CrossMark

An asymptotic-preserving scheme for the semiconductor Boltzmann equation toward the energy-transport limit

Jingwei Hu [a,1], Li Wang [b,∗]

[a] *The Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin, 201 East 24th St, Stop C0200, Austin, TX 78712, United States*
[b] *Department of Mathematics, University of California, Los Angeles, 520 Portola Plaza, Los Angeles, CA 90095, United States*

ARTICLE INFO

ABSTRACT

We design an asymptotic-preserving scheme for the semiconductor Boltzmann equation which leads to an energy-transport system for electron mass and energy as mean free path goes to zero. As opposed to the classical drift-diffusion limit where the stiff collisions are all in one scale, new difficulties arise in the *two-scale* stiff collision terms because the simple BGK penalization [15] fails to drive the solution to the correct limit. We propose to set up *a spatially dependent threshold on the penalization of the stiffer collision operator* such that the evolution of the solution resembles a Hilbert expansion at the continuous level. Formal asymptotic analysis and numerical results confirm the efficiency and accuracy of our scheme.

1. Introduction

The semiconductor Boltzmann equation describes the transport of charge carriers in semiconductor devices. By incorporating the quantum mechanical effects semiclassically, it provides accurate description of the physics at the kinetic level [32,8,27]. A non-dimensional form of this equation usually contains small parameters such as the mean free path or time, which pose tremendous computational challenge since one has to numerically resolve the small scales. To save the computational cost, in the past decades various macroscopic models were derived from the Boltzmann equation based on different dominating effects. One of the widely-used models is the drift-diffusion equation (a mass continuity equation for electrons or holes) [34,17]. Ideally, if the parameters in the kinetic equation are uniformly small in the entire domain of interest, then it suffices to solve the macroscopic models [29,9,35]. In practical applications, however, the validity of these models may break down in part of the domain and one has to resort to kinetic equations [5,6]. To handle such multiscale phenomena, a typical approach is to use domain decomposition [4,30], but the task of setting up a good interface condition coupling the two models at different scales can be highly non-trivial. An alternative approach seeks a unified scheme for the kinetic equation such that when the parameter goes to zero it automatically becomes a macroscopic solver. This design concept leads to the asymptotic-preserving (AP) scheme [22], first proposed by S. Jin for transport equations in diffusive regimes [21].

In the semiconductor framework, the first AP scheme was introduced in [23] for the Boltzmann equation with an anisotropic collision term. The scheme was further improved in [11]. Recently a high-order scheme was constructed in [13],

---

which relaxed the strict parabolic stability condition to a hyperbolic one. All these works deal with a *single, linear* collision operator with *smoothed kernel* which uniquely defines an equilibrium state. As a result, the corresponding macroscopic equation is in the form of a drift-diffusion equation. Although this equation gives satisfactory simulation results for semiconductor devices on the micrometer scale, it is not able to capture the hot-electron effects in submicron devices [28]. Some works in the high field scaling considered this effect, but it only makes sense for the situation where the field effect is strong enough to balance the collision [25,7].

In this work, we consider a more realistic semiconductor Boltzmann equation [1,10]. By assuming the elastic collision as dominant, and electron–electron correlation as sub-dominant effects, one can pass on the asymptotic limit to obtain an energy-transport (ET) model, which is a system of conservation laws for electron mass and energy with fluxes defined through a constitutive relation [26]. To design an AP scheme for such kinetic equation, we face twofold challenge: 1. the convection terms are stiff; 2. two stiff collision terms live on different scales. The convection terms can be treated by an even–odd decomposition as in [24,23]. For the collision terms, due to their complicated forms, we choose to penalize them with a suitable BGK operator [15]. However, unlike the usual collision operator with smoothed kernel, the leading elastic operator has non-unique null space. Only when the electron–electron operator at next level takes into effect, the solution can be eventually driven to a fixed equilibrium − a Fermi–Dirac distribution. A closer examination of the asymptotic behavior of the numerical solution reveals that the penalization should be performed wisely, otherwise it won't capture the correct limit. To this end, we propose a *thresholded penalization* scheme. Simply speaking, when a certain threshold is satisfied, we turn off the leading order mechanism and move to the next order, which in some sense resembles the Hilbert expansion at the continuous level. We will show that this new scheme, under certain assumptions, satisfies the following four properties required in a standard AP scheme: let α denote the small parameter; Δt, Δx, and Δk be the time step and mesh size in spatial and wave vector (momentum) domain, then

1. for fixed α, it is consistent to the kinetic equation when $\Delta t, \Delta x, \Delta k \to 0$;
2. for fixed $\Delta t, \Delta x, \Delta k$, it becomes a discretization to the limiting ET system when $\alpha \to 0$;
3. it is uniformly stable for a wide range of α, from $\alpha = O(1)$ to $\alpha \ll 1$;
4. implicit terms can be implemented explicitly (free of Newton-type solvers).

The rest of the paper is organized as follows. In the next section we give a brief review of the semiconductor Boltzmann equation and the derivation of its ET limit in diffusive regimes. Section 3 is devoted to a detailed description of our AP schemes. We first consider the spatially homogeneous case with an emphasis on the two-scale collision terms, and then include the spatial dependence to treat the full problem. In either case, the asymptotic property of the numerical solution is carefully analyzed. In Section 4 we present several numerical examples to illustrate the efficiency, accuracy, and AP properties of the schemes. Finally, the paper is concluded in Section 5.

2. The semiconductor Boltzmann equation and its energy-transport limit

The Boltzmann transport equation that describes the evolution of electrons in the conduction band of a semiconductor reads [32,8,27]

$$\partial_t f + \frac{1}{\hbar}\nabla_k \varepsilon(k) \cdot \nabla_x f + \frac{q}{\hbar}\nabla_x V(x,t) \cdot \nabla_k f = \mathcal{Q}(f), \quad x \in \Omega \subset \mathbb{R}^d,\ k \in B \subset \mathbb{R}^d,\ d = 2, 3, \tag{2.1}$$

where $f(x, k, t)$ is the electron distribution function of position x, wave vector k, and time t. $\hbar$ is the reduced Planck constant, and q is the positive elementary charge. The first Brillouin zone B is the primitive cell in the reciprocal lattice of the crystal. For simplicity, we will restrict to the parabolic band approximation, where B can be extended to the whole space $B = \mathbb{R}^d$, and the energy-band diagram $\varepsilon(k)$ is given by

$$\varepsilon(k) = \frac{\hbar^2}{2m^*}|k|^2,$$

with m^* being the effective mass of electrons.

In principle, the electrostatic potential $V(x, t)$ is produced self-consistently by the electron density with a fixed ion background of doping profile $h(x)$ through the Poisson equation:

$$\epsilon_0 \nabla_x \big(\epsilon_r(x)\nabla_x V(x,t)\big) = q\big(\rho(x,t) - h(x)\big), \tag{2.2}$$

where

$$\rho(x,t) := \int\limits_{\mathbb{R}^d} f(x,k,t)\frac{g}{(2\pi)^d}\,\mathrm{d}k$$

is the electron density (the spin degeneracy $g = 2s + 1$, with $s = 1/2$ being the spin of electrons). ϵ_0 and $\epsilon_r(x)$ are the vacuum and the relative material permittivities. The doping profile $h(x)$ takes into account the impurities due to acceptor and donor ions in the semiconductor device.

The collision operator $\mathcal{Q}$ models three different effects:

$$\mathcal{Q} = \mathcal{Q}_{\text{imp}} + \mathcal{Q}_{\text{ph}} + \mathcal{Q}_{\text{ee}},$$

where $\mathcal{Q}_{\text{imp}}$ and $\mathcal{Q}_{\text{ph}}$ account for the interactions between electrons and the lattice defects caused by ionized impurities and crystal vibrations (also called phonons); $\mathcal{Q}_{\text{ee}}$ describes the correlations between electrons themselves. We will specify their forms later.

2.1. The nondimensionalized semiconductor Boltzmann equation

In this paper, we are interested in a high energy scale [2,1] at which the relative energy gain or loss of electron energy during a phonon collision is very small, then it is valid to consider an elastic approximation of the electron–phonon interactions $\mathcal{Q}_{\text{ph}}$. Combining elastic collisions ($\mathcal{Q}_{\text{imp}}$, which is elastic by nature, and the elastic part of $\mathcal{Q}_{\text{ph}}$) and performing careful nondimensionalization on (2.1), one arrives at the following scaled semiconductor Boltzmann equation in diffusive regime [10]:

$$\partial_t f + \frac{1}{\alpha}(\nabla_k \varepsilon \cdot \nabla_x f + \nabla_x V \cdot \nabla_k f) = \frac{1}{\alpha^2}\mathcal{Q}_{\text{el}}(f) + \frac{1}{\alpha}\mathcal{Q}_{\text{ee}}(f) + \mathcal{Q}_{\text{ph}}^{\text{inel}}(f), \tag{2.3}$$

where the parameter α can be interpreted as the scaled mean-free path for elastic collisions. The energy-band diagram $\varepsilon(k)$ is now simply

$$\varepsilon(k) = \frac{1}{2}|k|^2. \tag{2.4}$$

The elastic collision $\mathcal{Q}_{\text{el}}$ and the electron–electron collision $\mathcal{Q}_{\text{ee}}$ are given by (the exact form of $\mathcal{Q}_{\text{ph}}^{\text{inel}}$ will not be needed in the following discussion and is thus omitted):

$$\mathcal{Q}_{\text{el}}(f)(k) = \int_{\mathbb{R}^d} \Phi_{\text{el}}(k, k')\delta(\varepsilon' - \varepsilon)(f' - f)\,\mathrm{d}k', \tag{2.5}$$

$$\mathcal{Q}_{\text{ee}}(f)(k) = \int_{\mathbb{R}^{3d}} \Phi_{\text{ee}}(k, k_1, k', k_1')\delta(\varepsilon' + \varepsilon_1' - \varepsilon - \varepsilon_1)\delta(k' + k_1' - k - k_1)$$

$$\times \left[f'f_1'(1 - \eta f)(1 - \eta f_1) - ff_1(1 - \eta f')(1 - \eta f_1') \right]\mathrm{d}k_1\mathrm{d}k'\mathrm{d}k_1', \tag{2.6}$$

where δ is the Dirac measure. ε', ε_1, f', f_1, $\cdots$ are short notations for $\varepsilon(k')$, $\varepsilon(k_1)$, $f(x, k', t)$, $f(x, k_1, t)$, $\cdots$. η is the typical distribution function scale characterizing the degree of degeneracy of the system. The scattering matrix $\Phi_{\text{el}}(k, k')$ is symmetric in k and k': $\Phi_{\text{el}}(k, k') = \Phi_{\text{el}}(k', k)$; $\Phi_{\text{ee}}(k, k_1, k', k_1')$ is pairwise symmetric in four variables: $\Phi_{\text{ee}}(k, k_1, k', k_1') = \Phi_{\text{ee}}(k_1, k, k', k_1') = \Phi_{\text{ee}}(k', k_1', k, k_1)$. They are determined by the underlying interaction laws. The Poisson equation (2.2) becomes

$$\mathcal{D}_0\nabla_x(\epsilon_r(x)\nabla_x V(x, t)) = \rho(x, t) - h(x), \tag{2.7}$$

where $\mathcal{D}_0$ is the square of the scaled Debye length, and

$$\rho(x, t) = \int_{\mathbb{R}^d} f(x, k, t)\,\mathrm{d}k.$$

2.2. The energy-transport limit

This subsection is devoted to a formal derivation of the asymptotic limit of (2.3) as $\alpha \to 0$. Our approach, following that of [10], is a combination of the Hilbert expansion and the moment method which mainly rely on the properties of the collision operators $\mathcal{Q}_{\text{el}}$ and $\mathcal{Q}_{\text{ee}}$. To this end, we list them at first. They will also be useful in designing numerical schemes.

