-v^2}, \quad \rho^0(z) = \begin{cases} z + 1 & \text{for } z \in [-1, 0], \\ -z + 1 & \text{for } z \in (0, 1]. \end{cases} \tag{48}$$

A periodic boundary condition is assumed in x. Figure 10 shows fourth-order gPC-SG solutions and the reference solutions obtained by the collocation method. The two solutions are in good agreement at different time T.

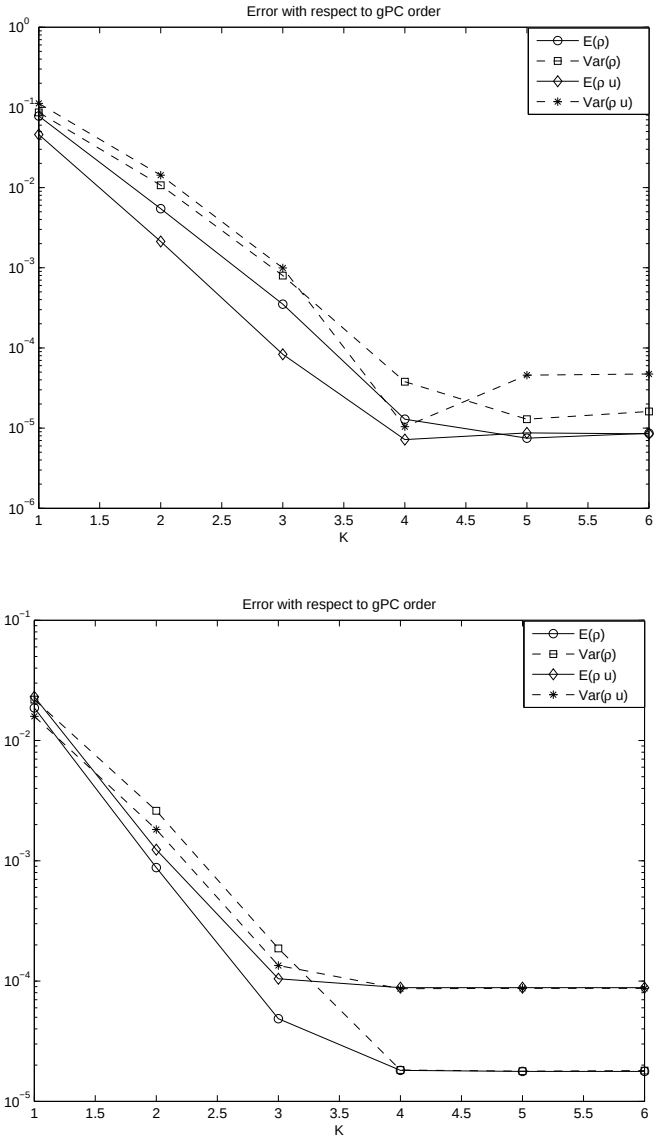

FIG. 9. *Test* I. L^2 *error in space between gPC-SG and the reference solutions for the mean and variance of* ρ *and* ρu *with respect to gPC order* K *at* $T = 0.1$, $\Delta x = 0.01$, $\Delta t = 5 \times 10^{-5}$. $\epsilon = 10^{-5}$ *and* $\epsilon = 10^{-2}$ *in the top and bottom figures, respectively.*

Now $\rho^0(z)$ in (48) is a continuous function in z on $[-1, 1]$, but $\partial_z \rho^0(z)$ is discontinuous. To check how regularity in z of the initial data affects the convergence rate, we compare the result with a C^∞ initial data, namely,

$$f^0(x, v, z) = \frac{\rho^0(z)}{\sqrt{\pi}} e^{-v^2}, \quad \rho^0(z) = \sin\left[\frac{\pi}{2}(z+1)\right]. \tag{49}$$

