FAST CONVERGENCE AND ASYMPTOTIC PRESERVING OF THE GENERAL SYNTHETIC ITERATIVE SCHEME [*]

WEI SU[†], LIANHUA ZHU[†], AND LEI WU[‡]

Abstract. Recently the general synthetic iteration scheme (GSIS) is proposed to find the steady-state solution of the Boltzmann equation [22], where various numerical simulations have shown that (i) the steady-state solution can be found within dozens of iterations at any Knudsen number K, and (ii) the solution is accurate even when the spatial cell size in the bulk region is much larger than the molecular mean free path, i.e. Navier-Stokes solutions are recovered at coarse grids. The first property indicates that the error decay rate between two consecutive iterations decreases to zero with K, while the second one implies that the GSIS is asymptotically preserving the Navier-Stokes limit. This paper is dedicated to the rigorous proof of both properties.

Key words. gas kinetic equation, convergence rate, asymptotic preserving

AMS subject classifications. 76P05, 65L04, 65M12

1. Introduction. Rarefied gas flows are encountered in many engineering problems, including space vehicles' reentry, micro-electromechanical systems, and shale gas extraction. Although several macroscopic equations are proposed [4, 8, 17], they only work when the Knudsen number K, defined as the ratio of the molecular mean free path to the characteristic system length, is small. To describe the rarefied gas dynamics from the continuum to free molecular flow regimes, the Boltzmann equation from the gas kinetic theory is needed. However, it is a grand challenge to solve the Boltzmann equation, as the numerical scheme should be carefully designed, especially when K is small so that the Boltzmann collision operator becomes stiff. For example, in the direct simulation Monte Carlo method [3], it is required that the numerical scale (i.e. the spatial cell size Δx and time step Δt) is smaller than the corresponding kinetic scale (i.e. the mean free path and mean collision time of gas molecules), otherwise the numerical dissipation will lead to inaccurate solutions. This would become prohibitive in the multiscale simulation involving the near-continuum flow where the Knudsen number is small.

Many efforts have been devoted to solving the Boltzmann equation or its simplified model equations under a numerical scale larger than the kinetic one [15, 6, 7, 5, 12, 28, 9]. The common feature of these numerical methods is that they are time-dependent. Some schemes asymptotically preserve the Euler limit, i.e. they become a consistent discretization of Euler equations when K goes to zero [6, 7], while some asymptotically preserve the Navier-Stokes limit when Δt is much larger than the mean collision time [5, 12], or when both Δt and Δx are much larger than the corresponding kinetic scales [28, 9, 10].

It should be noted that in continuum flows the turbulence is common and time-dependent numerical solvers for Navier-Stokes equations are highly demanded. However, in rarefied gas dynamics the turbulence is absent and steady-state solutions are

[†]James Weir Fluids Laboratory, Department of Mechanical and Aerospace Engineering, University of Strathclyde, Glasgow G1 1XJ, UK

[‡]Department of Mechanics and Aerospace Engineering, Southern University of Science and Technology, Shenzhen 518055, China (wul@sustech.edu.cn)

of particular interest. Therefore, the implicit solver is strongly recommended, especially in multiscale problems where the Knudsen layer (a few mean free path away from the solid surface) must be resolved (see the thermal flow in section 5.4 for an example). In designing the implicit solver, one has to consider the following two major problems: (i) how to find the steady-state solution quickly and (ii) how to recover the solution of Navier-Stokes equations in the continuum flow regime, on coarse spatial grids that are adequate to capture the hydrodynamics but much larger than the mean free path?

A simple way to obtain the steady-state solution of gas kinetic equations is to use the conventional iterative scheme (CIS), where the collision operator is splitted into the gain term and the loss term. The streaming operator and loss term are calculated at the current iteration step, while the gain term is evaluated at the previous iteration step. Therefore, the resultant ordinary differential equation can be effectively solved by the sweep technique after discretization in physical space, without involving the inverse of large-size matrix [11]. The CIS is efficient and accurate when the Knudsen number is large, where converged solution can be obtained within a few iterations. However, the iteration number increases significantly when K is small. Moreover, the CIS does not asymptotically preserve the Navier-Stokes limit, so a huge number of spatial cells should be used to capture the flow dynamics when K is small [25].

It is noted that these two major deficiencies of CIS not only appear in gas system, but also in other kinetic systems for neutron transport, thermal radiation, and phonon transport in crystals, to name just a few. The exciting breakthrough to tackle these deficiencies is firstly achieved probably in neutron transport, where the diffusion synthetic acceleration (DSA) is developed, see the review [1]. In DSA, the mesoscopic kinetic equation is solved together with a diffusion equation for the macroscopic variable appearing in the gain part of the collision operator. This diffusion equation is exactly the limiting equation derived from the kinetic equation in optical thick regions (or equivalently when K is small) [14]. Not only fast convergence to steady state but also asymptotic preserving are achieved on coarse physical grids, in the whole range of Knudsen number.

The DSA scheme has been extended to some special rarefied gas flows, such as the Poiseuille [24, 27] and Couette [19] flows, where the flow velocity appearing in the collision operator is perpendicular to the computational domain, thus its macroscopic governing equation is an inhomogeneous diffusion equation. In the continuum flow regime the inhomogeneous term vanishes and the macroscopic equation is reduced to the diffusion equation, while in rarefied regimes the inhomogeneous term dominates and captures the rarefaction effects exactly. However, for general rarefied gas flows, the limiting macroscopic equations are not the diffusion equation but the Navier-Stokes equation as according to the Chapman-Enskog expansion [4]. Thus, in a recent paper we proposed the general synthetic iterative scheme (GSIS), where the macroscopic synthetic equations take the form of linearized Navier-Stokes equations but with source terms describing high-order rarefaction effects [22].

Since the synthetic equations contain the continuity equation (obtained from mass conservation) that is not of diffusion type, GSIS may have different behaviors when compared to other DSA schemes. Although the numerical examples in Ref. [22] have shown that GSIS has the properties of fast convergence to steady state and asymptotic Navier-Stokes preserving, the mathematical prove is lacking. It is therefore the aim of this paper to rigorously investigate the two properties of GSIS.

The remainder of this paper is organized as follows. In sections 2 and 3 the convergence rates of CIS and GSIS for the linearized BGK kinetic equation [2] are

calculated, respectively. The reason of explicitly including Newton's law of shear stress and Fourier's law of heat conduction in the macroscopic synthetic equations is analyzed. In section 4 the asymptotic preserving property of GSIS is studied and it is found that GSIS preserves the Navier-Stokes limit with $\Delta x \sim O(1)$, as long as this spatial resolution is able to capture the hydrodynamic behavior. In section 5, several numerical simulations, including the most challenging simulation of thermally-induced flow in the near-continuum regime, are preformed to show the fast convergence and asymptotic Navier-Stokes preserving properties of the GSIS. A summary is given in section 6.

2. The convergence rate of CIS and false convergence. To make the mathematical derivation simple but keep the essential flow physics, we consider the linearized BGK equation in general two-dimensional problems:

$$(2.1) \qquad \boldsymbol{v} \cdot \frac{\partial h(\boldsymbol{x}, \boldsymbol{v})}{\partial \boldsymbol{x}} = \frac{h_{eq}(\boldsymbol{x}, \boldsymbol{v}) - h(\boldsymbol{x}, \boldsymbol{v})}{K},$$

where $h(\boldsymbol{x}, \boldsymbol{v})$ is the velocity distribution function at location $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{v} = (v_1, v_2)$ is the molecular velocity, K is the Knudsen number, and h_{eq} is the reference velocity distribution function:

$$(2.2) \qquad h_{eq}(\boldsymbol{x}, \boldsymbol{v}) = \varrho(\boldsymbol{x}) + 2\boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v} + \tau(\boldsymbol{x})\left(v^2 - 1\right),$$

with ϱ, $\boldsymbol{u} = (u_1, u_2)$ and τ being the density, flow velocity and temperature deviated from the global equilibrium state, respectively. They are related to the velocity distribution function h as

$$(2.3) \qquad M(\boldsymbol{x}) = [\varrho, u_1, u_2, \tau] = \int m h(\boldsymbol{x}, \boldsymbol{v}) E(\boldsymbol{v}) d\boldsymbol{v},$$

where $E(\boldsymbol{v}) = \exp(-v^2)/\pi$ is the velocity distribution function at global equilibrium and $m = [1, v_1, v_2, v^2 - 1]$ are the collision invariants. Also, the shear stress σ_{ij} (with $i, j = 1, 2$) and heat flux q_i are defined as

$$(2.4) \qquad \sigma_{ij}(\boldsymbol{x}) = 2\int \left(v_i v_j - \frac{v^2}{2}\delta_{ij}\right) h(\boldsymbol{x}, \boldsymbol{v}) E(\boldsymbol{v}) d\boldsymbol{v},$$

$$(2.5) \qquad q_i(\boldsymbol{x}) = \int v_i \left(v^2 - 2\right) h(\boldsymbol{x}, \boldsymbol{v}) E(\boldsymbol{v}) d\boldsymbol{v},$$

with δ_{ij} the Kronecker delta.

In CIS, given the value of velocity distribution function $h^{(k)}(\boldsymbol{x}, \boldsymbol{v})$ at the k-th iteration step, its value at the next iteration step is calculated by solving the following equation:

$$(2.6) \qquad h^{(k+1)} + K\boldsymbol{v} \cdot \frac{\partial h^{(k+1)}}{\partial \boldsymbol{x}} = h_{eq}(h^{(k)}),$$

and the process is repeated until the maximum relative difference between successive estimates of macroscopic quantities, i.e.

$$(2.7) \qquad \epsilon = max\left\{\sqrt{\frac{\int |M^{(k+1)} - M^{(k)}|^2 d\boldsymbol{x}}{\int |M^{(k)}|^2 d\boldsymbol{x}}}\right\},$$

is less than a preassigned value.