Proposition 1. *(See [1].)*

1. *For any "regular" test function $g(k)$,*

$$\int_{\mathbb{R}^d} \mathcal{Q}_{\text{el}}(f)g\,\mathrm{d}k = -\frac{1}{2}\int_{\mathbb{R}^{2d}} \Phi_{\text{el}}(k, k')\delta(\varepsilon' - \varepsilon)(f' - f)(g' - g)\,\mathrm{d}k\mathrm{d}k'.$$

In particular, for any $g(\varepsilon(k))$,

$$\int_{\mathbb{R}^d} \mathcal{Q}_{\text{el}}(f)g(\varepsilon)\,\mathrm{d}k = 0.$$

2. $\mathcal{Q}_{el}(f)$ is a self-adjoint, non-positive operator on $L^2(\mathbb{R}^d)$.
3. The null space of $\mathcal{Q}_{el}(f)$ is given by

$$\mathcal{N}(\mathcal{Q}_{el}) = \big\{ f\big(\varepsilon(k)\big), \ \forall f \big\}.$$

4. The orthogonal complement of $\mathcal{N}(\mathcal{Q}_{el})$ is

$$\mathcal{N}(\mathcal{Q}_{el})^\perp = \left\{ g(k) \ \Big| \ \int_{S_\varepsilon} g(k)\, dN_\varepsilon(k) = 0, \ \forall \varepsilon \in R \right\},$$

where the integral is defined through the "coarea formula" [14]: for any "regular" functions f and $\varepsilon(k): \mathbb{R}^d \to R$, it holds

$$\int_{\mathbb{R}^d} f(k)\, dk = \int_R \left(\int_{S_\varepsilon} f(k)\, dN_\varepsilon(k) \right) d\varepsilon. \tag{2.8}$$

Here $S_\varepsilon = \{k \in \mathbb{R}^d, \ \varepsilon(k) = \varepsilon\}$ denotes the surface of constant energy ε, $dN_\varepsilon(k)$ is the coarea measure, and $N(\varepsilon) := \int_{S_\varepsilon} dN_\varepsilon(k)$ is the energy density-of-states. Under the parabolic band approximation (2.4), (2.8) is just a spherical transformation:

$$\int_{\mathbb{R}^d} f(k)\, dk = \int_0^\infty \left(\int_{\mathbb{S}^{d-1}} f(|k|\sigma)|k|^{d-2}\, d\sigma \right) d\varepsilon,$$

and $N(\varepsilon) = \int_{\mathbb{S}^{d-1}} |k|^{d-2}\, d\sigma = m(\mathbb{S}^{d-1})|k|^{d-2}$.

5. The range of $\mathcal{Q}_{el}$ is $\mathcal{R}(\mathcal{Q}_{el}) = \mathcal{N}(\mathcal{Q}_{el})^\perp$. The operator is invertible as an operator from $\mathcal{N}(\mathcal{Q}_{el})^\perp$ onto $\mathcal{R}(\mathcal{Q}_{el}) = \mathcal{N}(\mathcal{Q}_{el})^\perp$. Its inverse is denoted by $\mathcal{Q}_{el}^{-1}$.
6. For any $\psi(\varepsilon(k))$, we have $\mathcal{Q}_{el}(f\psi) = \psi \mathcal{Q}_{el}(f)$ and $\mathcal{Q}_{el}^{-1}(f\psi) = \psi \mathcal{Q}_{el}^{-1}(f)$.

Proposition 2. *(See [2].)*

1. For any "regular" test function $g(k)$,

$$\int_{\mathbb{R}^d} \mathcal{Q}_{ee}(f) g\, dk = -\frac{1}{4} \int_{\mathbb{R}^{4d}} \Phi_{ee}(k, k_1, k', k_1') \delta(\varepsilon' + \varepsilon_1' - \varepsilon - \varepsilon_1) \delta(k' + k_1' - k - k_1)$$
$$\times \big[f' f_1'(1 - \eta f)(1 - \eta f_1) - f f_1(1 - \eta f')(1 - \eta f_1') \big] (g' + g_1' - g - g_1)\, dk dk_1 dk' dk_1'.$$

In particular, we have the conservation of mass and energy

$$\int_{\mathbb{R}^d} \mathcal{Q}_{ee}(f)\, dk = \int_{\mathbb{R}^d} \mathcal{Q}_{ee}(f)\varepsilon\, dk = 0.$$

2. H-theorem: let $H(f) = \ln(f/(1 - \eta f))$, then $\int_{\mathbb{R}^d} \mathcal{Q}_{ee}(f) H(f)\, dk \leq 0$, and if $f = f(\varepsilon(k))$,

$$\int_{\mathbb{R}^d} \mathcal{Q}_{ee}(f) H(f)\, dk = 0 \quad \Longleftrightarrow \quad f = M_q(\varepsilon(k)) \quad \Longleftrightarrow \quad \mathcal{Q}_{ee}(f) = 0,$$

where

$$M_q\big(\varepsilon(k)\big) = \frac{1}{\eta} \frac{1}{z^{-1} e^{\frac{\varepsilon(k)}{T}} + 1} \tag{2.9}$$

is the Fermi–Dirac distribution function [33]. The variables z and T are the fugacity and temperature. Alternatively, M_q can be defined in terms of the chemical potential $\mu = T \ln z$ and T.

We also need the properties of the operator $\overline{\mathcal{Q}_{ee}}$, an energy space counterpart of $\mathcal{Q}_{ee}$: for any $F(\varepsilon(k))$,

$$\overline{\mathcal{Q}_{ee}}(F)(\varepsilon) := \int_{S_\varepsilon} \mathcal{Q}_{ee}(F)(k)\, dN_\varepsilon(k).$$

Proposition 3. *(See [1].)*

1. *Conservation of mass and energy*

$$\int_R \overline{\mathcal{Q}_{ee}}(F)\,d\varepsilon = \int_R \overline{\mathcal{Q}_{ee}}(F)\varepsilon\,d\varepsilon = 0.$$

2. *H-theorem: let $H(F) = \ln(F/(1 - \eta F))$, then $\int_R \overline{\mathcal{Q}_{ee}}(F)H(F)\,d\varepsilon \leq 0$, and*

$$\int_R \overline{\mathcal{Q}_{ee}}(F)H(F)\,d\varepsilon = 0 \iff F = M_q(\varepsilon) \iff \overline{\mathcal{Q}_{ee}}(F) = 0.$$

Now that the mathematical preliminaries are set up, we are ready to derive the macroscopic limit. The main result is summarized in the following theorem.

Theorem 4. *(See [10].) In Eq. (2.3), when $\alpha \to 0$, the solution f formally tends to a Fermi–Dirac distribution function (2.9), with the position and time dependent fugacity $z(x, t)$ and temperature $T(x, t)$ satisfying the so-called energy-transport (ET) model:*

$$\partial_t \begin{pmatrix} \rho \\ \rho\mathcal{E} \end{pmatrix} + \begin{pmatrix} \nabla_x \cdot j_\rho \\ \nabla_x \cdot j_\mathcal{E} \end{pmatrix} - \begin{pmatrix} 0 \\ \nabla_x V \cdot j_\rho \end{pmatrix} = \begin{pmatrix} 0 \\ W_{ph}^{inel} \end{pmatrix}, \tag{2.10}$$

where the density ρ and energy $\mathcal{E}$ are defined as

$$\rho(z, T) = \int_{\mathbb{R}^d} f\,dk = \int_{\mathbb{R}^d} M_q\,dk, \qquad \mathcal{E}(z, T) = \frac{1}{\rho}\int_{\mathbb{R}^d} f\varepsilon\,dk = \frac{1}{\rho}\int_{\mathbb{R}^d} M_q\varepsilon\,dk; \tag{2.11}$$

the fluxes j_ρ and $j_\mathcal{E}$ are given by

$$\begin{pmatrix} j_\rho(z, T) \\ j_\mathcal{E}(z, T) \end{pmatrix} = -\begin{pmatrix} \mathcal{D}_{11} & \mathcal{D}_{12} \\ \mathcal{D}_{21} & \mathcal{D}_{22} \end{pmatrix} \begin{pmatrix} \frac{\nabla_x z}{z} - \frac{\nabla_x V}{T} \\ \frac{\nabla_x T}{T^2} \end{pmatrix} \tag{2.12}$$

with the diffusion matrices

$$\mathcal{D}_{ij} = \int_{\mathbb{R}^d} \nabla_k \varepsilon \otimes \mathcal{Q}_{el}^{-1}(-\nabla_k \varepsilon)M_q(1 - \eta M_q)\varepsilon^{i+j-2}\,dk; \tag{2.13}$$

and the energy relaxation operator W_{ph}^{inel} is

$$W_{ph}^{inel}(z, T) = \int_{\mathbb{R}^d} \mathcal{Q}_{ph}^{inel}(M_q)\varepsilon\,dk. \tag{2.14}$$

Proof. Inserting the Hilbert expansion $f = f_0 + \alpha f_1 + \alpha^2 f_2 + \ldots$ into Eq. (2.3) and collecting equal powers of α leads to

$$O(\alpha^{-2}): \quad \mathcal{Q}_{el}(f_0) = 0, \tag{2.15}$$

$$O(\alpha^{-1}): \quad \mathcal{Q}_{el}(f_1) = \nabla_k \varepsilon \cdot \nabla_x f_0 + \nabla_x V \cdot \nabla_k f_0 - \mathcal{Q}_{ee}(f_0). \tag{2.16}$$

From (2.15) and Proposition 1 (3), we know there exists some function F such that

$$f_0(x, k, t) = F(x, \varepsilon(k), t).$$

Plugging f_0 into (2.16):

$$\mathcal{Q}_{el}(f_1) = \nabla_k \varepsilon \cdot (\nabla_x F + \nabla_x V \partial_\varepsilon F) - \mathcal{Q}_{ee}(F). \tag{2.17}$$

The solvability condition for f_1 (Proposition 1 (4–5)) implies

$$\int_{S_\varepsilon} \nabla_k \varepsilon \cdot (\nabla_x F + \nabla_x V \partial_\varepsilon F)\,dN_\varepsilon(k) = \int_{S_\varepsilon} \mathcal{Q}_{ee}(F)\,dN_\varepsilon(k).$$

Clearly the left hand side of above equation is equal to zero ($\nabla_k \varepsilon$ is odd in k), so

$$\int_{S_\varepsilon} \mathcal{Q}_{ee}(F)\,dN_\varepsilon(k) = \overline{\mathcal{Q}_{ee}}(F) = 0.$$

By Proposition 3 (2), we know that F is a Fermi–Dirac distribution M_q with position and time dependent $z(x,t)$ and $T(x,t)$. Therefore, $\mathcal{Q}_{ee}(F)$ itself is equal to zero by Proposition 2 (2), and (2.17) reduces to

$$\mathcal{Q}_{el}(f_1) = \nabla_k \varepsilon \cdot (\nabla_x M_q + \nabla_x V \partial_\varepsilon M_q).$$

Then using Proposition 1 (5–6), we have

$$f_1 = -\mathcal{Q}_{el}^{-1}(-\nabla_k \varepsilon) \cdot (\nabla_x M_q + \nabla_x V \partial_\varepsilon M_q). \tag{2.18}$$

Going back to the original equation (2.3), multiplying both sides by $(1, \varepsilon(k))^T$ and integrating w.r.t. k gives:

$$\int_{\mathbb{R}^d} \left[\partial_t f + \frac{1}{\alpha}(\nabla_k \varepsilon \cdot \nabla_x f + \nabla_x V \cdot \nabla_k f) \right] \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} dk = \int_{\mathbb{R}^d} \left[\frac{1}{\alpha^2}\mathcal{Q}_{el}(f) + \frac{1}{\alpha}\mathcal{Q}_{ee}(f) + \mathcal{Q}_{ph}^{inel}(f) \right] \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} dk. \tag{2.19}$$

Terms involving $\mathcal{Q}_{el}(f)$ and $\mathcal{Q}_{ee}(f)$ vanish due to Propositions 1 (1) and 2 (1). Since $\mathcal{Q}_{ph}^{inel}$ conserves mass [8], (2.19) simplifies to

$$\partial_t \begin{pmatrix} \rho \\ \rho\mathcal{E} \end{pmatrix} + \int_{\mathbb{R}^d} \frac{1}{\alpha}(\nabla_k \varepsilon \cdot \nabla_x f + \nabla_x V \cdot \nabla_k f) \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} dk = \begin{pmatrix} 0 \\ W_{ph}^{inel}(f) \end{pmatrix}, \tag{2.20}$$

where $W_{ph}^{inel}(f) = \int_{\mathbb{R}^d} \mathcal{Q}_{ph}^{inel}(f)\varepsilon \, dk$.