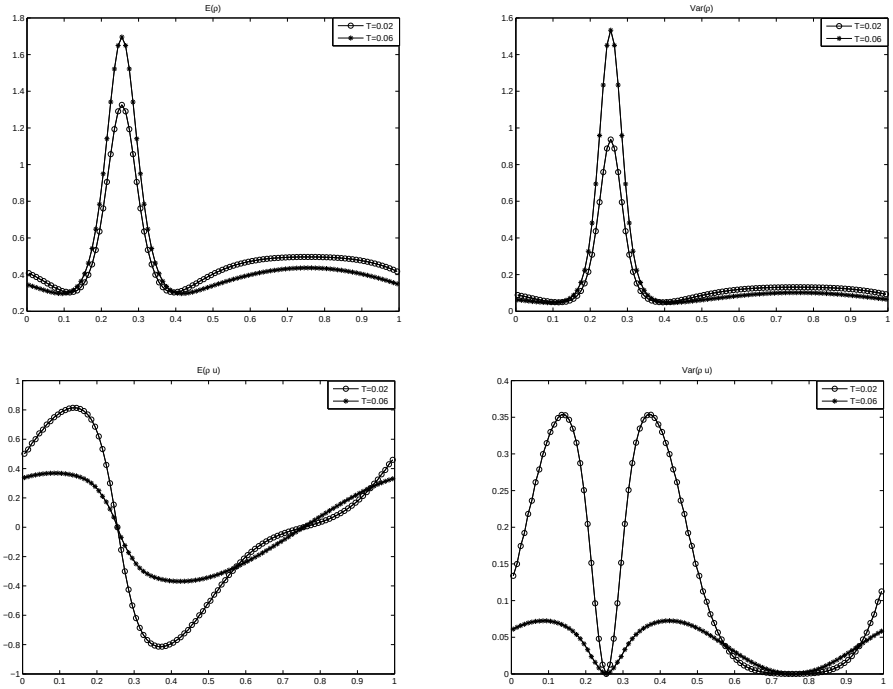

FIG. 10. *Test* II. *Mean and variance of* ρ *and* ρu *at different time* T. $\Delta x = 0.01$, $\Delta t = 10^{-5}$, $\epsilon = 0.001$. *Star: gPC-SG with* $K = 4$. *Solid line: the reference solutions.*

We follow by comparing the difference in f between Kth- and $\frac{K}{2}$th-order gPC-SG solutions. The error is measured by

$$\begin{aligned}\left|\left|f_K - f_{\frac{K}{2}}\right|\right|^2_{L^2(x,v)} &= \int_{I_z}\int_\Omega\int_{\mathbb{R}} (f_K - f_{\frac{K}{2}})^2 dv dx \pi(z) dz \\ &\approx \sum_{i=1}^{N_x}\sum_{j=1}^{2N_v}\left[\sum_{k=0}^{K/2}\left(\alpha_k(x_i,v_j) - \tilde{\alpha}_k(x_i,v_j)\right)^2 + \sum_{K/2+1}^{K}\alpha_k(x_i,v_j)^2\right]\omega_j \Delta x, \end{aligned} \tag{50}$$

where α_k $(k = 0, \ldots, K)$ and $\tilde{\alpha}_k$ $(k = 0, \ldots, K/2)$ are the gPC coefficients of f_K and $f_{K/2}$, respectively. $\alpha_k(v_k)$ for $v_k < 0$ are similarly defined as in (47).

Figure 11 shows the error defined in (50) versus gPC order K in a log-log scale. For initial data given in (48), it shows a linear decay in K. A straight line with slope 1 is included as the reference. For initial data given in (49), a spectral convergence of the solution in K is observed. This validates our analysis in Theorem 4.3, which suggests that the decay rate in K depends on the regularity of the initial data in z.