Although in practical numerical simulations the streaming operator $\partial/\partial\boldsymbol{x}$ is discretized, e.g. by the finite-difference, finite-volume, or discontinuous Galerkin (DG) methods [21, 20], here it is kept intact when calculating the convergence rate of iteration; that of the discretized version of Eq. (2.6) will be tested in numerical simulations in section 5 below.

To calculate the convergence rate, we first define the following error function between velocity distribution functions at two consecutive iteration steps:

$$(2.8) \qquad Y^{(k+1)}(\boldsymbol{x},\boldsymbol{v}) = h^{(k+1)}(\boldsymbol{x},\boldsymbol{v}) - h^{(k)}(\boldsymbol{x},\boldsymbol{v}),$$

where according to Eq. (2.6) the error function $Y^{(k+1)}(\boldsymbol{x},\boldsymbol{v})$ is found to satisfy

$$(2.9) \qquad Y^{(k+1)} + K\boldsymbol{v}\cdot\frac{\partial Y^{(k+1)}}{\partial\boldsymbol{x}} = \Phi_\varrho^{(k)} + 2\Phi_{u_1}^{(k)}v_1 + 2\Phi_{u_2}^{(k)}v_2 + \Phi_\tau^{(k)}\left(v^2-1\right),$$

with the following error functions for macroscopic quantities between two consecutive iteration steps:

$$(2.10) \qquad \Phi_M^{(k+1)}(\boldsymbol{x}) = M^{(k+1)}(\boldsymbol{x}) - M^{(k)}(\boldsymbol{x}) = \int mY^{(k+1)}(\boldsymbol{x},\boldsymbol{v})E(\boldsymbol{v})d\boldsymbol{v}.$$

Second, to determine the convergence rate ω we perform the Fourier stability analysis by seeking the eigenvalue ω and eigenfunctions $y(\boldsymbol{v})$ and $\alpha_M = [\alpha_\varrho, \alpha_{u_1}, \alpha_{u_2}, \alpha_\tau]$ of the following forms:

$$(2.11) \qquad Y^{(k+1)}(\boldsymbol{x},\boldsymbol{v}) = \omega^k y(\boldsymbol{v})\exp(i\boldsymbol{\theta}\cdot\boldsymbol{x}),$$

$$(2.12) \qquad \Phi_M^{(k+1)}(\boldsymbol{x}) = \omega^{k+1}\alpha_M\exp(i\boldsymbol{\theta}\cdot\boldsymbol{x}),$$

where i is the imaginary unit and $\boldsymbol{\theta} = (\theta_1,\theta_2)$ is the wave vector of perturbance. It should be noted that the factor ω^k emerges in Eq. (2.11) rather than the factor ω^{k+1} in Eq. (2.12) because from Eq. (2.9) we know $Y^{(k+1)}$ is determined by macroscopic quantities in the k-th iteration step.

Substituting Eqs. (2.11) and (2.12) into Eqs. (2.9) and (2.10) we obtain the following expression for $y(\boldsymbol{v})$:

$$(2.13) \qquad y(\boldsymbol{v}) = \frac{\alpha_\varrho + 2\alpha_{u_1}v_1 + 2\alpha_{u_2}v_2 + \alpha_\tau(v^2-1)}{1 + iK\boldsymbol{\theta}\cdot\boldsymbol{v}}.$$

In the following we assume the wave vector of perturbation satisfies $|\boldsymbol{\theta}|^2 = \theta_1^2 + \theta_2^2 = 1$. Although in reality the perturbation may have various values of $\boldsymbol{\theta}$, their corresponding convergence rates do not interact because the kinetic equation is linear. Moreover from the denominator in Eq. (2.13) we see that the convergence rate depends only on the product of $K\boldsymbol{\theta}$. If $|\boldsymbol{\theta}| \neq 1$, the convergence rate at specific values of $\boldsymbol{\theta}$ and K can be calculated by replacing the Knudsen number with $K\boldsymbol{\theta}$.

Finally, with Eqs. (2.10), (2.11), (2.12) and (2.13), we obtain four linear algebraic equations for unknowns α_M that can be written in the matrix form as

$$(2.14) \qquad C\alpha_M^\top = \omega\alpha_M^\top,$$

where the superscript $\top$ is the transpose operator, and the 4×4 matrix C reads

$$(2.15) \qquad C = \begin{bmatrix} \int y_0 d\boldsymbol{v} & 2\int v_1 y_0 d\boldsymbol{v} & 2\int v_2 y_0 d\boldsymbol{v} & \int V y_0 d\boldsymbol{v} \\ \int v_1 y_0 d\boldsymbol{v} & 2\int v_1^2 y_0 d\boldsymbol{v} & 2\int v_1 v_2 y_0 d\boldsymbol{v} & \int v_1 V y_0 d\boldsymbol{v} \\ \int v_2 y_0 d\boldsymbol{v} & 2\int v_1 v_2 y_0 d\boldsymbol{v} & 2\int v_2^2 y_0 d\boldsymbol{v} & \int v_2 V y_0 d\boldsymbol{v} \\ \int V y_0 d\boldsymbol{v} & 2\int V v_1 y_0 d\boldsymbol{v} & 2\int V v_2 y_0 d\boldsymbol{v} & \int V^2 y_0 d\boldsymbol{v} \end{bmatrix},$$
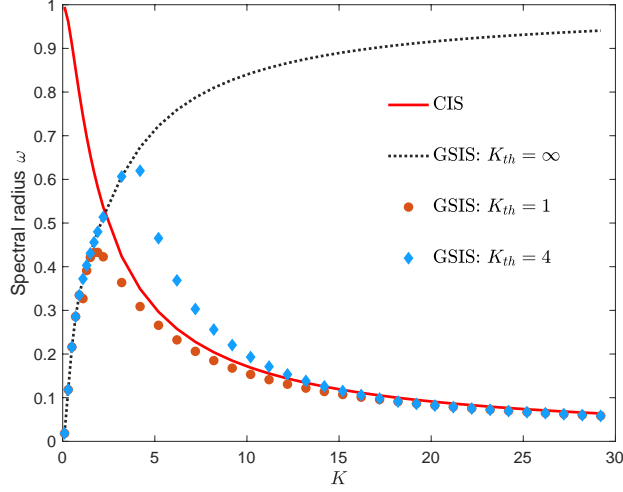
FIG. 1. *The spectral radius ω as a function of the Knudsen number K in both CIS and GSIS.*

with $V = v^2 - 1$ and

$$(2.16) \qquad\qquad y_0(\boldsymbol{v}) = \frac{E(\boldsymbol{v})}{1 + iK\boldsymbol{\theta} \cdot \boldsymbol{v}}.$$

In general, the spectral radius (i.e. the convergence rate, or the error decay rate between two consecutive iterations) can be obtained by numerically computing the eigenvalues ω of matrix C and taking the maximum absolute value of ω. The result of spectral radius as a function of the Knudsen number is shown in Figure 1, from which we find that in the limiting case of $K \to 0$, the spectral radius satisfies

$$(2.17) \qquad\qquad \omega_{CIS} \to 1 - \frac{K^2}{2}.$$

The fact that ω_{CIS} approaches one when $K \to 0$ means that errors defined in Eqs. (2.8) and (2.10) decay rather slowly. Worse still, the CIS encounters the problem

of "false convergence". According to the analysis of Adam and Larsen for radiation transfer problem [1] (same for the BGK equation here), if the iteration is terminated at the $(k+1)$-th step with the following convergence criterion

$$(2.18) \qquad |\Phi_M^{(k+1)} - \Phi_M^{(k)}| < \epsilon,$$

then the relative difference from the true steady-state solution Φ_M of Eq. (2.1) satisfies

$$(2.19) \qquad |\Phi_M^{(k+1)} - \Phi_M| < \frac{\omega_{CIS}}{1 - \omega_{CIS}}\epsilon \to \frac{2\epsilon}{K^2}, \quad \text{when } K \to 0.$$

The asymptotic expression at $K \to 0$ shows that, if the convergence criterion is chosen as Eq. (2.18), the error in final iteration can be much larger than the preassigned value ϵ; thus false convergence occurs if ϵ is not small enough. To reach the same convergence criterion for $|\Phi_M^{(k+1)} - \Phi_M|$, the total number of iterations must scale as $O(K^{-N})$ with $N > 2$. This proves that in CIS it is very hard to obtain the converged solution in the near-continuum flow regime; one numerical example is given in Figure 3 of Ref. [25].

3. GSIS and its convergence rate. In GSIS, additional macroscopic equations are coupled with CIS to boost convergence [22]. The algorithm is summarized below. First, given the value of velocity distribution function $h^{(k)}(\boldsymbol{x}, \boldsymbol{v})$ at the k-th iteration step, the velocity distribution function at the intermediate $(k+1/2)$-th step is obtained by solving the following kinetic equation:

$$(3.1) \qquad h^{(k+1/2)} + K\boldsymbol{v} \cdot \frac{\partial h^{(k+1/2)}}{\partial \boldsymbol{x}} = h_{eq}(h^{(k)}).$$

Second, macroscopic quantities $\bar{M} = (\bar{\varrho}, \bar{u}_1, \bar{u}_2, \bar{\tau})$ are introduced to update M at the (k+1)-th iteration step in the following manner:

$$(3.2) \qquad M^{(k+1)} = \beta \bar{M} + (1 - \beta)M^{(k+1/2)},$$

where $0 \le \beta \le 1$ is the relaxation coefficient: when $\beta = 0$ the scheme is reduced to the CIS with no fast convergence at small values of Knudsen number. The macroscopic quantities $\bar{M}$ can be obtained by solving the following partial differential equations (note that the Einstein summation is used):

$$(3.3) \qquad \frac{\partial \bar{u}_i}{\partial x_i} = 0,$$

$$(3.4) \qquad \frac{\partial \bar{\varrho}}{\partial x_i} + \frac{\partial \bar{\tau}}{\partial x_i} + \frac{\partial \bar{\sigma}_{ij}}{\partial x_j} = 0,$$

$$(3.5) \qquad \frac{\partial \bar{q}_i}{\partial x_i} = 0,$$

which are obtained by taking the velocity moments of the BGK equation (2.1). However, in GSIS the shear stress $\bar{\sigma}_{ij}$ and heat flux $\bar{q}_i$ are not directly calculated from the velocity distribution function $h^{(k+1/2)}$ according to Eqs. (2.4) and (2.5), but are reconstructed as

$$(3.6) \qquad \bar{\sigma}_{ij} = -2K_e \frac{\partial \bar{u}_{<i}}{\partial x_{j>}} + \text{HoT}_{\sigma_{ij}},$$

$$(3.7) \qquad \bar{q}_i = -K_e \frac{\partial \bar{\tau}}{\partial x_i} + \text{HoT}_{q_i},$$

such that Eqs. (3.3), (3.4) and (3.5) are reduced to the linearized Navier-Stokes equations when the high-order terms $\text{HoT}_{\sigma_{ij}}$ and HoT_{q_i} vanish, and when the effective Knudsen number K_e takes the value of $K_e = K$; the role of K_e will be discussed below in section 3.2, where we will show the inclusion of Newton's law for shear stress and Fourier's law for heat conduction in Eqs. (3.6) and (3.7) is critical to realize the fast convergence for near-continuum flows. Note that the high-order terms $\text{HoT}_{\sigma_{ij}}$ and HoT_{q_i} should be derived exactly from the kinetic equation (2.1), so that the rarefaction effects beyond the Navier-Stokes level are all captured.