From the previous discussion, we know $f = f_0 + \alpha f_1 + \dots$ with f_0 being the Fermi–Dirac distribution (2.9), and f_1 given by (2.18). Plugging f into (2.20), to the leading order we have ($O(\alpha^{-1})$ term drops out since f_0 is an even function in k):

$$\partial_t \begin{pmatrix} \rho \\ \rho\mathcal{E} \end{pmatrix} + \int_{\mathbb{R}^d} (\nabla_k \varepsilon \cdot \nabla_x f_1 + \nabla_x V \cdot \nabla_k f_1) \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} dk = \begin{pmatrix} 0 \\ W_{ph}^{inel}(M_q) \end{pmatrix}. \tag{2.21}$$

Utilizing the special form of M_q, one can rewrite f_1 as

$$f_1 = -\mathcal{Q}_{el}^{-1}(-\nabla_k \varepsilon) \cdot \left(\frac{\nabla_x z}{z} - \frac{\nabla_x V}{T} + \varepsilon \frac{\nabla_x T}{T^2} \right) M_q(1 - \eta M_q).$$

Then a simple manipulation of (2.21) yields the ET system (2.10). □

The ET model (2.10) is widely used in practical and industrial applications (see [28] for a review and references therein). It can also be derived from the Boltzmann equation through the so-called SHE (Spherical Harmonics Expansion) model [16]. In either case, the rigorous theory behind the formal limit is still an open question.

Remark 5. If the electron–electron interaction $\mathcal{Q}_{ee}$ is assumed as one of the dominant terms in (2.3), i.e., the same order as elastic collision $\mathcal{Q}_{el}$ (which could be the physically relevant situation for very dense electrons), one can still derive the ET model (2.10) via a similar procedure [2], but the diffusion coefficients $\mathcal{D}_{ij}$ are different. A rigorous result is available in this case [3].

Remark 6. If we consider even longer time scale such that the energy relaxation term W_{ph}^{inel} equilibrates the electron temperature T to lattice temperature T_L, and further assume that M_q is the classical Maxwellian, then (2.10) reduces to a single equation

$$\partial_t \rho + \nabla_x \cdot \left[-\mathcal{D}_{11}\left(\frac{\nabla_x \rho}{\rho} - \frac{\nabla_x V}{T_L} \right) \right] = 0.$$

This is the classical drift-diffusion model [36,34].

Remark 7. Unlike the classical statistics, given macroscopic variables ρ and $\mathcal{E}$ (2.11), finding the corresponding Fermi–Dirac distribution (2.9) is not a trivial issue. Under the parabolic approximation (2.4), ρ and $\mathcal{E}$ are related to z and T via [18]

$$\begin{cases} \rho = \dfrac{(2\pi T)^{\frac{d}{2}}}{\eta} \mathcal{F}_{\frac{d}{2}}(z), \\[2ex] \mathcal{E} = \dfrac{d}{2} T \dfrac{\mathcal{F}_{\frac{d}{2}+1}(z)}{\mathcal{F}_{\frac{d}{2}}(z)}, \end{cases} \tag{2.22}$$

where $\mathcal{F}_\nu(z)$ is the Fermi–Dirac function of order ν

$$\mathcal{F}_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}}{z^{-1}e^x + 1} \, dx, \quad 0 < z < \infty, \tag{2.23}$$

and $\Gamma(\nu)$ is the Gamma function. For small z ($0 < z < 1$), the integrand in (2.23) can be expanded in powers of z:

$$\mathcal{F}_\nu(z) = \sum_{n=1}^\infty (-1)^{n+1} \frac{z^n}{n^\nu} = z - \frac{z^2}{2^\nu} + \frac{z^3}{3^\nu} - \dots.$$

Thus, when $z \ll 1$, $\mathcal{F}_\nu(z)$ behaves like z itself and one recovers the classical limit.

3. Asymptotic-preserving (AP) schemes for the semiconductor Boltzmann equation

Eq. (2.3) contains three different scales. To overcome the stiffness induced by the $O(\alpha^{-2})$ and $O(\alpha^{-1})$ terms, a fully implicit scheme would be desirable. However, neither the collision operators nor the convection terms are easy to solve implicitly. Our goal is to design an appropriate numerical scheme that is uniformly stable in both kinetic and diffusive regimes, i.e., works for all values of α ranging from $\alpha \sim O(1)$ to $\alpha \ll 1$, while the implicit terms can be treated explicitly.

We will first consider a spatially homogeneous case with an emphasis on the collision operators, and then include the spatial dependence to treat the convection terms. To facilitate the presentation, we always make the following assumptions without further notice:

1. The inelastic collision operator $\mathcal{Q}_{ph}^{inel}$ in (2.3) is assumed to be zero, since it is the weakest effect and its appearance won't bring extra difficulties to numerical schemes.
2. The scattering matrices Φ_{el} and Φ_{ee} are *rotationally invariant*:

$$\Phi_{el}(k, k') = \Phi_{el}(|k|, |k'|), \qquad \Phi_{ee}(k, k_1, k', k_1') = \Phi_{ee}(|k|, |k_1|, |k'|, |k_1'|).$$

Then it is not difficult to verify that (see Proposition 1 (4))

$$\mathcal{Q}_{el}(f)(k) = \lambda_{el}(\varepsilon)([f](\varepsilon) - f(k)), \tag{3.1}$$

where

$$\lambda_{el}(\varepsilon(k)) := \int_{\mathbb{R}^d} \Phi_{el}(k, k')\delta(\varepsilon' - \varepsilon) \, dk' = \Phi_{el}(|k|, |k|)N(\varepsilon),$$

and $[f](\varepsilon(k))$ is the mean value of f over sphere S_ε:

$$[f](\varepsilon(k)) := \frac{1}{N(\varepsilon)} \int_{S_\varepsilon} f(k) \, dN_\varepsilon(k).$$

In particular, for any odd function $f(k)$,

$$\mathcal{Q}_{el}(f)(k) = -\lambda_{el}(\varepsilon)f(k), \qquad \mathcal{Q}_{el}^{-1}(f)(k) = -\frac{1}{\lambda_{el}(\varepsilon)}f(k).$$

This observation is crucial in designing our AP schemes.

3.1. The spatially homogeneous case

In the spatially homogeneous case, Eq. (2.3) reduces to

$$\partial_t f = \frac{1}{\alpha^2}\mathcal{Q}_{el}(f) + \frac{1}{\alpha}\mathcal{Q}_{ee}(f), \tag{3.2}$$

where f only depends on k and t. An explicit discretization of (3.2), e.g., the forward Euler scheme, suffers from severe stability constraints: Δt has to be smaller than $O(\alpha^2)$. Implicit schemes do not have such a restriction, but require some sort of iteration solvers for $\mathcal{Q}_{el}$ and $\mathcal{Q}_{ee}$ which can be quite complicated.

To tackle these two stiff terms, we adopt the penalization idea in [15], i.e., penalize both $\mathcal{Q}_{el}$ and $\mathcal{Q}_{ee}$ by their corresponding "BGK" operators:

$$\partial_t f = \frac{\mathcal{Q}_{el}(f) - \beta_{el}(M_{el} - f)}{\alpha^2} + \frac{\beta_{el}(M_{el} - f)}{\alpha^2} + \frac{\mathcal{Q}_{ee}(f) - \beta_{ee}(M_{ee} - f)}{\alpha} + \frac{\beta_{ee}(M_{ee} - f)}{\alpha}. \tag{3.3}$$

Using the properties of $\mathcal{Q}_{ee}$ and $\mathcal{Q}_{el}$ from the last section, M_{ee} can be naturally chosen as the Fermi–Dirac distribution M_q (in the homogeneous case ρ and $\mathcal{E}$ are conserved, so M_q is an absolute Maxwellian and can be obtained from the initial

condition), whereas M_{el} can in principle be any function of ε such that it shares the same density and energy as f (the choice of M_{el} is not essential as we shall see, and we will get back to this when we consider the spatially inhomogeneous case). As the goal of penalization is to make the residue $\mathcal{Q}_\bullet(f) - \beta_\bullet(M_\bullet - f)$ as small as possible so that it is non-stiff or less stiff, and $\mathcal{Q}_{el}$, $\mathcal{Q}_{ee}$ can be expressed symbolically as

$$\mathcal{Q}_{el}(f)(k) = \mathcal{Q}_{el}^+(f)(\varepsilon) - \lambda_{el}(\varepsilon)f(k); \qquad \mathcal{Q}_{ee}(f)(k) = \mathcal{Q}_{ee}^+(f)(k) - \mathcal{Q}_{ee}^-(f)(k)f(k),$$

we choose

$$\beta_{el} \approx \max_\varepsilon \lambda_{el}(\varepsilon); \qquad \beta_{ee} \approx \max_k \mathcal{Q}_{ee}^-(f)(k). \tag{3.4}$$

Other choices are also possible [15]. Generally speaking, we only need the coefficient to be a rough estimate of the Frechet derivative of the collision operator around equilibrium.