6.3. Test III: Random boundary data. In this test, we assume some randomness for the boundary conditions.

$$x \in [0,1], \quad F_L(v) = (1 + 0.5z)M(v), \quad F_R(-v) = (1 + 0.5z)M(v),$$
$$\sigma(v,w) = 1, \quad \phi = \exp\left(-50\exp(1)(1/4 - x)^2\right).$$

Consider the initial distribution $f(x, v, t = 0) = M(v)$. In this numerical test, the inflow boundary conditions involve a random fluctuation around the equilibrium at

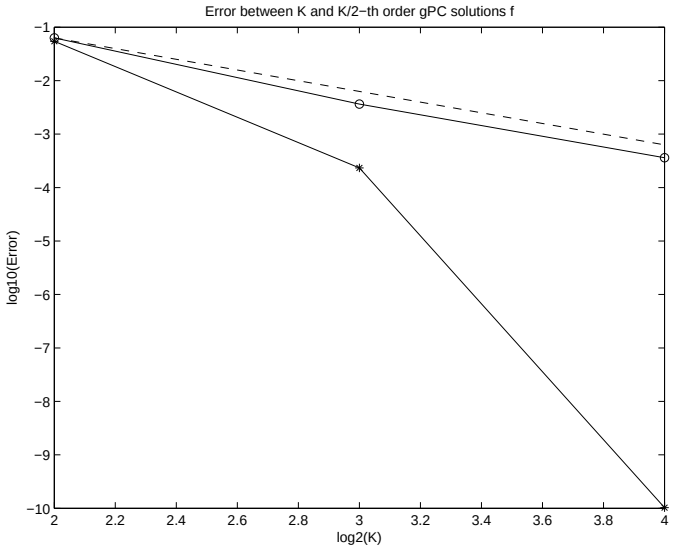

FIG. 11. *Test* II *using* $\sigma = 2 + z$. *Error of* f *between* K*th- and* $\frac{K}{2}$*th-order gPC-SG solutions, defined in* (50). $T = 0.05$, $\Delta x = 0.02$, $\Delta t = 5 \times 10^{-5}$, $\epsilon = 10^{-3}$. *Solid line with circles: the solutions with initial data* (48). *Solid line with stars: the solutions with initial data* (49). *Dotted line is a straight line with slope* 1 *as the reference solution.*

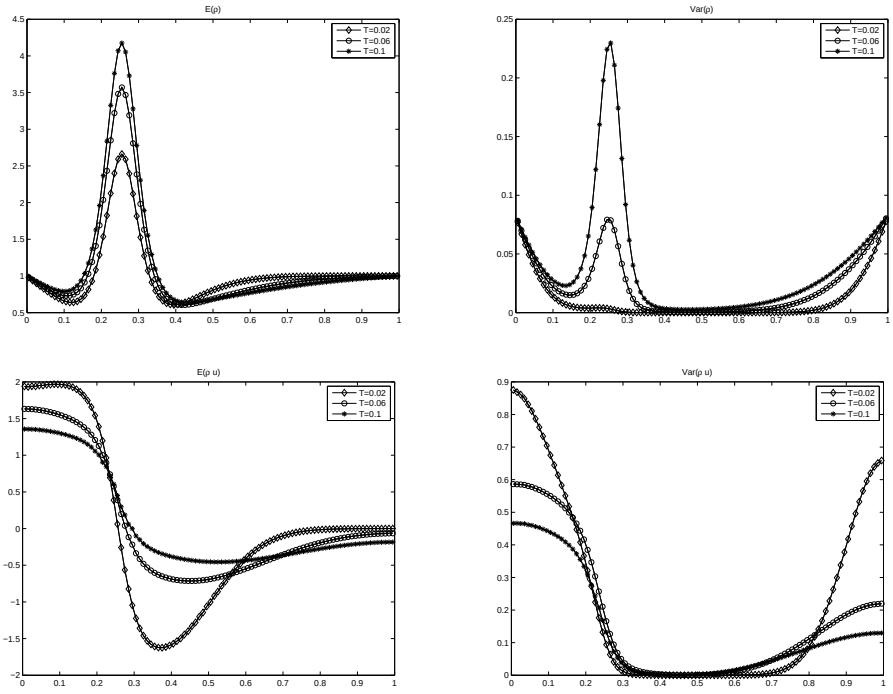

FIG. 12. *Test* III. *Mean and variance of density* ρ *and momentum density* ρu *at different time* T *with* $\Delta x = 0.01$, $\Delta t = 5 \times 10^{-5}$, $\epsilon = 0.001$.