There are many ways to construct the expressions for these high-order terms. In our paper [22], they are constructed by multiplying Eq. (2.1) with $\left(v_i v_j - v^2 \delta_{ij}/2\right) E$ and $v_i \left(v^2 - 2\right) E$, respectively, and integrating the resultant equations with respect to the molecular velocity $\boldsymbol{v}$:

$$(3.8) \qquad \text{HoT}_{\sigma_{ij}} = 2K_e \frac{\partial u_{<i}^{(k+1/2)}}{\partial x_{j>}} - 2K \int \left(v_i v_j - \frac{v^2}{2}\delta_{ij}\right)\boldsymbol{v} \cdot \frac{\partial h^{(k+1/2)}}{\partial \boldsymbol{x}} d\boldsymbol{v},$$

$$(3.9) \qquad \text{HoT}_{q_i} = K_e \frac{\partial \tau^{(k+1/2)}}{\partial x_i} - K \int v_i (v^2 - 2)\boldsymbol{v} \cdot \frac{\partial h^{(k+1/2)}}{\partial \boldsymbol{x}} d\boldsymbol{v}.$$

Clearly, the above high-order terms are correct when the steady state is reached. However, during iteration when the steady state is not reached, Eqs. (3.6) and (3.7) are not correct since macroscopic quantities $\bar{M}$ are evaluated at the $(k+1)$-th iteration step, while high-order terms $\text{HoT}_{\sigma_{ij}}$ and HoT_{q_i} are computed using the intermediate velocity distribution function at the $(k + 1/2)$-th iteration step.

3.1. The case of $M^{(k+1)} = \bar{M}$ and $K_e = K$. To calculate the convergence rate of GSIS, we introduce the error function for the velocity distribution function as

$$(3.10) \qquad Y^{(k+1/2)}(\boldsymbol{x}, \boldsymbol{v}) = h^{(k+1/2)}(\boldsymbol{x}, \boldsymbol{v}) - h^{(k)}(\boldsymbol{x}, \boldsymbol{v}) = \omega^k y(\boldsymbol{v}) \exp(i\boldsymbol{\theta} \cdot \boldsymbol{x}),$$

while the error functions Φ_M are still defined as $\Phi_M^{(k+1)}(\boldsymbol{x}) = M^{(k+1)}(\boldsymbol{x}) - M^{(k)}(\boldsymbol{x})$. It is clear that, according to Eq. (3.1), the solution of $y(\boldsymbol{v})$ is still given by Eq. (2.13). It should be emphasized that, in GSIS the error functions Φ_M are not directly calculated from Y. Instead, macroscopic quantities $M^{(k+1)}(\boldsymbol{x})$ are governed by Eqs. (3.3), (3.4) and (3.5), from which one can obtain the governing equations for $\Phi_M^{(k+1)}(\boldsymbol{x})$.

On substituting Eqs. (3.10), (2.10) and (2.12) into the governing equations for $\Phi_M^{(k+1)}(\boldsymbol{x})$, we obtain the following four linear algebraic equations:

$$(3.11) \qquad \omega(i\theta_1 \alpha_{u_1} + i\theta_2 \alpha_{u_2}) = 0,$$

$$(3.12) \qquad \omega[i\theta_1(\alpha_\varrho + \alpha_\tau) + K_e \alpha_{u_1}] = S_1,$$

$$(3.13) \qquad \omega[i\theta_2(\alpha_\varrho + \alpha_\tau) + K_e \alpha_{u_2}] = S_2,$$

$$(3.14) \qquad \omega K_e \alpha_\tau = S_\tau,$$

where the source terms, due to high-order terms in Eqs. (3.8) and (3.9), are also linear functions of α_M: $S_1 = \int s_1 y E d\boldsymbol{v}$, $S_2 = \int s_2 y E d\boldsymbol{v}$, and $S_\tau = \int s_\tau y E d\boldsymbol{v}$, with

$$(3.15) \qquad s_1 = K_e v_1 + K(i\theta_1 v_1 + i\theta_2 v_2)\left[i\theta_1(v_1^2 - v_2^2) + 2i\theta_2 v_1 v_2\right],$$

$$(3.16) \qquad s_2 = K_e v_2 + K(i\theta_1 v_1 + i\theta_2 v_2)\left[2i\theta_1 v_1 v_2 + i\theta_2(v_2^2 - v_1^2)\right],$$

$$(3.17) \qquad s_\tau = K_e V + K(i\theta_1 v_1 + i\theta_2 v_2)(i\theta_1 v_1 + i\theta_2 v_2)(V - 1).$$

The spectral radius of GSIS can be obtained by solving Eqs. (3.11)-(3.17). That is, these equations are firstly rewritten in the matrix form as

$$(3.18) \qquad L\omega \alpha_M^\top = R\alpha_M^\top,$$

where the 4×4 matrix L is obtained from the left-hand side of equations (3.11)-(3.14) as:

$$(3.19) \qquad L = \begin{bmatrix} 0 & i\theta_1 & i\theta_2 & 0 \\ i\theta_1 & K_e & 0 & i\theta_1 \\ i\theta_2 & 0 & K_e & i\theta_2 \\ 0 & 0 & 0 & K_e \end{bmatrix},$$

and due to the fact that $y(\boldsymbol{v})$ in Eq. (2.13) a linear combination of $\alpha_\varrho, \alpha_{u_1}, \alpha_{u_2}$ and α_τ, the 4×4 matrix R is obtained from Eqs. (3.15)-(3.17) as

$$(3.20) \qquad R = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \int y_0 s_1 d\boldsymbol{v} & 2\int v_1 y_0 s_1 d\boldsymbol{v} & 2\int v_2 y_0 s_1 d\boldsymbol{v} & \int V y_0 s_1 d\boldsymbol{v} \\ \int y_0 s_2 d\boldsymbol{v} & 2\int v_1 y_0 s_2 d\boldsymbol{v} & 2\int v_2 y_0 s_2 d\boldsymbol{v} & \int V y_0 s_2 d\boldsymbol{v} \\ \int y_0 s_\tau d\boldsymbol{v} & 2\int v_1 y_0 s_\tau d\boldsymbol{v} & 2\int v_2 y_0 s_\tau d\boldsymbol{v} & \int V y_0 s_\tau d\boldsymbol{v} \end{bmatrix}.$$

By introducing $G = L^{-1}R$ and numerically computing the eigenvalues of the matrix G we obtain the spectral radius ω of GSIS. The results are shown in the dotted black line in Figure 1. It is seen that the spectral radius goes to zero as $\omega \sim K^2$ when $K \to 0$. This demonstrates that the GSIS is able to boost convergence significantly in the near-continuum flow regime. However, the spectral radius increases to one when $K \to \infty$. This stands in sharp contrast to DSA schemes for rarefied Poiseuille-type flows, where the flow velocity (only in one direction) is perpendicular to spatial variables [24]. As a consequence, the synthetic equation for the only flow velocity is of diffusion-type [24, 27, 19], so that the spectral radius goes to zero when $K \to \infty$ and $K \to 0$. However, in GSIS we have multiple flow velocities whose directions are inside the spatial domain, and the synthetic equations contain the continuity equation (3.3) which is not of diffusion type. We believe the continuity equation makes the spectral radius goes to one when the Knudsen number is large.

To make the spectral radius in GSIS goes to zero when Knudsen number is large, in Eq. (3.2) we choose the relaxation coefficient as

$$(3.21) \qquad \beta = \frac{min(K, K_{th})}{K},$$

where K_{th} is the threshold Knudsen number, so that GSIS is reduced to CIS when K is large, where the spectral radius approaches zero. The spectral radius of this GSIS can be obtained by computing the eigenvalue of the matrix $G = \beta L^{-1}R + (1-\beta)C$, where the results at the threshold Knudsen number of values 1 and 4 are shown in Figure 1. Clearly, by choosing approximate value of β, we can make the maximum value of ω less than 0.5; this means that after 10 iterations, the error will be decreased by at least three orders of magnitude. Thus, theoretically, GSIS can reach fast convergence in the whole range of Knudsen number, which has been demonstrated by the numerical examples in our recent paper [22].