Therefore, a first-order IMEX scheme for (3.3) is written as

$$\frac{f^{n+1} - f^n}{\Delta t} = \frac{\mathcal{Q}_{el}(f^n) - \beta_{el}(M_{el} - f^n)}{\alpha^2} + \frac{\beta_{el}(M_{el} - f^{n+1})}{\alpha^2} + \frac{\mathcal{Q}_{ee}(f^n) - \beta_{ee}(M_q - f^n)}{\alpha} + \frac{\beta_{ee}(M_q - f^{n+1})}{\alpha}. \tag{3.5}$$

3.1.1. Asymptotic properties of the numerical solution

To better understand the asymptotic behavior of the numerical solution, in this subsection we assume that $\mathcal{Q}_{ee}(f) = M_q - f$. Then scheme (3.5) becomes

$$\frac{f^{n+1} - f^n}{\Delta t} = \frac{\mathcal{Q}_{el}(f^n) - \beta_{el}(M_{el} - f^n)}{\alpha^2} + \frac{\beta_{el}(M_{el} - f^{n+1})}{\alpha^2} + \frac{(1 - \beta_{ee})(M_q - f^n)}{\alpha} + \frac{\beta_{ee}(M_q - f^{n+1})}{\alpha}.$$

This is equivalent to

$$f^{n+1} = \frac{1 + \frac{\Delta t}{\alpha^2}(\beta_{el} - \lambda_{el}) + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} f^n + \frac{\frac{\Delta t}{\alpha^2}}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} \mathcal{Q}_{el}^+(f^n) + \frac{\frac{\Delta t}{\alpha}}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} M_q$$

$$= \frac{1 + \frac{\Delta t}{\alpha^2}(\beta_{el} - \lambda_{el}) + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} f^n + \text{ some function of } \varepsilon.$$

Iteratively, it yields

$$f^n = \left(\frac{1 + \frac{\Delta t}{\alpha^2}(\beta_{el} - \lambda_{el}) + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} \right)^n f^0 + \text{ some function of } \varepsilon.$$

So when α is small, for *arbitrary initial data* f^0 and any integer m, there exists some integer N, s.t.

$$f^n \leq O(\alpha^m) + \text{ some function of } \varepsilon, \quad \text{for } n > N, \tag{3.6}$$

which means that f^n can be arbitrarily close to the null space of $\mathcal{Q}_{el}$:

$$\mathcal{Q}_{el}(f^n) \leq O(\alpha^m), \quad \text{for } n > N. \tag{3.7}$$

At this stage, if we examine the distance between f and M_q, we found that

$$f^{n+1} - M_q = \frac{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}}(f^n - M_q) + \frac{\frac{\Delta t}{\alpha^2}}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} \mathcal{Q}_{el}(f^n), \tag{3.8}$$

so

$$\left| f^{n+1} - M_q \right| \leq r \left| f^n - M_q \right| + O(\alpha^m), \quad \text{for } n > N,$$

where

$$r = \left| \frac{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} \right|.$$

Then for proper β_{ee}, we have $0 < r < 1$ and

$$\left| f^{n+n_1} - M_q \right| \leq r^{n_1} \left| f^n - M_q \right| + O(\alpha^m), \quad \text{for } n > N. \tag{3.9}$$

This implies that no matter what the initial condition is, f will eventually be driven to the desired Fermi–Dirac distribution, but the convergence rate can be rather slow for small α. What even worse is, when $\alpha \to 0$, $r \to 1$, we see from (3.8), (3.6) that f will stay around some function of ε, but nothing guarantees it is M! This violates the property 2 mentioned in the Introduction.

On the other hand, we know from (3.7) that $\mathcal{Q}_{el}(f^n)$ can be arbitrarily small after some time. What if we just set it equal to zero afterwards? Dropping this term as well as its penalization leads to

$$\frac{f^{n+1} - f^n}{\Delta t} = \frac{(1 - \beta_{ee})(M_q - f^n)}{\alpha} + \frac{\beta_{ee}(M_q - f^{n+1})}{\alpha},$$

which is

$$f^{n+1} - M_q = \frac{1 + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha}\beta_{ee}}(f^n - M_q). \tag{3.10}$$

Hence as long as $\beta_{ee} > 1/2$, we have $\left|\frac{1 + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha}\beta_{ee}}\right| < 1$, and f^n will converge to M_q regardless of the initial condition. Compared with (3.8), this one has much faster convergence rate.

3.1.2. The thresholded AP scheme

The above simple analysis illustrates that the penalization idea [15] should be applied wisely, especially when there are stiff terms on different scales. Back to the original equation (3.2), we propose to solve it as follows.

At time step t^{n+1}, check the norm of $\mathcal{Q}_{el}(f^n)$ in k:

- if $\|\mathcal{Q}_{el}(f^n)\| > \delta$, apply scheme (3.5);
- otherwise, apply (3.5) with $\mathcal{Q}_{el}(f^n) = \beta_{el} = 0$.

As explained previously, the threshold can be chosen based on the property we expect in (3.7), say, $\delta = \alpha^m$, some $m > 2$. However, in practice, as any numerical solver of $\mathcal{Q}_{el}$ has certain accuracy, we therefore set

$$\delta = O(\Delta k^l), \tag{3.11}$$

where l denotes the order of accuracy of the numerical solver for $\mathcal{Q}_{el}$.

Remark 8. The choice of threshold (3.11) does not violate the consistency of our method, since when $\Delta k \to 0$, we have $\delta \to 0$, and we are back to the original scheme (3.5) whose consistency to (3.2) is not a problem (i.e. the scheme satisfies the property 1 in the Introduction).

3.2. The spatially inhomogeneous case

We now include the spatial dependence to treat the full problem (2.3). To handle the newly added stiff convection terms, we follow the idea of [23] to form it into a set of parity equations. Denote $f^+ = f(x, k, t)$, $f^- = f(x, -k, t)$, then they solve

$$\partial_t f^+ + \frac{1}{\alpha}(\nabla_k \varepsilon \cdot \nabla_x f^+ + \nabla_x V \cdot \nabla_k f^+) = \frac{1}{\alpha^2}\mathcal{Q}_{el}(f^+) + \frac{1}{\alpha}\mathcal{Q}_{ee}(f^+),$$

$$\partial_t f^- - \frac{1}{\alpha}(\nabla_k \varepsilon \cdot \nabla_x f^- + \nabla_x V \cdot \nabla_k f^-) = \frac{1}{\alpha^2}\mathcal{Q}_{el}(f^-) + \frac{1}{\alpha}\mathcal{Q}_{ee}(f^-).$$

Now write

$$r(x, k, t) = \frac{1}{2}(f^+ + f^-), \qquad j(x, k, t) = \frac{1}{2\alpha}(f^+ - f^-),$$

we have

$$\partial_t r + \nabla_k \varepsilon \cdot \nabla_x j + \nabla_x V \cdot \nabla_k j = \frac{\mathcal{Q}_{el}(r)}{\alpha^2} + \frac{\mathcal{Q}_{ee}(f^+) + \mathcal{Q}_{ee}(f^-)}{2\alpha}, \tag{3.12}$$

$$\partial_t j + \frac{1}{\alpha^2}(\nabla_k \varepsilon \cdot \nabla_x r + \nabla_x V \cdot \nabla_k r) = -\frac{\lambda_{el}}{\alpha^2}j + \frac{\mathcal{Q}_{ee}(f^+) - \mathcal{Q}_{ee}(f^-)}{2\alpha^2}, \tag{3.13}$$

where we used the fact that j is an odd function in k, thus $\mathcal{Q}_{el}(j) = -\lambda_{el}j$.

For (3.12)–(3.13), the same penalization as in the homogeneous case suggests

$$\partial_t r + \nabla_k \varepsilon \cdot \nabla_x j + \nabla_x V \cdot \nabla_k j = \frac{\mathcal{Q}_{el}(r) - \beta_{el}(M_{el} - r)}{\alpha^2} + \frac{\beta_{el}(M_{el} - r)}{\alpha^2}$$
$$+ \frac{\mathcal{Q}_{ee}(f^+) + \mathcal{Q}_{ee}(f^-) - 2\beta_{ee}(M_q - r)}{2\alpha} + \frac{\beta_{ee}(M_q - r)}{\alpha},$$

$$\partial_t j + \frac{1}{\alpha^2}(\nabla_k \varepsilon \cdot \nabla_x r + \nabla_x V \cdot \nabla_k r) = -\frac{\lambda_{el}}{\alpha^2}j + \frac{\mathcal{Q}_{ee}(f^+) - \mathcal{Q}_{ee}(f^-) + 2\beta_{ee}\alpha j}{2\alpha^2} - \frac{\beta_{ee}}{\alpha}j.$$

Note here $M_q = M_q(x, \varepsilon, t)$ is the local Fermi–Dirac distribution, and $M_{el} = M_{el}(x, \varepsilon, t)$ is some function of ε whose form will be specified later. The coefficients β_{el} and β_{ee} are chosen the same as in (3.4), except that β_{ee} can also be made space dependent.

The above equations can be formed into a diffusive relaxation system [24]:

$$\partial_t r + \nabla_k \varepsilon \cdot \nabla_x j + \nabla_x V \cdot \nabla_k j = G_1(r, j) + \frac{\beta_{el}(M_{el} - r)}{\alpha^2} + \frac{\beta_{ee}(M_q - r)}{\alpha}, \tag{3.14}$$

$$\partial_t j + \theta(\nabla_k \varepsilon \cdot \nabla_x r + \nabla_x V \cdot \nabla_k r) = G_2(r, j) - \frac{\beta_{ee}}{\alpha} j - \frac{1}{\alpha^2}\left[\lambda_{el} j + (1 - \alpha^2 \theta)(\nabla_k \varepsilon \cdot \nabla_x r + \nabla_x V \cdot \nabla_k r)\right], \tag{3.15}$$

where

$$G_1(r, j) = \frac{\mathcal{Q}_{el}(r) - \beta_{el}(M_{el} - r)}{\alpha^2} + \frac{\mathcal{Q}_{ee}(f^+) + \mathcal{Q}_{ee}(f^-) - 2\beta_{ee}(M_q - r)}{2\alpha}, \tag{3.16}$$

$$G_2(r, j) = \frac{\mathcal{Q}_{ee}(f^+) - \mathcal{Q}_{ee}(f^-) + 2\beta_{ee}\alpha j}{2\alpha^2}, \tag{3.17}$$

and $0 \leq \theta(\alpha) \leq 1/\alpha^2$ is a control parameter simply chosen as $\theta(\alpha) = \min\{1, 1/\alpha^2\}$.

A first-order IMEX scheme for the system (3.14)–(3.15) thus reads

$$\frac{r^{n+1} - r^n}{\Delta t} + \nabla_k \varepsilon \cdot \nabla_x j^n + \nabla_x V^n \cdot \nabla_k j^n = G_1(r^n, j^n) + \frac{\beta_{el}(M_{el}^{n+1} - r^{n+1})}{\alpha^2} + \frac{\beta_{ee}(M_q^{n+1} - r^{n+1})}{\alpha}, \tag{3.18}$$

$$\frac{j^{n+1} - j^n}{\Delta t} + \theta(\nabla_k \varepsilon \cdot \nabla_x r^n + \nabla_x V^n \cdot \nabla_k r^n) = G_2(r^n, j^n) - \frac{\beta_{ee}}{\alpha} j^{n+1}$$
$$- \frac{1}{\alpha^2}\left[\lambda_{el} j^{n+1} + (1 - \alpha^2 \theta)(\nabla_k \varepsilon \cdot \nabla_x r^{n+1} + \nabla_x V^{n+1} \cdot \nabla_k r^{n+1})\right]. \tag{3.19}$$

3.2.1. Asymptotic properties of the numerical solution

Leaving aside other issues such as the spatial discretization, let us for the moment again assume that $\mathcal{Q}_{ee}(f) = M_q - f$, then (3.18)–(3.19) simplify to

$$\frac{r^{n+1} - r^n}{\Delta t} + \nabla_k \varepsilon \cdot \nabla_x j^n + \nabla_x V^n \cdot \nabla_k j^n = \frac{\mathcal{Q}_{el}(r^n) - \beta_{el}(M_{el}^n - r^n)}{\alpha^2} + \frac{\beta_{el}(M_{el}^{n+1} - r^{n+1})}{\alpha^2}$$
$$+ \frac{(1 - \beta_{ee})(M_q^n - r^n)}{\alpha} + \frac{\beta_{ee}(M_q^{n+1} - r^{n+1})}{\alpha},$$

$$\frac{j^{n+1} - j^n}{\Delta t} + \theta(\nabla_k \varepsilon \cdot \nabla_x r^n + \nabla_x V^n \cdot \nabla_k r^n) = \frac{\beta_{ee} - 1}{\alpha} j^n - \frac{\beta_{ee}}{\alpha} j^{n+1}$$
$$- \frac{1}{\alpha^2}\left[\lambda_{el} j^{n+1} + (1 - \alpha^2 \theta)(\nabla_k \varepsilon \cdot \nabla_x r^{n+1} + \nabla_x V^{n+1} \cdot \nabla_k r^{n+1})\right].$$