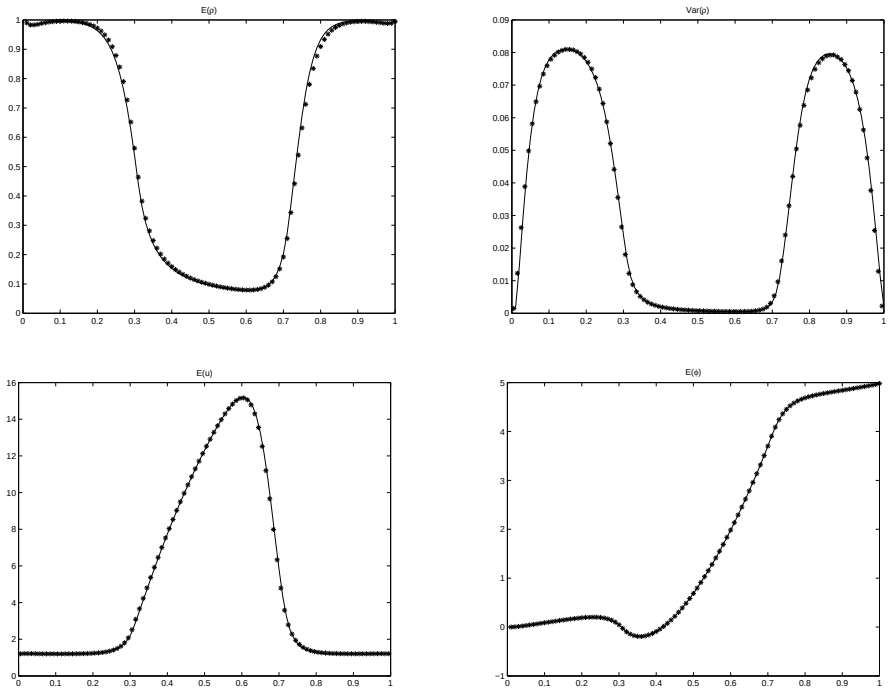

FIG. 13. *Test* IV. *First row: mean and variance of* ρ. *Second row: mean of velocity* u *and potential* ϕ. $T = 0.05$, $\Delta x = 0.01$, $\Delta t = 10^{-5}$, $\epsilon = 0.001$. *Star: gPC-SG with* $K = 4$. *Solid line: the reference solutions.*

both boundaries. The fourth-order gPC solutions and reference solutions are given in Figure 12.

6.4. Test IV: The Boltzmann–Poisson equation with random parameters. We consider the case where the electric potential is given by the solution of a Poisson equation, namely,

$$\beta(z)\partial_{xx}\phi = \rho - c(x,z), \qquad \phi(0) = 0, \qquad \phi(1) = V,$$

where $\beta(z)$ is the scaled Debye length, V the applied bias voltage, and $c(x,z)$ the doping profile. In this test, we involve random perturbations on the parameters $\beta(z)$ and $c(x,z)$, which are given by $\beta(z) = 0.002(1 + 0.5z)$, and

$$c(x,z) = \left(1 - (1-M)\rho(0,t=0)\left[\tanh\left(\frac{x-x_1}{s}\right) - \tanh\left(\frac{x-x_2}{s}\right)\right]\right)(1+0.5z)$$

with $s = 0.02$, $M = (1 - 0.001)/2$, $x_1 = 0.3$, $x_2 = 0.7$, and $V = 5$. The numerical data is chosen from [10], with

$$x \in [0,1], \quad F_L(v) = M(v), \quad F_R(-v) = M(v),$$
$$f(x,v,t=0) = M(v), \quad \epsilon = 0.001.$$

The electric field $E = -\partial_x\phi$ is obtained by central difference approximation. The fourth-order gPC solutions and the reference solutions are shown in Figure 13, and they are in good agreement.