In addition to the fast converge at large Knudsen numbers, Eq. (3.21) with $\beta < 1$ can make the algorithm more stable. This is because the high-orders terms defined in Eqs. (3.8) and (3.9) may have strong variations around sharp (i.e. rectangular) corners, which leads to unphysical variations of density, velocity and temperature when solving Eqs. (3.3), (3.4) and (3.5). By updating the macroscopic quantities
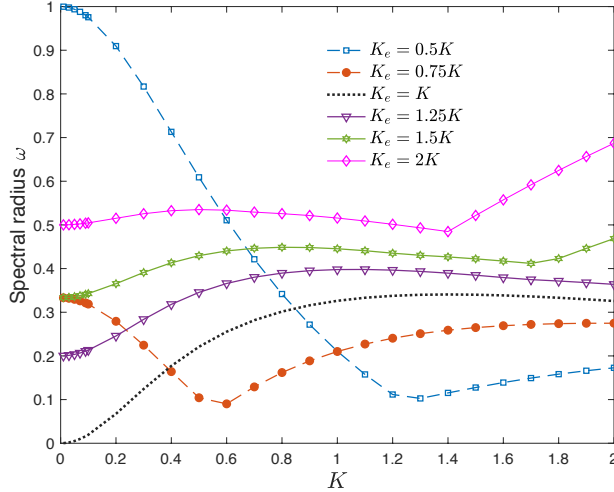
Fig. 2. *The spectral radius ω as a function of the Knudsen number K in GSIS, when K_e in Eqs.* (3.6) *and* (3.7) *takes different values and the relaxation coefficient in Eq.* (3.2) *is $\beta = 1$.*

according to Eq. (3.2), the error at sharp boundaries is reduced and stable algorithm is guaranteed [22].

3.2. Why are the Navier-Stokes constitutive relations important. Note that when the Knudsen number is small, we choose $K_e = K$ in Eqs. (3.6) and (3.7); this means that in the near-continuum regime, the constitutive relations in Navier-Stokes equations, i.e. Newton's law for shear stress and Fourier's law for heat conduction are explicitly included in the macroscopic synthetic equations. This turns out to be extremely important for the fast convergence of GSIS, as the spectral radius goes to zero when $K \to 0$. To demonstrate this superiority, we calculate the spectral radius by choosing different values of K_e when the Knudsen number is small; the results in Figure 2 show that only when $K_e = K$ can the spectral radius go to zero when $K \to 0$. Moreover, when the Knudsen is small, e.g. $K < 0.2$, the further the effective Knudsen number K_e deviates away from K, the larger the spectral radius and hence the slower

the convergence. These results confirm that the fastest convergence can be realized by including the Navier-Stokes constitutive relations explicitly in synthetic equations. If the shear stress and heat flux in Eqs. (3.4) and (3.5) are directly computed from the $h^{(k+1/2)}$ according to Eqs. (2.4) and (2.5), our Fourier stability analysis finds that the scheme does not permit fast convergence in the near-continuum flow regime.

4. Asymptotic preserving property of GSIS. Note that in GSIS the kinetic equation and macroscopic synthetic equations can be solved by different numerical methods with different order of accuracy. Now we consider the influence of spatial discretization in the gas kinetic solver on the accuracy of GSIS, based on the assumptions that (i) synthetic equations can be solved exactly and (ii) the spatial cell size Δx is able to capture the physical solution. To this end, we consider whether the Navier-Stokes equations can be exactly derived, through the Chapman-Enskog expansion [4], from the discretized gas kinetic equation with the following scaling:

$$\Delta x \sim K^{1/\alpha}. \tag{4.1}$$

Here α denotes the order of accuracy in the asymptotic preserving of Navier-Stokes equations. Clearly, the larger the value of α, the better the numerical scheme, as we can use less number of spatial cells to achieve the same accuracy. It has been rigorously proven that the discrete unified gas kinetic scheme has $\alpha = 2$, so that Δx can be much larger than the mean free path to capture the hydrodynamics in bulk regions [10]. If $\alpha = \infty$, the scheme will capture the hydrodynamical behavior of gas when Δx is approximately the system size (no matter what the value of K is), as long as this size is adequate to capture the flow physics, for example the dynamics of a sound wave whose wavelength is approximately the system size.

Since the spectral radius of GSIS approaches zero when $K \to 0$, the converged solution can be found within a few number of iterations, and we have $h_{eq}^{(k+1)} = h_{eq}^{(k)}$ and $M^{(k+1)} = M^{(k)}$. Thus, when the iteration is converged, the kinetic equation (3.1) in the discretized form can be written as

$$\boldsymbol{v} \cdot \frac{\partial h}{\partial \boldsymbol{x}} + O(\Delta x^n)\delta(h) = \frac{h_{eq} - h}{K}, \tag{4.2}$$

where n is the order of approximation for the spatial derivative $\partial h/\partial x$, and $\delta(h)$ is the $(n+1)$-th order derivative of h. For instance, if we use the second-order upwind finite difference scheme, $n = 2$. If the third-order approximating polynomials are used in the DG method [21, 20], we have $n = 4$.

In the Chapman-Enskog expansion the velocity distribution function is approximated by the Taylor expansion $h = h_0 + Kh_1 + \cdots$. By subsisting this expansion into Eq. (4.2) and collecting terms with the order of K^{-1}, we have $h_0 = f_{eq}$, when the following largest scaling is chosen

$$\Delta x \sim K^{1/\infty} = O(1). \tag{4.3}$$

Under the scaling (4.3), by collecting terms with the order K^0, we have $h_1 = -\boldsymbol{v} \cdot \partial h_{eq}/\partial \boldsymbol{x} - \delta(h_{eq})$. Thus, according to Eqs. (3.6), (3.7), (3.8) and (3.9), Newton's law for shear stress and Fourier's law for heat conduction are recovered in macroscopic synthetic equations with accuracy $O(K^2)$:

$$\sigma_{ij} = -2K\frac{\partial u_{<i}}{\partial x_{j>}} + O(K^2), \quad q_i = -K\frac{\partial \tau}{\partial x_i} + O(K^2). \tag{4.4}$$

Thus, as long as the spatial resolution $\Delta x = O(1)$ is able to capture the physical solution, GSIS is able to recover the linearized Navier-Stokes equations when $K \to 0$. In this sense the scheme is better than the discrete unified gas kinetic scheme where $\Delta x = O(\sqrt{K})$ [10]. Therefore, the overall order of accuracy of GSIS depends only on the order to solve macroscopic synthetic equations, that is, n in Eq. (4.2). In reality, however, such a large spatial cell size $\Delta x = O(1)$ cannot be used in regions with Knudsen layers or shock structures, where the physical solutions require a spatial resolution of $O(K)$. Fortunately, these kinetic layers only take up a small fraction of the computational domain, say, in the vicinity of solid walls, which can be captured by implicit schemes with non-uniform spatial discretization. This will be tested in the following numerical examples.

5. Numerical examples. To assess the analytical results in previous sections, we simulate the one-dimensional coherent Rayleigh-Brillouin scattering, sound wave propagation, and Couette flows, as well as two-dimensional thermally-induced flow.

5.1. Coherent Rayleigh-Brillouin scattering. In coherent Rayleigh-Brillouin scattering the wavelike density perturbation in gas is created by a moving optical lattice, where the governing equation for the velocity distribution function h' can be written as

$$(5.1) \qquad \frac{\partial h'}{\partial t} + v_2 \frac{\partial h'}{\partial x_2} = \frac{h'_{eq} - h'}{K} + 2v_2 \cos(2\pi x_2 + 2\pi f_s t),$$

where f_s is the frequency of the moving optical lattice; we choose $f_s = \sqrt{5/6}$ (i.e. the sound speed normalized by the most probable speed of gas molecules) so that the amplitude of perturbed density will scale as $1/K$ when the Knudsen number is small. If the numerical scheme cannot preserve the Navier-Stokes limit when $K \to 0$, then a huge number of spatial cells is needed to keep a low numerical dissipation. Since the interaction between the gas and solid wall is absent [26], this problem is ideal to test the accuracy and efficiency of the kinetic scheme.

The problem is periodic in both temporal and spatial directions. Therefore, using the transform $h'(x_2, \boldsymbol{v}, t) = h(x_2, \boldsymbol{v}) \exp(2\pi i f_s t)$, the final perturbed density profile can be obtained by solving the following equation [26]:

$$(5.2) \qquad 2\pi i f_s h + v_2 \frac{\partial h}{\partial x_2} = \frac{h_{eq} - h}{K} + 2v_2 \cos(2\pi x_2),$$

which can be handled by implicit solvers.

Note that in order to keep the calculation more realistic, in the following the linearized Shakhov kinetic model is used and the velocity distribution function contains the three-dimensional molecular velocity space [16]. That is, the reference distribution function in Eq. (5.2) now becomes

$$(5.3) \qquad h_{eq}(\boldsymbol{x}, \boldsymbol{v}) = \varrho(\boldsymbol{x}) + 2\boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v} + \tau(\boldsymbol{x}) \left(v^2 - 1\right) + \frac{4}{15} \boldsymbol{q} \cdot \boldsymbol{v} \left(v^2 - \frac{5}{2}\right),$$

where the perturbed density is $\rho = \int E_3(\boldsymbol{v}) h d\boldsymbol{v}$ with the equilibrium velocity distribution function $E_3(\boldsymbol{v}) = \exp(-v^2)/\pi^{1.5}$, the flow velocity is $\boldsymbol{u} = \int \boldsymbol{v} E_3(\boldsymbol{v}) h d\boldsymbol{v}$, the perturbed temperature is $\tau = \int (2v^2/3 - 1) E_3(\boldsymbol{v}) h d\boldsymbol{v}$, and the heat flux is $\boldsymbol{q} = \int \boldsymbol{v} \left(v^2 - 5/2\right) E_3(\boldsymbol{v}) h d\boldsymbol{v}$.