The scheme for j is actually

$$j^{n+1} = \frac{1 + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha^2}\lambda_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} j^n - \frac{\frac{\Delta t}{\alpha^2}}{1 + \frac{\Delta t}{\alpha^2}\lambda_{el} + \frac{\Delta t}{\alpha}\beta_{ee}}\left(\nabla_k \varepsilon \cdot \nabla_x r^{n+1} + \nabla_x V^{n+1} \cdot \nabla_k r^{n+1}\right)$$
$$+ \frac{\theta \Delta t}{1 + \frac{\Delta t}{\alpha^2}\lambda_{el} + \frac{\Delta t}{\alpha}\beta_{ee}}\left[(\nabla_k \varepsilon \cdot \nabla_x r^{n+1} + \nabla_x V^{n+1} \cdot \nabla_k r^{n+1}) - (\nabla_k \varepsilon \cdot \nabla_x r^n + \nabla_x V^n \cdot \nabla_k r^n)\right].$$

When $\alpha \ll 1$ (so $\theta = 1$), suppose all functions are smooth, we have

$$j^{n+1} = -\frac{1}{\lambda_{el}}\left(\nabla_k \varepsilon \cdot \nabla_x r^{n+1} + \nabla_x V^{n+1} \cdot \nabla_k r^{n+1}\right) + O(\alpha).$$

The scheme for r results in

$$r^{n+1} = \frac{1 + \frac{\Delta t}{\alpha^2}(\beta_{el} - \lambda_{el}) + \frac{\Delta t}{\alpha}(\beta_{ee} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} r^n - \frac{\Delta t}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}}\left(\nabla_k \varepsilon \cdot \nabla_x j^n + \nabla_x V^n \cdot \nabla_k j^n\right)$$
$$+ \frac{\frac{\Delta t}{\alpha^2}}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}} \mathcal{Q}_{el}^+(r^n) + \frac{\frac{\Delta t}{\alpha}\beta_{ee}}{1 + \frac{\Delta t}{\alpha^2}\beta_{el} + \frac{\Delta t}{\alpha}\beta_{ee}}\left(M_q^{n+1} - M_q^n\right)$$

$$+ \frac{\frac{\Delta t}{\alpha}}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} M_{\text{q}}^n + \frac{\frac{\Delta t}{\alpha^2}\beta_{\text{el}}}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left(M_{\text{el}}^{n+1} - M_{\text{el}}^n \right)$$

$$= \frac{1 + \frac{\Delta t}{\alpha^2}(\beta_{\text{el}} - \lambda_{\text{el}}) + \frac{\Delta t}{\alpha}(\beta_{\text{ee}} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} r^n + O\left(\alpha^2\right) + \text{ some function of } \varepsilon.$$

Iteratively, this gives

$$r^n = \left(\frac{1 + \frac{\Delta t}{\alpha^2}(\beta_{\text{el}} - \lambda_{\text{el}}) + \frac{\Delta t}{\alpha}(\beta_{\text{ee}} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \right)^n r^0 + O\left(\alpha^2\right) + \text{ some function of } \varepsilon.$$

Clearly the first term on the right hand side will be damped down as time goes by *no matter what the initial condition is.* After several steps, we have

$$r^n = O\left(\alpha^2\right) + \text{ some function of } \varepsilon, \quad \text{for } n > N,$$

which implies

$$\mathcal{Q}_{\text{el}}\left(r^n\right) = O\left(\alpha^2\right), \quad \text{for } n > N.$$

At this stage, if we continue to penalize $\mathcal{Q}_{\text{el}}(r^n)$, similarly as before,

$$r^{n+1} - M_{\text{q}}^{n+1} = \frac{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}(\beta_{\text{ee}} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left(r^n - M_{\text{q}}^n \right) + \frac{\frac{\Delta t}{\alpha^2}}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \mathcal{Q}_{\text{el}}\left(r^n\right)$$

$$- \frac{\Delta t}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left[\left(\nabla_k \varepsilon \cdot \nabla_x j^n + \nabla_x V^n \cdot \nabla_k j^n \right) + \frac{M_{\text{q}}^{n+1} - M_{\text{q}}^n}{\Delta t} \right]$$

$$- \frac{\frac{\Delta t^2}{\alpha^2}\beta_{\text{el}}}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left(\frac{M_{\text{q}}^{n+1} - M_{\text{q}}^n}{\Delta t} - \frac{M_{\text{el}}^{n+1} - M_{\text{el}}^n}{\Delta t} \right)$$

$$= \frac{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}(\beta_{\text{ee}} - 1)}{1 + \frac{\Delta t}{\alpha^2}\beta_{\text{el}} + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left(r^n - M_{\text{q}}^n \right) + O\left(\alpha^2 + \Delta t\right), \quad \text{for } n > N. \tag{3.20}$$

From (3.20) we see that r will converge to M_{q} with a dominant $O(\Delta t)$ error. Therefore, a good choice of M_{el} is

$$M_{\text{el}} = M_{\text{q}},$$

which will not only simplify the scheme, but also improve the asymptotic error from $O(\Delta t)$ to $O(\alpha^2)$.

Moreover, like the spatially homogeneous case, we suffer from the same problem that the convergence rate is too slow for small but finite α. To accelerate the convergence, we again choose some threshold δ as in (3.11) such that once $\|\mathcal{Q}_{\text{el}}(r^n)\|(x) < \delta$ (check it for every x), we set $\mathcal{Q}_{\text{el}}(r^n) = 0$ and turn off its penalization. Then (3.20) becomes

$$r^{n+1} - M_{\text{q}}^{n+1} = \frac{1 + \frac{\Delta t}{\alpha}(\beta_{\text{ee}} - 1)}{1 + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left(r^n - M_{\text{q}}^n \right) - \frac{\Delta t}{1 + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left[\left(\nabla_k \varepsilon \cdot \nabla_x j^n + \nabla_x V^n \cdot \nabla_k j^n \right) + \frac{M_{\text{q}}^{n+1} - M_{\text{q}}^n}{\Delta t} \right]$$

$$= \frac{1 + \frac{\Delta t}{\alpha}(\beta_{\text{ee}} - 1)}{1 + \frac{\Delta t}{\alpha}\beta_{\text{ee}}} \left(r^n - M_{\text{q}}^n \right) + O\left(\alpha\right),$$

and we gain much faster convergence rate.

3.2.2. The thresholded semi-discrete AP scheme

Based on the discussion above, we integrate the thresholding idea into (3.18)–(3.19) to propose the following semi-discrete AP scheme.

At time step t^{n+1}, given $f^n = (f^+)^n$, $(f^-)^n$, r^n, j^n, ρ^n, $\mathcal{E}^n$, and V^n:

- *Step 1*: Compute M_{q}^{n+1} used in (3.18) and V^{n+1} in (3.19).
  Although (3.18) appears implicit (recall $M_{\text{el}} = M_{\text{q}}$), M_{q}^{n+1} can be computed explicitly similarly as in [15]. Specifically, we multiply both sides of (3.18) by $(1, \varepsilon(k))^T$ and integrate w.r.t. k. Utilizing the conservation properties of $\mathcal{Q}_{\text{el}}$, $\mathcal{Q}_{\text{ee}}$, and the BGK operator, we get

$$\int_{\mathbb{R}^d} r^{n+1} \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \, dk = \int_{\mathbb{R}^d} r^n \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \, dk - \Delta t \int_{\mathbb{R}^d} \left(\nabla_k \varepsilon \cdot \nabla_x j^n + \nabla_x V^n \cdot \nabla_k j^n \right) \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \, dk.$$

Note that by definition

$$\int_{\mathbb{R}^d} r \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \mathrm{d}k = \int_{\mathbb{R}^d} f \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \mathrm{d}k = \begin{pmatrix} \rho \\ \rho \mathcal{E} \end{pmatrix},$$

so the preceding scheme just gives an evolution of the macroscopic variables

$$\begin{pmatrix} \rho^{n+1} \\ \rho^{n+1} \mathcal{E}^{n+1} \end{pmatrix} = \begin{pmatrix} \rho^n \\ \rho^n \mathcal{E}^n \end{pmatrix} - \Delta t \int_{\mathbb{R}^d} \left(\nabla_k \varepsilon \cdot \nabla_x j^n + \nabla_x V^n \cdot \nabla_k j^n \right) \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \mathrm{d}k.$$

Once ρ^{n+1} and $\mathcal{E}^{n+1}$ are computed, we can invert the system (2.22) to get z^{n+1} and T^{n+1} (details see [18]). Plugging them into (2.9) then defines M_q^{n+1}. Given ρ^{n+1}, V^{n+1} can be easily obtained by solving the Poisson equation (2.7).

- *Step 2*: Compute r^{n+1}.
  At every spatial point x, check the norm of $\mathcal{Q}_{el}(r^n)$ in k:
  – if $\|\mathcal{Q}_{el}(r^n)\|(x) > \delta$, apply scheme (3.18);
  – otherwise, apply (3.18) with $\mathcal{Q}_{el}(r^n)(x) = \beta_{el} = 0$.
- *Step 3*: Employ scheme (3.19) to get j^{n+1}.
- *Step 4*: Reconstruct $f^{n+1} = (f^+)^{n+1} = r^{n+1} + \alpha j^{n+1}$ and $(f^-)^{n+1} = r^{n+1} - \alpha j^{n+1}$.

3.2.3. Space discretization

We finally include the spatial discretization to the previous semi-discrete scheme to construct a fully-discrete scheme. We will show that when $\alpha \to 0$, it automatically becomes a discretization for the limiting ET model (2.10), thus is Asymptotic Preserving (satisfies the property 2 in the Introduction).

For the sake of brevity, we will present the method in a splitting framework, namely, separating the explicit and implicit parts in (3.18)–(3.19). This is equivalent to an unsplit version since our scheme is of an IMEX type. We also assume a slab geometry: $x \in \Omega \subset \mathbb{R}^1$. Extension to higher dimensions is straightforward.

Let $r_{l,m}^n$, $j_{l,m}^n$ denote the numerical approximation of $r(x_l, k_m, t^n)$ and $j(x_l, k_m, t^n)$, where $0 < l \le N_x$, $0 < m = (m_1, ..., m_d) \le N_k^d$, N_x and N_k are the number of points in x and k directions respectively. We have at time step t^{n+1}:

- *Step 1*: Solve the explicit part of (3.18)–(3.19) to get $r_{l,m}^*$ and $j_{l,m}^*$. The thresholding idea is embedded in this step when computing $G_1(r_{l,m}^n, j_{l,m}^n)$: if $\|\mathcal{Q}_{el}(r^n)\| < \delta$ at x_l, set $\mathcal{Q}_{el}(r^n)(x_l) = 0$. For convection terms, we use the upwind scheme with a slope limiter [31] on the Riemann invariants.
- *Step 2*: Solve for the macroscopic quantities ρ_l^* and $\mathcal{E}_l^*$ and thus define $(M_q)_{l,m}^*$ and $V_{l,m}^*$.

$$\rho_l^* = \sum_m r_{l,m}^* \Delta k^d, \qquad \mathcal{E}_l^* = \frac{1}{2} \sum_m r_{l,m}^* |k_m|^2 \Delta k^d / \rho_l^*.$$

Finding $(M_q)_{l,m}^*$ is then exactly the same as in the semi-discrete scheme. $V_{l,m}^*$ is solved from a simple finite-difference discretization of the Poisson equation.