Acknowledgment. The authors thank the two referees for their constructive comments which helped to improve the quality of the paper.

REFERENCES

[1] E. F. A. K. Prinja and J. S. Warsa, *Stochastic methods for uncertainty quantification in radiation transport*, in Proceedings of the International Conference on Mathematics, Computational Methods & Reactor Physics, 2009.

[2] C. Cercignani, *The Boltzmann Equation and Its Applications*, Appl. Math. Sci., Springer-Verlag, New York, 1988, https://dx.doi.org/10.1007/978-1-4612-1039-9.

[3] E. Gabetta, L. Pareschi, and G. Toscani, *Relaxation schemes for nonlinear kinetic equations*, SIAM J. Numer. Anal., 34 (1997), pp. 2168–2194.

[4] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991, https://dx.doi.org/10.1007/978-1-4612-3094-6.

[5] 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, https://dx.doi.org/10.1017/S0962492914000075.

[6] J. Hu and S. Jin, *A stochastic Galerkin method for the Boltzmann equation with uncertainty*, J. Comput. Phys., 315 (2016), pp. 150–168.

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

[8] S. Jin, *Asymptotic preserving schemes for multiscale kinetic and hyperbolic equations: A review*, Riv. Mat. Univ. Parma, 3 (2012), pp. 177–216.

[9] S. Jin, J.-G. Liu, and Z. Ma, *Uniform Spectral Convergence of the Stochastic Galerkin Method for the Linear Transport Equations with Random Inputs in Diffusive Regime and a Micro-Macro Decomposition Based Asymptotic Preserving Method*, preprint, 2016.

[10] S. Jin and L. Pareschi, *Discretization of the multiscale semiconductor Boltzmann equation by diffusive relaxation schemes*, J. Comput. Phys., 161 (2000), pp. 312–330, https://dx.doi.org/10.1006/jcph.2000.6506.

[11] 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.

[12] S. Jin, D. Xiu, and X. Zhu, *Asymptotic-preserving methods for hyperbolic and transport equations with random inputs and diffusive scalings*, J. Comput. Phys, 289 (2014), pp. 35–52.

[13] A. Klar, *An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit*, SIAM J. Numer. Anal., 35 (1998), pp. 1073–1094, https://dx.doi.org/10.1137/S0036142996305558.

[14] E. W. Larsen and J. E. Morel, *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes II*, J. Comput. Phys., 83 (1989), pp. 212–236.

[15] O. P. Le Maître and O. M. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Sci. Comput., Springer, New York, 2010, https://dx.doi.org/10.1007/978-90-481-3520-2.

[16] M. Lo'eve, *Probability Theory*, 4th ed., Springer-Verlag, New York, 1977.

[17] P. Markowich, C. Ringhofer, and C. Schmeiser, *Semiconductor Equations*, Springer-Verlag, New York, 1989.

[18] E. Poupaud, *Diffusion approximation of the linear semiconductor Boltzmann equation: Analysis of boundary layers*, Asymptot. Anal., 4 (1991), pp. 293–317.

[19] C. Ringhofer, C. Schmeiser, and A. Zwirchmayr, *Moment methods for the semiconductor Boltzmann equation on bounded position domains*, SIAM J. Numer. Anal., 39 (2001), pp. 1078–1095.

[20] D. Xiu, *Numerical Methods for Stochastic Computations*, Princeton University Press, Princeton, NJ, 2010.

[21] D. Xiu and J. S. Hesthaven, *High-order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput., 27 (2005), pp. 1118–1139, https://dx.doi.org/10.1137/040615201.

[22] D. Xiu and J. Shen, *Efficient stochastic Galerkin methods for random diffusion equations*, J. Comput. Phys., 228 (2009), pp. 226–281.

[23] A. Zwirchmayr and C. Schmeiser, *Convergence of moment methods for linear kinetic equations*, SIAM J. Numer. Anal., 36 (1998), pp. 74–88.