The presence of the first term in Eq. (5.2) modifies the synthetic equations (3.3), (3.4) and (3.5) to the following forms [22]:

$$(5.4) \qquad 2\pi i f_s \bar{\varrho} + \frac{\partial \bar{u}_i}{\partial x_i} = 0,$$

$$(5.5) \qquad 2\pi i f_s \bar{u}_i + \frac{\partial \bar{\varrho}}{\partial x_i} + \frac{\partial \bar{\tau}}{\partial x_i} + \frac{\partial \bar{\sigma}_{ij}}{\partial x_j} = 0,$$

$$(5.6) \qquad 2\pi i f_s \bar{\tau} + \frac{\partial \bar{q}_i}{\partial x_i} + \frac{\partial \bar{u}_i}{\partial x_i} = 0,$$

while the shear stress and heat flux are modified into the following forms:

$$(5.7) \qquad 2\pi i f_s \sigma_{22} + \mathrm{HoT}_{\sigma_{22}} + \frac{4}{3}\frac{\partial \bar{u}_2}{\partial x_2} = -\delta_{\mathrm{rp}}\sigma_{22},$$

$$(5.8) \qquad 2\pi i f_s q_2 + \mathrm{HoT}_{q_2} + \frac{5}{4}\frac{\partial \bar{\tau}}{\partial x_2} = -\frac{2}{3}\delta_{\mathrm{rp}}q_2.$$

The comparison between CIS and GSIS in terms of efficiency and accuracy is shown in Figure 3, when the Knudsen number is $K = 0.005$ and the synthetic equations in GSIS are solved by the Fourier spectral method exactly. Due to the small value of convergence rate ω, GSIS yields converged solution even after only 3 iterations and when $\epsilon = 10^{-3}$, which is consistent with the analytical solution (2.19). That is,

$$(5.9) \qquad |\Phi_M^{(k+1)} - \Phi_M| < \frac{\omega_{GSIS}}{1 - \omega_{GSIS}}\epsilon \to \epsilon K^2, \quad \text{when } K \to 0.$$

However, since ω_{CIS} approaches one when $K \to 0$, CIS needs huge number of iterations and the convergence criterion has to be set very small, i.e. $\epsilon = 10^{-10}$ in this case. This is also consistent with the prediction in Eq. (2.19). In terms of spatial accuracy, Figure 3(b) proves that GSIS can asymptotically preserve the Navier-Stokes limit with $\Delta x \sim O(1)$ provide this spatial resolution is adequate to describe the physical problem (here it is the density wave profile), while the CIS needs much smaller spatial cell size in order to keep the numerical dissipation small: when the cell size is large, the numerical dissipation leads to smaller amplitude of density wave when compared to the true solution.

To further show the importance of spatial resolution, we consider the same problem when the Knudsen number is $K = 10^{-8}$ in Figure 4. Even in this case the Euler equation cannot be used, otherwise the amplitude of perturbation will go to infinity. The extremely small value of K poses a real challenge to the kinetic scheme, as any small value of numerical dissipation could easily contaminate the final solution, which leads to a significant small amplitude of density perturbation. Even at such a small Knudsen number, if the macroscopic equation is solved exactly (by the Fourier spectral method) and $\Delta x \sim O(1)$ is able to capture the spatial variation, GSIS has infinite order of accuracy, see the circles in Figure 4. However, when the second-order finite difference scheme is used to solve the synthetic equations, we need more than 2000 spatial points to capture the density perturbation; this is understandable since the numerical dissipation (or numerical viscosity) $(\Delta x)^2 = 2.5 \times 10^{-7}$ is larger than the physical viscosity (here it is reflected by the Knudsen number K). When the synthetic equations are solved by the fourth-order finite difference, we see that $N = 100$ leads to wrong solutions, while $N = 300$ yields accurate solution. This is because the numerical dissipations are about $(\Delta x)^4 = 10^{-8}$ and 1.2×10^{-10}, respectively, so that the former is too dissipative while the latter is accurate.
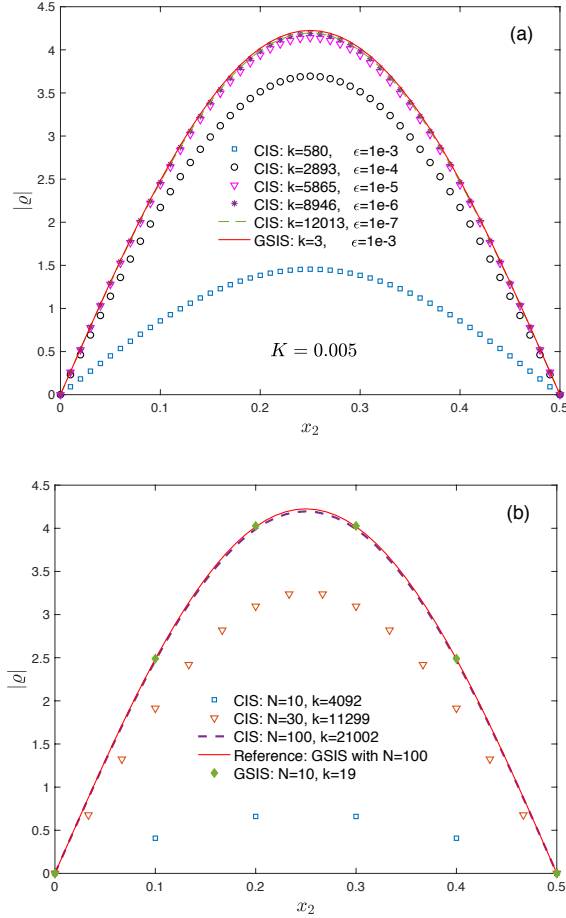
FIG. 3. *Comparisons of the density profile in coherent Rayleigh-Brillouin scattering obtained from CIS and GSIS, when the Knudsen number is $K = 0.005$. (a) The convergence history of CIS, at different values of the error $\epsilon = \left| \int |\varrho^{(k+1)} dx_2| / \int |\varrho^{(k)}| dx_2 - 1 \right|$. The spatial region $x_2 \in [0,1]$ is divided into $N = 100$ uniform cells, and due to symmetry only half of the density wave profile is plotted. The result from GSIS is converged even when $\epsilon = 10^{-3}$, after three iterations. (b) The influence of spatial cell number on the density profile, where the iteration is terminated when $\epsilon < 10^{-10}$. The kinetic equation is solved by the second-order upwind finite difference, while the synthetic equations are solved by the Fourier spectral method.*

5.2. One-dimensional sound wave propagation. The governing equation in this problem is the same as that in coherent Rayleigh-Brillouin scattering, except here we have no external driving force and the diffuse boundary condition is imposed on two walls: the one at $x_2 = 0$ is oscillating in the x_2 direction with a velocity $u_2 = \cos(2\pi f_s t)$ while the one at $x_2 = 1$ is stationary, see the schematic in Figure 5; detailed form of the diffuse boundary condition for the linearized kinetic equation can be found in Ref. [22]. Here we choose $f_s = 1/2\pi$ and $\sqrt{\pi} K/2 = 0.001$, and the kinetic and macroscopic equations are solved by DG with different orders of accuracy. This test is used to show that, when the Knudsen layer near the solid walls are well resolved, the accuracy of GSIS is mainly determined by the accuracy in solving macroscopic synthetic equations.
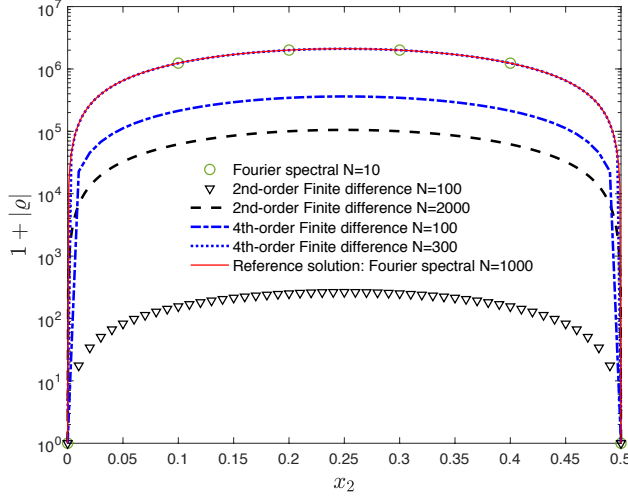
FIG. 4. *Comparisons of the density profile in coherent Rayleigh-Brillouin scattering, where the kinetic equation is solved by the second-order upwind finite difference, while macroscopic equations are solved by various schemes with different number of spatial points (N). The Knudsen number is $K = 10^{-8}$. The density amplitude is shifted by one in order to show it in the log scale.*

In the numerical simulation we assume the thickness of Knudsen layer is about 4 times the mean free path in the immediate vicinity of walls [19], which is divided into $N_\lambda = 16$ equal cells each, see Figure 5(a). The bulk region $0.004 < x_2 < 0.996$ is partitioned by N_b uniform cells. The order of DG scheme for the kinetic equation is denoted by n_K while the one for synthetic equations is n_S. In the constitutive relations (3.6) and (3.7), the effective Knudsen number is $K_e = K$. During iteration, the macroscopic quantities $M^{(k+1)}$ are updated based on Eq. (3.2), and the relaxation coefficient is chosen as

$$(5.10) \qquad \beta = \frac{min\,(K_{\mathrm{loc}}, K_{th})}{K_{\mathrm{loc}}},$$

where K_{th} is set as $min\left(1, \frac{5}{n_K - 1}\right)$ and K_{loc} is the local Knudsen number estimated
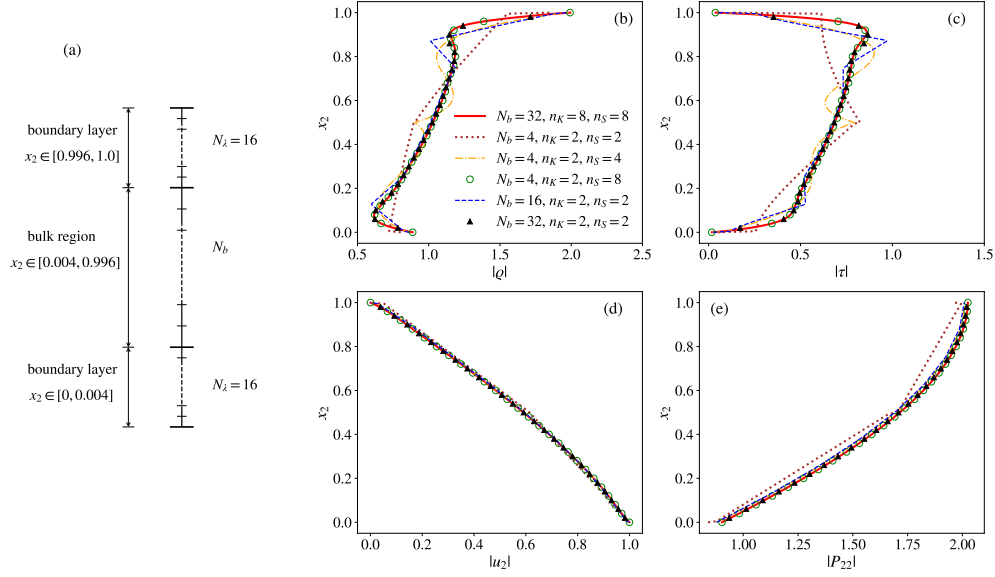
FIG. 5. *One-dimensional sound wave propagation induced by an oscillating wall at $x_2 = 0$, when the Knudsen number is $K = 0.001$. (a) schematic illustration of the bulk region and boundary layer, and spatial discretization along x_2. (b)-(e) are respectively the amplitudes of density $|\varrho|$, temperature $|\tau|$, longitudinal velocity $|u_2|$ and normal pressure $|P_{22}| = |\sigma_{22} + \varrho + \tau|$ obtained from schemes with different orders of accuracy and numbers of cells.*