- *Step 3*: Solve the implicit part of (3.18)–(3.19) to get $r_{l,m}^{n+1}$ and $j_{l,m}^{n+1}$ (if the threshold is satisfied in *Step 1*, set $\beta_{el} = 0$ in r's equation as well):

$$\frac{r_{l,m}^{n+1} - r_{l,m}^*}{\Delta t} = \frac{\beta_{el}[(M_q)_{l,m}^{n+1} - r_{l,m}^{n+1}]}{\alpha^2} + \frac{\beta_{ee}[(M_q)_{l,m}^{n+1} - r_{l,m}^{n+1}]}{\alpha}, \tag{3.21}$$

$$\frac{j_{l,m}^{n+1} - j_{l,m}^*}{\Delta t} = -\frac{1}{\alpha^2} \left[\lambda_{el} j_{l,m}^{n+1} + \left(1 - \alpha^2 \theta \right) \left(k_{m_1} \frac{r_{l+1,m}^{n+1} - r_{l-1,m}^{n+1}}{2\Delta x} + (\partial_x V)_l^{n+1} \frac{r_{l,m+1}^{n+1} - r_{l,m-1}^{n+1}}{2\Delta k} \right) \right] - \frac{\beta_{ee}}{\alpha} j_{l,m}^{n+1}. \tag{3.22}$$

First, it is easy to see that macroscopic quantities ρ_l^* and $\mathcal{E}_l^*$ remain unchanged during this step (the right hand side of (3.21) is conservative). Therefore, the previously obtained $(M_q)_{l,m}^*$ and $V_{l,m}^*$ are in fact $(M_q)_{l,m}^{n+1}$ and $V_{l,m}^{n+1}$. From (3.21) one can easily obtain $r_{l,m}^{n+1}$, and then (3.22) directly gives rise to $j_{l,m}^{n+1}$.

3.2.4. Asymptotic properties of the fully discrete scheme

As already shown in Section 3.2.1, sending α to zero in (3.18)–(3.19) for $\mathcal{Q}_{ee} = M_q - f$ leads to

$$r^{n+1} = M_q^{n+1},$$

$$j^{n+1} = -\frac{1}{\lambda_{el}} \left(\nabla_k \varepsilon \cdot \nabla_x r^{n+1} + \nabla_x V^{n+1} \cdot \nabla_k r^{n+1} \right),$$

which in 1-D fully discrete form read

$$r_{l,m}^{n+1} = (M_q)_{l,m}^{n+1},$$

$$j_{l,m}^{n+1} = -\frac{1}{\lambda_{\mathrm{el}}}\left(k_{m_1}\frac{r_{l+1,m}^{n+1} - r_{l-1,m}^{n+1}}{2\Delta x} + (\partial_x V)_l^{n+1}\frac{r_{l,m+1}^{n+1} - r_{l,m-1}^{n+1}}{2\Delta k}\right).$$

Plugging these relations into the discrete scheme of r resulted from the last subsection, we get (after multiplication by $(1, \varepsilon(k))^T$ and integration w.r.t. k):

$$\begin{aligned}
&\frac{1}{\Delta t}\begin{pmatrix}\rho^{n+1} - \rho^n \\ \rho^{n+1}\mathcal{E}^{n+1} - \rho^n\mathcal{E}^n\end{pmatrix} - \int_{\mathbb{R}^d}\frac{1}{\lambda_{\mathrm{el}}(\varepsilon)}\Bigg\{\frac{k_{m_1}}{2\Delta x}\bigg[k_{m_1}\frac{(M_q)_{l+2,m}^n - 2(M_q)_{l,m}^n + (M_q)_{l-2,m}^n}{2\Delta x} \\
&+ (\partial_x V)_{l+1}^n\frac{(M_q)_{l+1,m+1}^n - (M_q)_{l+1,m-1}^n}{2\Delta k} - (\partial_x V)_{l-1}^n\frac{(M_q)_{l-1,m+1}^n - (M_q)_{l-1,m-1}^n}{2\Delta k}\bigg] \\
&- \frac{|k_{m_1}|}{2\Delta x}\big((M_q)_{l+1,m}^n - 2(M_q)_{l,m}^n + (M_q)_{l-1,m}^n\big) + \frac{(\partial_x V)_l^n}{2\Delta k}\bigg[(\partial_x V)_l^n\frac{(M_q)_{l,m+2}^n - 2(M_q)_{l,m}^n + (M_q)_{l,m-2}^n}{2\Delta k} \\
&+ k_{m_1+1}\frac{(M_q)_{l+1,m+1}^n - (M_q)_{l-1,m+1}^n}{2\Delta x} - k_{m_1-1}\frac{(M_q)_{l+1,m-1}^n - (M_q)_{l-1,m-1}^n}{2\Delta x}\bigg] \\
&- \frac{|(\partial_x V)_l^n|}{2\Delta k}\big((M_q)_{l,m+1}^n - 2(M_q)_{l,m}^n + (M_q)_{l,m-1}^n\big)\Bigg\}\begin{pmatrix}1 \\ \varepsilon\end{pmatrix}\,dk = 0,
\end{aligned}\tag{3.23}$$

which is a *kinetic* scheme [12] for the ET model (2.10) (compare with (2.21) and (2.18)). Here for notation simplicity, we only consider the upwind scheme, the slope limiter can be added in the same manner.

4. Numerical examples

In this section, we present several numerical examples using our AP schemes. In order to perform the tests, we need accurate numerical solvers for the collision operators. Since the electron–electron interaction $\mathcal{Q}_{\mathrm{ee}}$ falls into a special case of the quantum Boltzmann operator, we adopt the fast spectral method developed in [20]. For the elastic collision $\mathcal{Q}_{\mathrm{el}}$, it is desirable to compute it in the same spectral framework but the direct evaluation would be very expensive. Here we introduce a fast approach by exploring the low-rank structure in the coefficient matrix. Details are gathered in Appendix A.

In the following, we always assume $k \in [-L_k, L_k]^2$ in 2-D and $x \in [0, L_x]$ in 1-D. N_k is the number of points in each k direction and N_x is the number of spatial discretization. For simplicity, we assume the scattering matrices $\Phi_{\mathrm{el}}(k, k') = \Phi_{\mathrm{ee}}(k, k_1, k', k_1') \equiv 1$. Then according to (3.4), β_{el} and β_{ee} can be chosen as $\beta_{\mathrm{el}} = 2\pi$ and $\beta_{\mathrm{ee}} = \pi\rho$ since $\mathcal{Q}_{\mathrm{ee}}^-(f) \le \frac{1}{2}\int_{\mathbb{R}^2}\int_{\mathbb{S}^1} ff_1 d\sigma\, dk_1 = \pi\rho f$ (see also (A.1)).

4.1. The spatially homogeneous case

We first check the behavior of the solution in the spatially homogeneous case. Consider the non-equilibrium initial data

$$f_0(k_1, k_2) = \frac{1}{\pi}\big[(k_1 - 1)^2 + (k_2 - 0.5)^2\big]e^{-[(k_1-1)^2 + (k_2-0.5)^2]}.$$

The parameters are chosen as $\alpha = 5e{-}3$ (diffusive regime), $\eta = 10$ (strong quantum effect; the equilibrium is very different from the classical Maxwellian), $L_k = 10.5$, $N_k = 64$. Under this condition, a stable explicit scheme would require $\Delta t = O(1e{-}6)$, while our scheme gives fairly good results with much coarser time step $\Delta t = 1$. Fig. 1 shows the AP error in L^∞-norm

$$\text{errorAP}_{L^\infty}^n = \max_{k_1,k_2}\big|f^n - M_q^n\big| \tag{4.1}$$

with time. Here we can see that during the initial period of time, this error decreases very slowly as explained in (3.6): f is only driven to some function of ε, not M_q, because of the dominating mechanism $\mathcal{Q}_{\mathrm{el}}$. Once the threshold comes into play as shown in (3.10), f will start to converge to M_q at a reasonable speed, which appears as a sharp transition in the asymptotic error. The magnitude of the threshold δ affects the position of this sharp transition in the way that bigger δ leads to earlier transition, as displayed in Fig. 1. As a comparison, the solid curve is obtained by the regular AP scheme without threshold, the error decreases very slowly as in (3.9). Fig. 2 displays the evolution of f at different times, where we clearly see that f first transits from non-radially symmetric to radially symmetric and then moves toward the desired Fermi–Dirac distribution M_q.

4.2. The spatially inhomogeneous case

In the rest of simulation, we always take $L_x = 1$, $L_k = 9.2$, and assume periodic boundary condition in x and zero boundary in k. The time step Δt is chosen to only satisfy the parabolic CFL condition: $\Delta t = O(\Delta x^2)$ (independent of α).
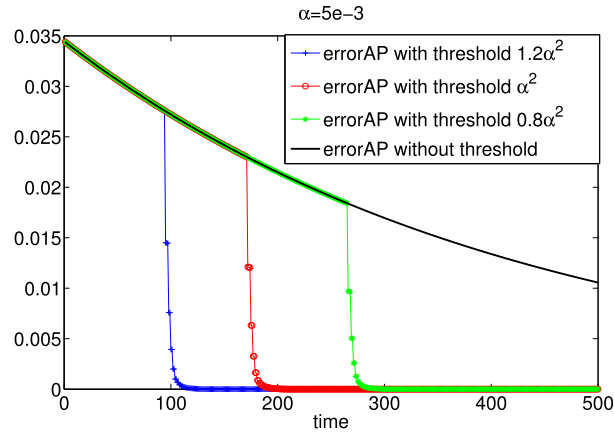
Fig. 1. Plots of the asymptotic error (4.1) versus time with different thresholds. Here $\eta = 10$, $\Delta t = 1$, $L_k = 10.5$, and $N_k = 64$.
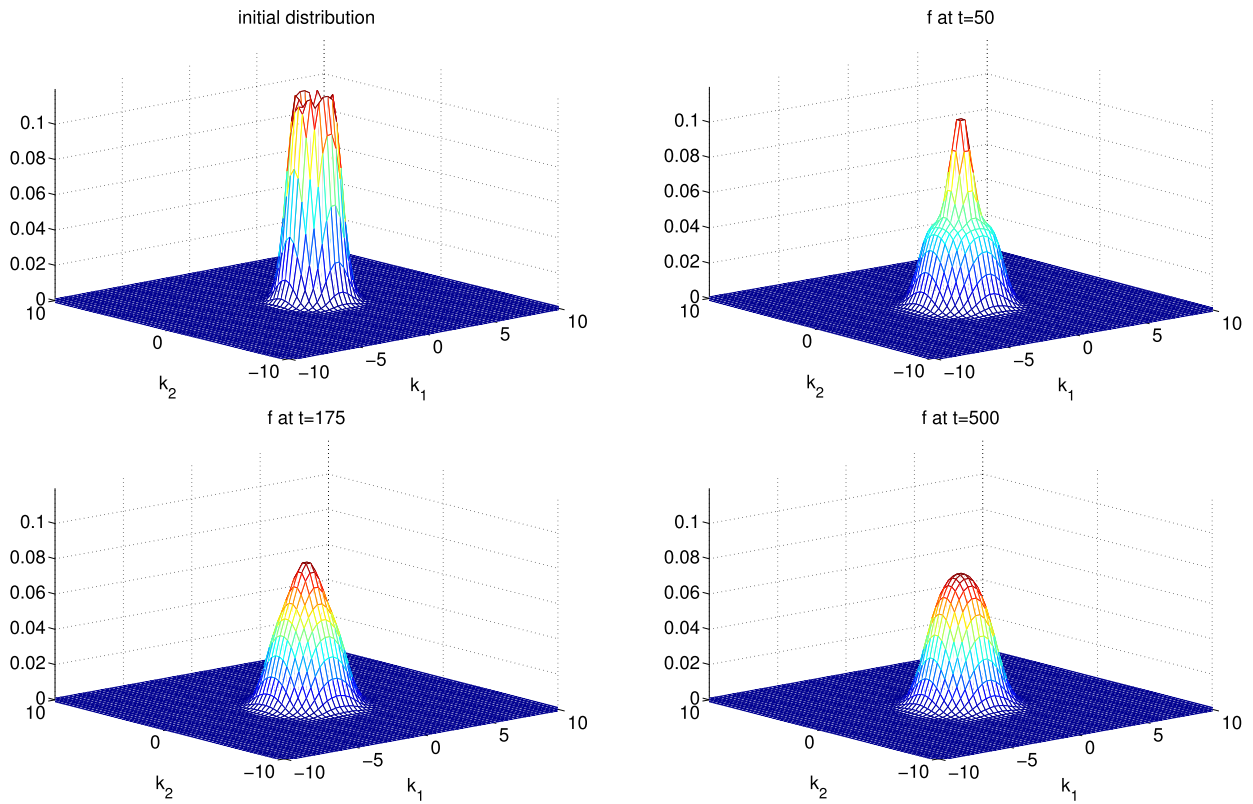

Fig. 2. Evolution of f at times $t = 0$, 50, 175 and 500. Here $\eta = 10$, $\Delta t = 1$, $L_k = 10.5$, and $N_k = 64$.

4.2.1. AP property

Consider Eq. (2.3) with non-equilibrium initial data

$$f_0(x, k_1, k_2) = \frac{1}{2\pi}\big(e^{-80(x - \frac{L_x}{2})^2} + 1\big)\big(e^{-[(k_1-1)^2 + k_2^2]} + e^{-[(k_1+1)^2 + k_2^2]}\big).$$

The electric field $\partial_x V$ is set to be one.

We check the asymptotic property by looking at the distance between r and M_q, i.e.,

$$\text{errorAP}_{L^1}^n = \sum_{x, k_1, k_2} \big|r^n - M_q^n\big| \Delta x \Delta k^2, \qquad \text{errorAP}_{L^\infty}^n = \max_{x, k_1, k_2} \big|r^n - M_q^n\big|. \tag{4.2}$$
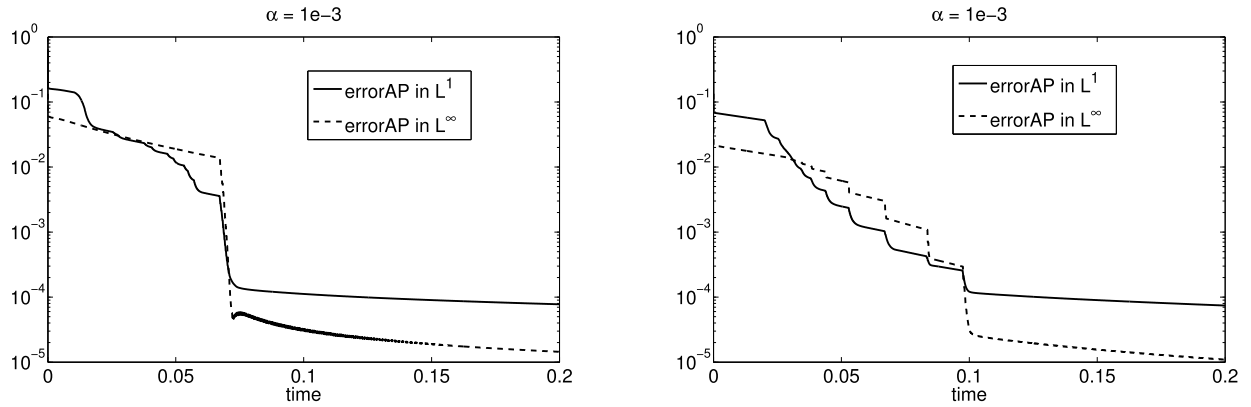
Fig. 3. Plots of the asymptotic error (4.2) versus time. Here $N_x = 40$, $\Delta t = 0.2 \Delta x^2$. Left: $\eta = 0.01$ (classical regime), $N_k = 32$. Right: $\eta = 3$ (quantum regime), $N_k = 64$.
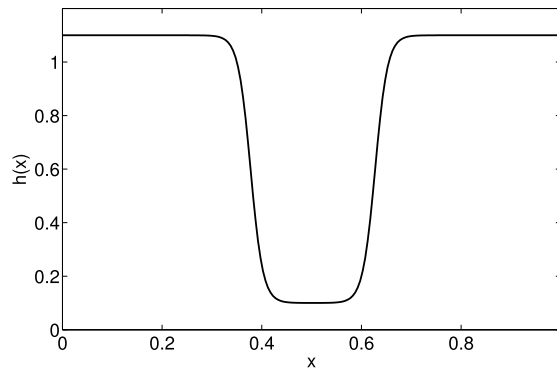

Fig. 4. Doping profile $h(x)$ for 1-D n^+–n–n^+ ballistic silicon diode.

The results are gathered in Fig. 3, where we observe a similar trend as in the space homogeneous case that r converges to M_q in two stages: first to a radially symmetric function (some function of ε) and then the local Maxwellian M_q. This in some sense mimics the Hilbert expansion in the derivation of ET model in Section 2.

4.2.2. 1-D n^+–n–n^+ ballistic silicon diode

We finally simulate a 1-D n^+–n–n^+ ballistic silicon diode, which is a simple model of the channel of a MOS transistor. The initial data is taken to be

$$f_0(x, k_1, k_2) = \left(1.1 + \frac{\tanh(40(x - \frac{5L_x}{8})) - \tanh(40(x - \frac{3L_x}{8}))}{2} \right) \times \left(e^{-[(k_1-1)^2 + k_2^2]} + e^{-[(k_1+1)^2 + k_2^2]} \right).$$

For Poisson equation (2.7), we choose $h(x) = \rho_0(x) = \int f_0 \, dk$, $\epsilon_r(x) \equiv 1$, $C_0 = 1/1000$, with boundary condition $V(0) = 0$, $V(L_x) = 1$. The doping profile $h(x)$ is shown in Fig. 4.

We consider two regimes: one is the kinetic regime with $\alpha = 1$, where we compare our solution with the one obtained by the explicit scheme (forward Euler); the other is the diffusive regime with $\alpha = 1e-3$, where our solution is compared with that of the ET system using the kinetic solver (3.23). Good agreements are obtained in Figs. 5, 6. Here the macroscopic quantities plotted are mass density ρ, energy $\mathcal{E}$ defined in (2.11), electron temperature T and fugacity z obtained through (2.22), electric field $E = -\partial_x V$, and mean velocity u defined as $u = j_\rho / \rho$.

5. Conclusion

We constructed an asymptotic preserving scheme for a multiscale semiconductor Boltzmann equation (coupled with Poisson equation) that in the diffusive regime captures the energy-transport limit. Besides the stiff convection terms, the *two-scale* collision operators pose new difficulties since the simple BGK penalization is unable to capture the correct limit. The key ingredient in our scheme is to set a suitable *threshold* for the leading elastic collision such that once this threshold is crossed, the less stiff collision begins to dominate. In this way, the convergence of the numerical solution to the local equilibrium resembles the Hilbert expansion at the continuous level. We analyzed this asymptotic behavior using a simplified BGK model. Several numerical results confirmed the asymptotic-preserving property for any non-equilibrium initial
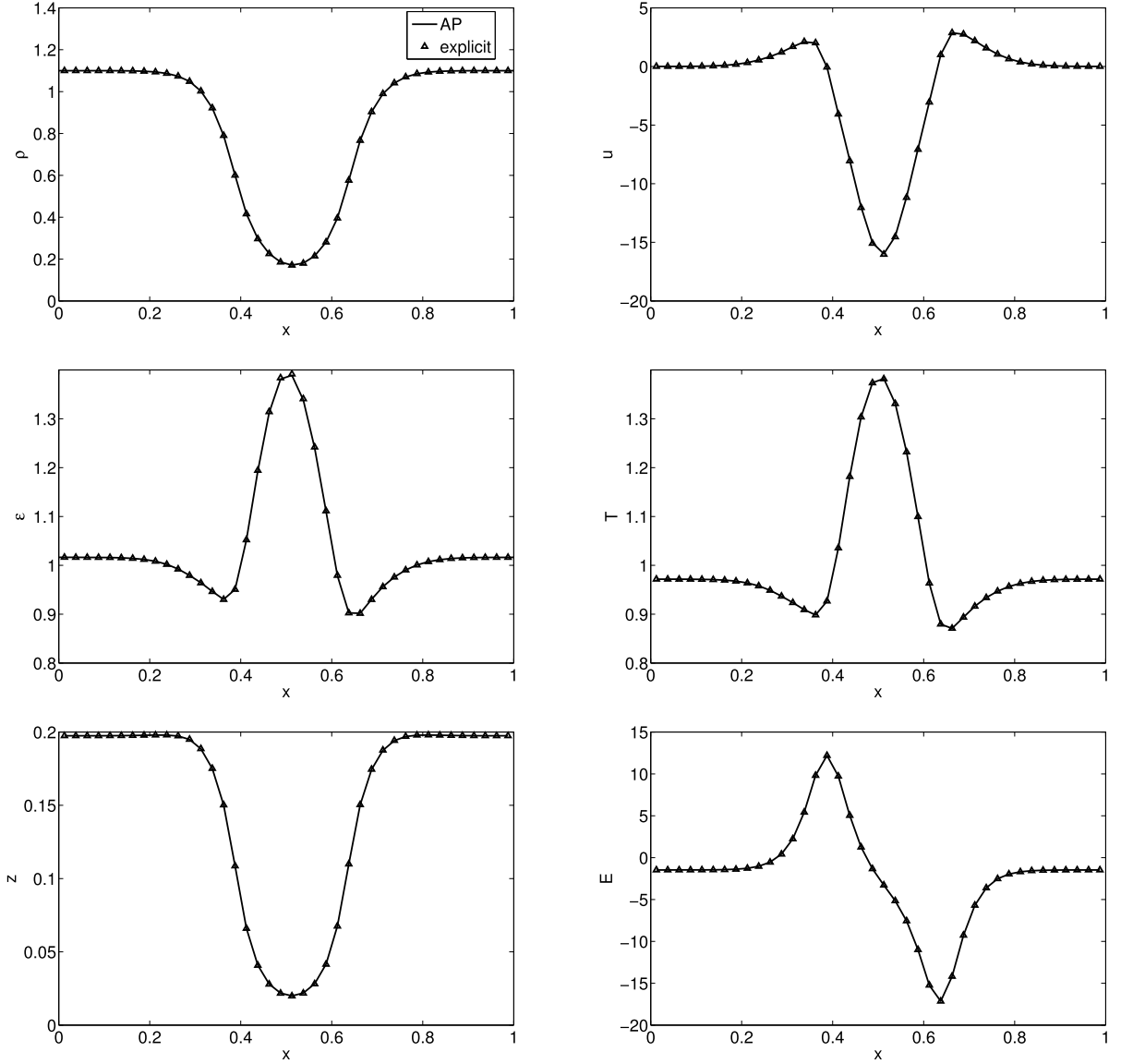
Fig. 5. Density ρ, velocity u, energy $\mathcal{E}$, temperature T, fugacity z, and electric field E at time $t = 0.05$. '——' is the AP scheme, '$\triangle$' is the forward Euler scheme. Here $\alpha = 1$, $\eta = 1$, $N_x = 40$, $N_k = 64$, $\Delta t = 0.2\Delta x^2$.

data, as well as the uniform stability of the scheme with respect to the mean free path, from kinetic regime to diffusive regime. As a variant of the scheme in this paper, we have devised another formulation using a splitting approach. The results are reported in a companion paper [19]. It is also interesting to include the inelastic electron–phonon interaction into the equation and we leave it to the future work.