from the local cell size: $K_{\mathrm{loc}} = K/L_{\mathrm{loc}}$ with L_{loc} the minimum height of the triangles used in DG, where the one-dimensional problem is actually simulated on two-dimensional domain partitioned by structured triangles and the periodic boundary condition is imposed on the lateral boundaries of the domain. Note that the Knudsen number is defined as the mean free path over the characteristic system length L and the spatial coordinate has been normalized by L.

Figure 5(b)-(e) show the amplitudes of density, temperature, velocity and normal pressure obtained from schemes with different orders of accuracy and numbers of cells. By choosing the numerical results with $N_{\mathrm{b}} = 32$ and $n_{\mathrm{K}} = n_{\mathrm{S}} = 8$ as the reference, we see that the increase of n_{S} or N_b leads to solutions converge toward the reference one, even when the kinetic equation is solved with second-order accuracy, i.e. $n_{\mathrm{K}} = 2$. This means that in GSIS the accuracy is controlled by the macroscopic solver, at least in the near-continuum flow regime.

Table 1 shows the total number of iterations and the amplitude of normal pressure at the oscillating wall for different combinations of N_{b}, n_{K} and n_{S}, when ϵ defined in Eq. (2.7) is less than 10^{-5}. For all test cases, the convergence is reached within 50 iteration steps, which demonstrates the efficiency of GSIS, even when the streaming operator is discretized. The fact that the total iteration step is larger than that in the coherent Rayleigh-Brillouin scattering is that, in this case, although K is small, the presence of Knudsen layer makes $K|\boldsymbol{\theta}|$ approximately one in Eq. (2.13). Thus, relative large iteration number is needed. On the other hand, Table 1 further confirms that the accuracy of solution in the bulk region is mainly determined by the accuracy in solving the macroscopic equations. As a consequence, the higher the order of DG for macroscopic equations, the more accurate the solution will be. Furthermore, accurate

TABLE 1

One-dimensional sound wave propagation when $\sqrt{\pi}K/2 = 0.001$: the number of iterative steps (Itr) when the residual $\epsilon < 10^{-5}$ and the amplitude of normal pressure $|P_{22}|$ at the oscillating wall ($x_2 = 0$) under different combinations of N_b, n_K and n_S. N_b is the number of cells in the bulk region and $L_b = 0.992/N_b$ is the cell size in the bulk region. n_K and n_S are the orders of DG to approximate the kinetic equation and synthetic equations, respectively.

| N_b | L_b | $n_K = 2$ | | | | | | $n_K = 8$, $n_S = 8$ | |
| | | $n_S = 2$ | | $n_S = 4$ | | $n_S = 8$ | | | |
		Itr	$|P_{22}|$	Itr	$|P_{22}|$	Itr	$|P_{22}|$	Itr	$|P_{22}|$
4	0.248	46	0.840	36	0.882	32	0.903	33	0.903
8	0.124	47	0.846	36	0.900	32	0.903	33	0.903
16	0.062	47	0.877	36	0.903	32	0.903	33	0.903
32	0.031	47	0.897	36	0.903	32	0.903	33	0.903

results can be obtained once macroscopic quantities are well resolved on coarse mesh, where the cell size is much larger than the mean free path of gas molecules. For instance, when the 8th-order DG for macroscopic equations is employed with $N_b = 4$, converged results are obtained on a mesh in the bulk region 248 times larger than the molecular mean free path. This is consistent with the results in section 4 that GSIS is able to asymptotically preserve the Navier-Stokes limit with a spatial size $\Delta x \sim O(1)$, provided this size is able to describe the physical solution.

5.3. One-dimensional Couette flow. We consider the steady Couette flow between two infinite parallel plates located at $x_2 = 0$ and $x_2 = 1$. The top plate moves along the horizontal direction with a velocity of $u_1 = 1$, while the bottom plate moves in the opposite direction with the same magnitude of speed. The Knudsen number is set as $\sqrt{\pi}K/2 = 0.01$. The diffuse boundary condition is imposed on the two plates. This test is used to show the importance of resolving the Knudsen layer [19].

The partition of the computational domain along the longitudinal direction (x_2) is similar as that in Figure 5(a). In order to guarantee the accuracy of resolution in the bulk region, the bulk region $x_2 \in [0.04, 0.96]$ is partitioned into $N_b = 32$ uniform cells, and the synthetic equations are solved with $N_S = 8$. Details to evaluate the constitutive relations in the synthetic equations and update the macroscopic quantities at $(k+1)$-th iteration step are given in section 5.2. We are interested in the velocity gradient at the center of the domain

$$(5.11) \qquad\qquad k_1 = \frac{\mathrm{d}u_1(x_2 = 0.5)}{\mathrm{d}x_2},$$

as well as the Knudsen layer function u_s [19]:

$$(5.12) \qquad\qquad u_s(x_2) = \frac{u_{NS}(x_2) - u_1(x_2)}{\sqrt{\pi}k_1 K/2}$$

where u_{NS} is the velocity in the bulk region approximated by $u_{NS}(x_2) = k_1(x_2 - 0.5)$.

Figure 6 illustrates the Knudsen layer functions obtained from DG schemes with different values of n_K and N_λ. The reference result is obtained when 16 nonuniform cells in the Knudsen layer with refinement near the wall are used, where the minimum and maximum cell sizes are 0.0089 and 0.46 mean free path, respectively; this nonuniform grid is more suitable to recover the divergence of velocity gradient at the
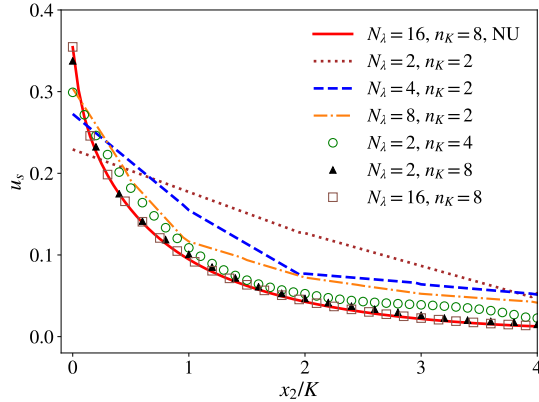
FIG. 6. *The Knudsen layer functions in one-dimensional Couette flow obtained from DG schemes with different orders of accuracy n_K (for the kinetic equation) and numbers of uniform cells N_λ. The Knudsen number is $\sqrt{\pi}K/2 = 0.01$. The order of DG scheme for macroscopic equations is $n_S = 8$, and the bulk region is partitioned by 32 uniform cells. Note that the reference result (red solid line labeled with 'NU') is obtained on the grid with 16 nonuniform cells in the Knudsen layer with refinement approaching to the wall.*

wall [13]. It is concluded that using very few cells and lower order scheme cannot predict the correct Knudsen layer function. Accurate solution can only be obtained by increasing the order of the scheme and/or the number of cells in the Knudsen layer.

Table 2 shows the number of iteration steps, the defect velocity at the wall and the velocity gradient k_1 for different combinations of N_λ and n_K, when ϵ defined in Eq. (2.7) is less than 10^{-5}. The convergence can be reached around 30 iteration steps when the cell size in the Knudsen layer L_λ is not smaller than 0.25 mean free path. However, as the cell size further reduces, the number of iteration increases. This is due to the fact that the local Knudsen number Kn_{loc} appearing in Eq. (5.10) that depends on cell size becomes large, and the coefficient β in Eq. (3.2) for the update of macroscopic quantities is less than one. Therefore, macroscopic quantities are not fully corrected by solutions from synthetic equations, which reduces the efficiency of GSIS a bit.

It is interesting to note that even when the Knudsen layer is not resolved, the velocity gradients (5.11) in Table 2 all approach the reference value. This is understandable since according to Eq. (5.12) the error introduced by inaccurate capture of Knudsen layer is $O(K)$, which is small compared to the wall velocity $O(1)$ in the near-continuum flow regimes. However, this case does not indicate that the capture of Knudsen layer is not important. In thermal effects where the velocity slip induced in the Knudsen layer is the origin of gas motion [17], the Knudsen layer must be resolved, in order to obtain accurate flow filed in the whole computational domain, see the spatial discretization in Ref. [23] for an example. In the following case, we will show evidence for this argument.

5.4. Two-dimensional thermally-induced flow. As depicted in Figure 7(a), we consider a two-dimensional flow induced by a hot beam that is encompassed in a cold chamber. Both the beam and chamber are square, with dimensions of 2×2 and 8×8, respectively. The beam with a (deviated) temperature of $\tau = 1$ is placed with distance of 1 away from the left and bottom walls of the enclosure. The temperature of the chamber is $\tau = 0$ and gas is filled between the beam and chamber. It is well

The number of iterative steps (Itr), the defect velocity u_s at the wall, and the Knudsen layer function $k_1 = \mathrm{d}u_1(x_2 = 0.5)/\mathrm{d}x_2$ under different combinations of N_λ and n_K, in one-dimensional Couette flow with $\sqrt{\pi}K/2 = 0.01$. $L_\lambda = 0.04/N_\lambda$ is the uniform cell size in the Knudsen layer. The bulk region is partitioned by 32 uniform triangles and the order of DG scheme for macroscopic equations is $n_S = 8$.

N_λ	$\frac{L_\lambda}{\sqrt{\pi}K/2}$	$n_K = 2$			$n_K = 4$			$n_K = 8$		
		Itr	u_s	k_1	Itr	u_s	k_1	Itr	u_s	k_1
2	2.00	43	0.229	1.957	27	0.299	1.955	27	0.338	1.955
4	1.00	32	0.273	1.957	26	0.322	1.955	30	0.349	1.955
8	0.50	28	0.304	1.956	26	0.338	1.955	32	0.354	1.955
16	0.25	36	0.324	1.956	36	0.349	1.955	43	0.355	1.955
32	0.13	49	0.339	1.956	49	0.353	1.955	55	0.354	1.955
64	0.06	62	0.348	1.956	48	0.354	1.955	66	0.353	1.955

known that no bulk flow is generated at Navier-Stokes limit and the gas temperature is governed by Fourier's law of heat conduction. However, under the non-equilibrium circumstances, thermal stress and thermal edge flows are induced by the rarefaction effects in the Knudsen layer [17].

We calculate the thermal flow at a challenge Knudsen number of $\sqrt{\pi}K/2 = 0.001$, when the characteristic flow length is chosen as the gap between the beam and chamber. In addition to the fact that the physical viscosity of gas is small, the magnitude of bulk velocity is also small; in this case high resolutions of both Knudsen layer and bulk region are necessary, otherwise, under-resolution of Knudsen layer and/or large numerical dissipation will introduce error at the same or even larger order as that of the bulk velocity and lead to completely wrong result.

The computational domain is partitioned by structured triangles with refinement in the vicinity of walls, see Figure 7(b), which is characterized by the total number of triangles N_Δ, the minimum cell size (the height of the local triangle) within Knudsen layer L_λ and the maximum cell size in bulk region L_b. The diffuse boundary condition is imposed on the solid surfaces. The DG scheme is employed in the numerical simulation, and details to implement GSIS are the same as those in the previous subsections, except here we use $K_{th} = 1$ for $n_K \leq 4$ and $K_{th} = 0.5$ for $n_K > 4$. The typical temperature field is shown in Figure 7(c) that are obtained on the mesh of $N_\Delta = 16104$, $L_\lambda = 0.16K$ and $L_b = 128K$ by DG schemes with $n_K = n_S = 6$.

We first consider the flow pattern in Figure 8, which are obtained by the DG scheme of different orders and on different spatial meshes. On the same spatial mesh with cell size of 0.08 mean free path in the Knudsen layer and 128 mean free path in the bulk region, completely different flow patterns are obtained by second-order and sixth-order DG schemes, see Figure 8(a) and (d). In the flow field predicted by the second-order scheme, four small vortices are generated around the lower-left corner of the beam; while another four are developed near the upper-right corner of the beam. At each of the lower-right and upper-left corners of the beam, three more vortices appear. Two small vortices are observed near the lower-right and upper-left corners of the chamber. On the other hand, from the sixth-order DG scheme, only eight vortices are developed, with each corner of the beam having two. Therefore, on this spatial grid ($L_\lambda = 0.08K$ and $L_b = 128K$), the second-order scheme produces larger errors; reducing the cell size by half is even not adequate, see Figure 8(b). It is also worth noticing that only increasing the order of DG scheme for synthetic equations,
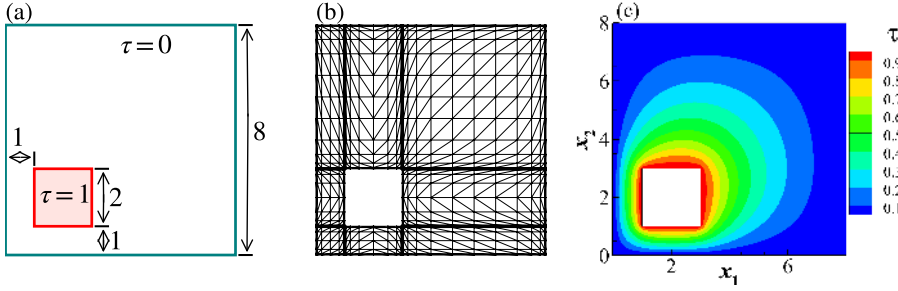
FIG. 7. *Two-dimensional thermal flow induced by squared hot beam encompassed in a squared cold chamber when $\sqrt{\pi}K/2 = 0.001$. Schematic illustrations of (a) geometry and (b) spatial discretization of structured triangles with refinement in the vicinity of solid walls. (c) typical temperature contours obtained by DG schemes of $n_K = n_S = 6$ and on spatial mesh of 16104 triangles. Note that all the flow properties are symmetric about the diagonal joining the lower-left and upper-right corners of the chamber.*
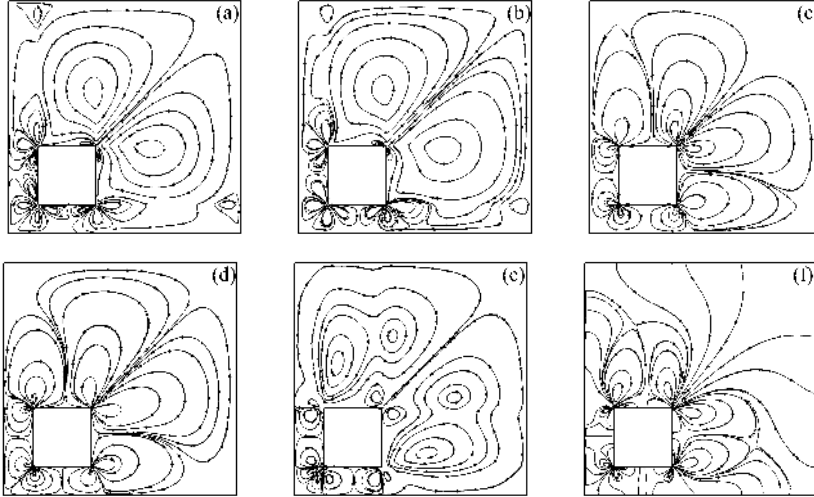

FIG. 8. *Streamlines obtained from DG schemes with different orders of accuracy and spatial meshes, in a two-dimensional thermal flow when $\sqrt{\pi}K/2 = 0.001$. (a) $L_\lambda = 0.08K$, $L_b = 128K$, and $n_K = n_S = 2$; (b) $L_\lambda = 0.04K$, $L_b = 64K$, and $n_K = n_S = 2$; (c) $L_\lambda = 0.32K$, $L_b = 256K$, and $n_K = n_S = 4$; (d) $L_\lambda = 0.08K$, $L_b = 128K$, and $n_K = n_S = 6$; (e) $L_\lambda = 0.08K$, $L_b = 128K$, $n_K = 2$ and $n_S = 6$; (f) $L_\lambda = 1.28K$, $L_b = 64K$, and $n_K = n_S = 4$.*

the predicted flow field is still quiet different from that when the mesoscopic and macroscopic equations are all solved using sixth-order DG scheme, see flow field in Figure 8(e) obtained on mesh of $L_\lambda = 0.08K$ and $L_b = 128K$ and schemes of $n_K = 2$ and $n_S = 6$. These results show the importance of resolving the Knudsen layer, where the non-equilibrium effects are the sources of gas motion.

To further demonstrate the importance of resolving the Knudsen layer, we use the fourth-order DG scheme to solve the kinetic and synthetic equations on two different meshes: Figure 8(c) is obtained when $L_\lambda = 0.32K$ and $L_b = 256K$, while Figure 8(f) is obtained when $L_\lambda = 1.28K$ and $L_b = 64K$. Figure 8(c) shows that, on the mesh where the Knudsen layer is resolved by relatively fine cells, the flow pattern with eight vortices are developed, which are very similar to the high resolution results in
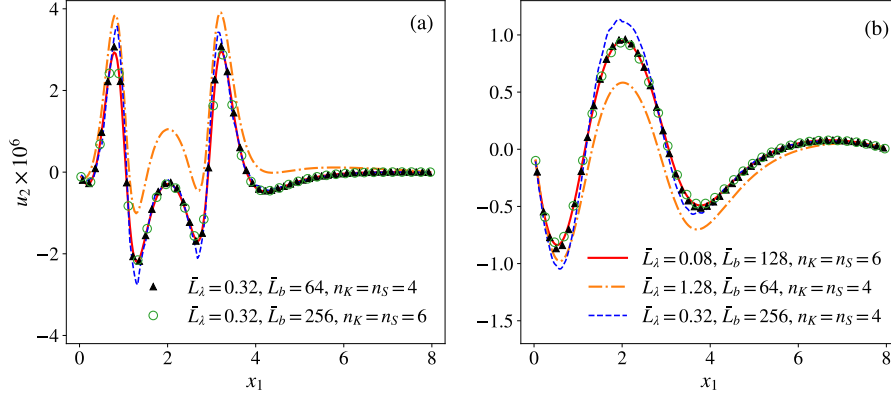
FIG. 9. *Flow velocities obtained from DG schemes with different orders of accuracy and meshes with different cell sizes, in a two-dimensional thermal flow when $\sqrt{\pi}K/2 = 0.001$. (a) vertical velocity u_2 along the horizontal line at $x_2 = 0.5$; (b) vertical velocity u_2 along the horizontal at $x_2 = 4.5$. $\bar{L}_\lambda$ and $\bar{L}_b$ are the minimum cell size in Knudsen layer and maximum cell size in bulk region, respectively, which are normalized by the mean free path of gas molecules, i.e. $\bar{L}_\lambda = L_\lambda/K$ and $\bar{L}_b = L_b/K$. The legends are applied to both figures.*

TABLE 3

The number of iterative steps (Itr), the mean velocity magnitude $|\bar{\boldsymbol{u}}|_{x_1=0.5}$ along the horizontal line $x_2 = 0.5$), as well as the computational time t_c under different combinations of spatial meshes and schemes, in a two-dimensional thermal flow with $\sqrt{\pi}K/2 = 0.001$. N_Δ is the number of triangles in a mesh, L_λ is the minimum cell size (triangle height) in Knudsen layer and L_b is the maximum cell size in bulk region. 'Err' is the relative error of the mean velocity magnitude $|\boldsymbol{u}|_{x_1=0.5}$ compared to the reference one obtained with $N_\Delta = 20424$, $L_\lambda = 0.08K$, $L_b = 128K$ and $n_K = n_S = 6$. The computational time t_c is the wall time cost by each case that runs on 8 Intel Xeon-E5-2680 CPUs using OpenMP for parallelism. Note that the the molecular velocity space is discretized by 8×8-point Gauss-Hermite quadrature [18].

N_Δ	L_λ/K	L_b/K	n_K	n_S	$|\bar{\boldsymbol{u}}| \times 10^6$	Err [%]	Itr	t_c [h]
8832	0.32	256	4	4	1.072	7.4	49	0.01
16104	0.16	128	4	4	1.060	6.2	46	0.02
24864	1.28	64	4	4	1.077	7.9	45	0.06
30176	0.32	64	4	4	1.013	1.5	48	0.04
33024	0.16	64	4	4	1.013	1.5	46	0.04
8832	0.32	256	6	6	0.982	1.6	57	0.07
10400	0.16	256	6	6	0.991	0.7	55	0.12
16104	0.16	128	6	6	0.991	0.7	54	0.18
20424	0.08	128	6	6	0.998	0	60	0.21

Figure 8(d), despite that the cell size in the bulk region is relatively large. However, when the bulk region is partitioned by fine cells but the Knudsen layer is under resolved by coarse cells, Figure 8(f) shows that although eight vortices can be observed, large errors appear near the walls, where gas seems to penetrate the solid walls.

To quantitatively show the accuracy and efficiency of numerical simulations, we plot the vertical velocity along the horizontal lines $x_2 = 0.5$ and $x_2 = 4.5$ in Figure 9. As the refinement of the spatial mesh and increment of the order of accuracy of the DG schemes, the flow properties converge to the reference ones (red solid lines) obtained on the mesh of $L_\lambda = 0.08K$ and $L_b = 128K$ and by sixth-order DG scheme. In Table 3 we list the mean velocity along the horizontal line: $|\bar{\boldsymbol{u}}| = \frac{1}{8} \int_0^8 |\boldsymbol{u}| (x_1, x_2 = 0.5) \, \mathrm{d}x_1$.

The relative error between the mean velocity from the mesh of $L_\lambda = 0.08K$ and $L_b = 128K$ by sixth-order DG scheme and the one from the mesh of $L_\lambda = 0.16K$ and $L_b = 256K$ by the same order schemes is within 1%, which indicates the accuracy of the reference result. The results for $|\bar{u}|$ with different spatial discretizations and different orders of DG schemes clearly demonstrate the importance of resolving the Knudsen layer. The iteration steps and computational time cost by each case are also listed in Table 3. For all the cases, GSIS can find the steady-state solution with 50 to 60 steps. The computational time to obtain accurate solution (with less than 2% error in $|\bar{u}|$ compared to the reference one) can be as little as several minutes. To the best of our knowledge, we are not aware of other numerical methods that are able to obtain accurate results within such a short time.

6. Conclusions. In summary, based on the Fourier stability analysis, we have rigorously proven that the GSIS can lead to fast convergence of steady-state solutions. Due to the presence of continuity equation in macroscopic synthetic equations, the convergence property is different to other diffusion synthetic schemes for neutron transport and thermal radiation. Therefore, in GSIS the macroscopic quantities can only be partly updated according to the solution of synthetic equations when the Knudsen number is large. On the other hand, we have proven that the GSIS asymptotically preserves the Navier-Stokes limit when the spatial cell size is able to capture the physical solution. That is, in the bulk region the spatial cell size can be $\Delta x \sim O(1)$ as long as this size is adequate to capture the hydrodynamics. However, in the vicinity of solid walls the physical solution requires $\Delta x \sim O(K)$ in order to capture the Knudsen layer structures. This does not pose any problems to GSIS since it is an implicit solver so non-uniform spatial grid can be used. These analytical results have been confirmed by several numerical examples in the present paper.

It should be emphasized that GSIS bears the advantage that the mesoscopic kinetic equation and macroscopic synthetic equations can be solved by different schemes. Since most of the cost will be spent on the kinetic solver due to additional discretization in molecular velocity space, in multi-scale simulation of rarefied gas dynamics, one can use high-order solver for synthetic equations, while low-order solver for kinetic equations, on a coarse spatial grid to save the computational time and cost,

REFERENCES

[1] M. L. Adams and E. W. Larsen, *Fast iterative methods for discrete-ordinates particle transport calculations*, Progr. Nucl. Energy, 40 (2002), pp. 3–159.
[2] P. L. Bhatnagar, E. P. Gross, and M. Krook, *A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems*, Phys. Rev., 94 (1954), pp. 511–525.
[3] G. A. Bird, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford Science Publications, Oxford University Press Inc, New York, 1994.
[4] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-uniform Gases*, Cambridge University Press, 1970.
[5] G. Dimarco and L. Pareschi, *Implicit-explicit linear multistep methods for stiff kinetic equations*, Siam J. Numer. Anal., 55 (2017), pp. 664–690.
[6] 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), p. 7625.
[7] F. Filbet and S. Jin, *An asymptotic preserving scheme for the ES-BGK model of the Boltzmann equation*, J. Sci. Comput., 46 (2011), pp. 204–224.
[8] H. Grad, *On the kinetic theory of rarefied gases*, Comm. Pure Appl. Math., 2 (1949), pp. 331–407.
[9] Z. Guo, K. Xu, and R. Wang, *Discrete unified gas kinetic scheme for all Knudsen number flows: Low-speed isothermal case*, Phys. Rev. E, 88 (2013), p. 033305.

[10] Z. L. Guo, J. Q. Li, and K. Xu, *On unified preserving properties of kinetic schemes*, arXiv:1909.04923, (2019).

[11] M. T. Ho, L. H. Zhu, L. Wu, P. Wang, Z. L. Guo, Z. H. Li, and Y. H. Zhang, *A multi-level parallel solver for rarefied gas flows in porous media*, Comput. Phys. Commun., 234 (2019), pp. 14–25.

[12] J. W. Hu and X. X. Zhang, *On a class of implicit-explicit Runge-Kutta schemes for stiff kinetic equations preserving the Navier-Stokes limit*, J. Sci. Comput., 73 (2017), pp. 797–818.

[13] S. Jiang and L. S. Luo, *Analysis and accurate numerical solutions of the integral equation derived from the linearized BGKW equation for the steady Couette flow*, J. Comput. Phys., 316 (2016), pp. 416–434.

[14] E. W. Larsen, J. Morel, and W. F. Miller, *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes*, J. Comput. Phys., 69 (1987), pp. 283 – 324.

[15] L. Mieussens, *Discrete velocity models and numerical schemes for the Boltzmann-BGK equation in plane and axisymmetric geometries*, J. Comput. Phys., 162 (2000), pp. 429–466.

[16] E. M. Shakhov, *Approximate kinetic equations in rarefied gas theory*, Fluid Dynamics, 3(1) (1968), pp. 112–115.

[17] Y. Sone, *Kinetic theory and fluid dynamics*, Birkhauser Boston, 2002.

[18] W. Su, S. Lindsay, H. H. Liu, and L. Wu, *Comparative study of the discrete velocity and lattice Boltzmann methods for rarefied gas flows through irregular channels*, Phys. Rev. E, 96 (2017), p. 023309.

[19] W. Su, P. Wang, H. Liu, and L. Wu, *Accurate and efficient computation of the Boltzmann equation for Couette flow: Influence of intermolecular potentials on Knudsen layer function and viscous slip coefficient*, J. Comput. Phys, 378 (2019), pp. 573–590.

[20] W. Su, P. Wang, Y. Zhang, and L. Wu, *Implicit Discontinuous Galerkin Method for the Boltzmann Equation*, Journal of Scientific Computing, 82 (2020), p. 39.

[21] W. Su, P. Wang, Y. H. Zhang, and L. Wu, *A high-order hybridizable discontinuous Galerkin method with fast convergence to steady-state solutions of the gas kinetic equation*, J. Comput. Phys., 376 (2019), pp. 973–991.

[22] W. Su, L. H. Zhu, P. Wang, Y. H. Zhang, and L. Wu, *Can we find steady-state solutions to multiscale rarefied gas flows within dozens of iterations?*, Journal of Computational Physics, 407 (2020), p. 109245.

[23] S. Taguchi and K. Aoki, *Rarefied gas flow around a sharp edge induced by a temperature field*, J. Fluid Mech., 694 (2012), pp. 191–224.

[24] D. Valougeorgis and S. Naris, *Acceleration schemes of the discrete velocity method: Gaseous flows in rectangular microchannels*, SIAM Journal on Scientific Computing, 25 (2003), pp. 534–552.

[25] P. Wang, M. T. Ho, L. Wu, Z. Guo, and Y. Zhang, *A comparative study of discrete velocity methods for low-speed rarefied gas flows*, Computers & Fluids, 161 (2018), pp. 33 – 46.

[26] L. Wu and X. J. Gu, *On the accuracy of macroscopic equations for linearized rarefied gas*, Advances in Aerodynamics, 2 (2020), p. 2.

[27] L. Wu, J. Zhang, H. H. Liu, Y. H. Zhang, and J. M. Reese, *A fast iterative scheme for the linearized Boltzmann equation*, J. Comput. Phys., 338 (2017), pp. 431–451.

[28] K. Xu and J. C. Huang, *A unified gas-kinetic scheme for continuum and rarefied flows*, J. Comput. Phys., 229 (2010), pp. 7747–7764.