Acknowledgements

Appendix A. Fast spectral methods for collision operators $\mathcal{Q}_{el}$ and $\mathcal{Q}_{ee}$

In this appendix, we briefly outline the numerical methods for computing the collision operators $\mathcal{Q}_{el}$ and $\mathcal{Q}_{ee}$. For numerical purpose, we assume the scattering matrices $\Phi_{el}(k, k') = \Phi_{ee}(k, k_1, k', k_1') \equiv 1$, and the wave vector $k \in \mathbb{R}^2$.
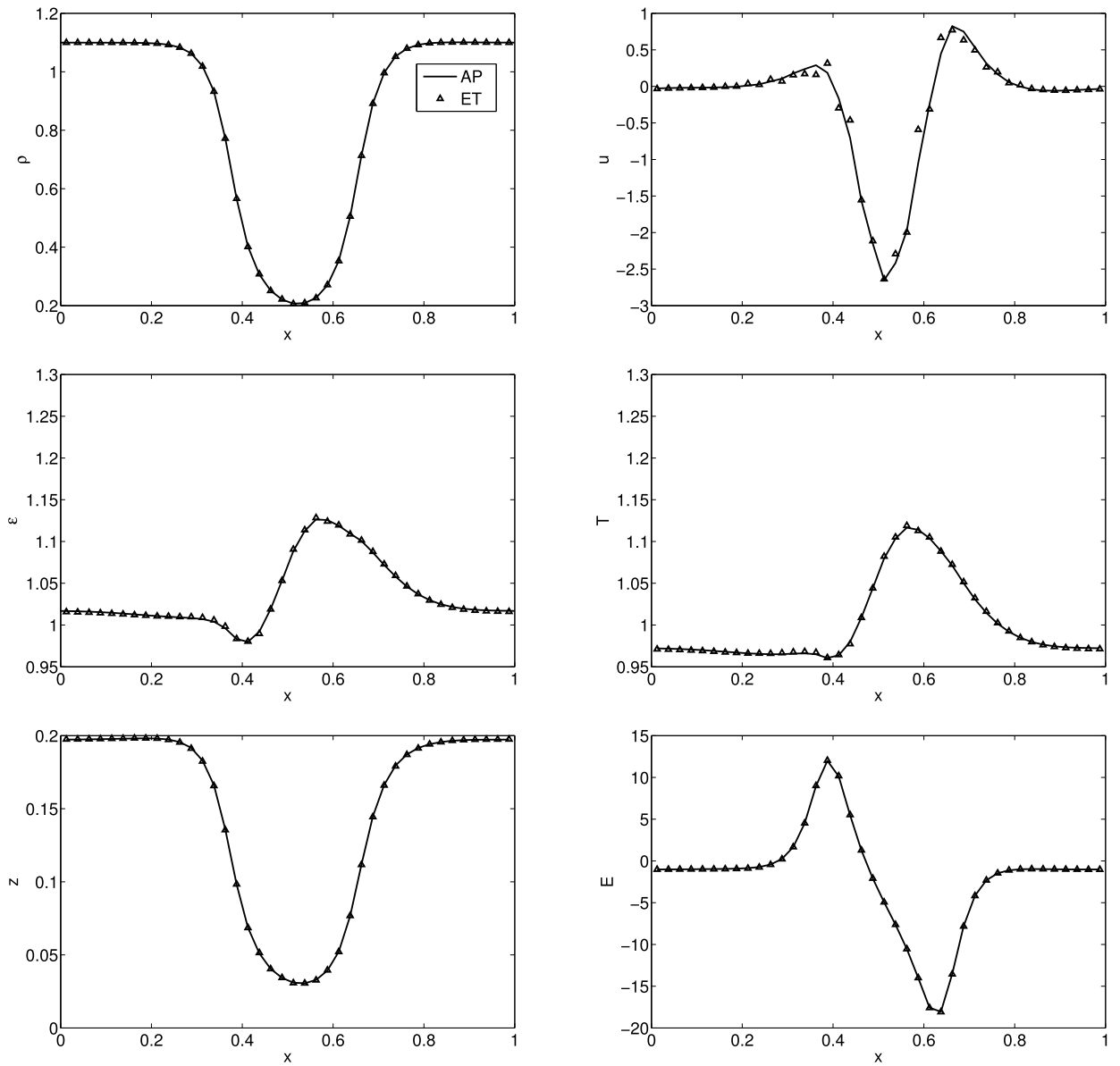
Fig. 6. Density ρ, velocity u, energy $\mathcal{E}$, temperature T, fugacity z, and electric field E at time $t = 0.05$. '——' is the AP scheme, '△' is the kinetic scheme (for the ET system). Here $\alpha = 1e{-}3$, $\eta = 1$, $N_x = 40$, $N_k = 64$, $\Delta t = 0.2\Delta x^2$.

A.1. Computation of $\mathcal{Q}_{ee}$

Under the parabolic band assumption, (2.6) reads

$$
\mathcal{Q}_{ee}(f)(k) = \int_{\mathbb{R}^6} \delta\big(k' + k'_1 - k - k_1\big)\delta\left(\frac{k'^2}{2} + \frac{k_1'^2}{2} - \frac{k^2}{2} - \frac{k_1^2}{2}\right)
$$

$$
\times \big[f'f'_1(1 - \eta f)(1 - \eta f_1) - ff_1\big(1 - \eta f'\big)\big(1 - \eta f'_1\big)\big]\,\mathrm{d}k_1 \mathrm{d}k' \mathrm{d}k'_1.
$$

By a change of variables, it is not difficult to rewrite the above integral in the center of mass reference system:

$$
\mathcal{Q}_{ee}(f)(k) = \frac{1}{2} \int_{\mathbb{R}^2} \int_{\mathbb{S}^1} \big[f'f'_1(1 - \eta f)(1 - \eta f_1) - ff_1\big(1 - \eta f'\big)\big(1 - \eta f'_1\big)\big]\,\mathrm{d}\sigma\,\mathrm{d}k_1, \tag{A.1}
$$

where

$$\begin{cases} k' = \dfrac{k+k_1}{2} + \dfrac{|k-k_1|}{2}\sigma, \\ k'_1 = \dfrac{k+k_1}{2} - \dfrac{|k-k_1|}{2}\sigma. \end{cases}$$

This is just the usual form of the quantum Boltzmann collision operator for a Fermi gas (of 2-D Maxwell molecules). Our way of computing $\mathcal{Q}_{ee}$ follows the fast spectral method in [20]. The starting point is to further transform (A.1) to a Carleman form

$$\mathcal{Q}_{ee}(f)(k) = \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} \delta(x\cdot y)\big[f' f'_1(1-\eta f)(1-\eta f_1) - f f_1\big(1-\eta f'\big)\big(1-\eta f'_1\big)\big]\,dx\,dy, \tag{A.2}$$

where $k_1 = k+x+y$, $k' = k+x$, and $k'_1 = k+y$. If $\text{Supp}(f(k)) \subset \mathcal{B}_S$ (a ball with radius S), we can truncate (A.2) as

$$\begin{aligned} \mathcal{Q}^R_{ee}(f)(k) &= \int\limits_{\mathcal{B}_R}\int\limits_{\mathcal{B}_R} \delta(x\cdot y)\big[f' f'_1(1-\eta f)(1-\eta f_1) - f f_1\big(1-\eta f'\big)\big(1-\eta f'_1\big)\big]\,dx\,dy \\ &= \int\limits_{\mathcal{B}_R}\int\limits_{\mathcal{B}_R} \delta(x\cdot y)\big[\big(f' f'_1 - f f_1\big) - \eta\big(f' f'_1 f_1 + f' f'_1 f - f' f_1 f - f_1 f'_1 f\big)\big]\,dx\,dy \end{aligned}$$

with $R = 2S$. We next choose a computational domain $\mathcal{D}_L = [-L, L]^2$ for k, and extend the function $f(k)$ periodically to the whole space $\mathbb{R}^2$. L is chosen such that $L \geq \frac{3\sqrt{2}+1}{2}S$ to avoid aliasing effect. We then approximate $f(k)$ by a truncated Fourier series

$$f(k) = \sum_{j=-N/2}^{N/2-1} \hat{f}_j e^{i\frac{\pi}{L} j\cdot k}, \quad \hat{f}_j = \frac{1}{(2L)^2}\int\limits_{\mathcal{D}_L} f(k) e^{-i\frac{\pi}{L} j\cdot k}\,dk. \tag{A.3}$$

Note here j is a 2D index rather than 1D. Inserting the Fourier expansion of f into $\mathcal{Q}^R_{ee}(f)$, and performing a spectral-Galerkin projection, we can get the governing equation for $\widehat{\mathcal{Q}^R_{ee}}(f)_j$. The computation is sped up by discovering a convolution structure and a separated expansion of the coefficient matrix. The final cost is roughly $O(MN^3 \log N)$, where N is the number of points in each k direction, and M is the number of angular discretization. More details can be found in [20].

A.2. Computation of $\mathcal{Q}_{el}$

Under the parabolic band assumption (2.4), (2.5) reads (see also (3.1))

$$\mathcal{Q}_{el}(f)(k) = \int\limits_{\mathbb{R}^2} \delta\big(\varepsilon' - \varepsilon\big)\big(f' - f\big)\,dk' = \int\limits_{\mathbb{S}^1} f\big(|k|\sigma\big)\,d\sigma - 2\pi f(k). \tag{A.4}$$

Compared to $\mathcal{Q}_{ee}$, this one is much easier to compute. For instance, one can do a direct numerical quadrature plus interpolation to approximate the integral over $\mathbb{S}^1$. To achieve better accuracy, we here present an efficient way to compute $\mathcal{Q}_{el}(f)$ based on the same spectral framework of $\mathcal{Q}_{ee}(f)$. Specifically, we still adopt the Fourier expansion (A.3). Inserting it into (A.4), we get

$$\mathcal{Q}_{el}(f)(k) = \sum_{j=-N/2}^{N/2-1} B(k, j)\hat{f}_j - 2\pi f(k), \tag{A.5}$$

where

$$B(k, j) = 2\pi J_0\left(\frac{\pi}{L}|k||j|\right).$$

Here L has to be $L \geq \frac{\sqrt{2}+2}{2}S$ to avoid aliasing. A direct computation of the above summation requires obviously $O(N^4)$ steps, which can be quite costly. But note that the coefficient matrix $B(k, j)$ only depends on the magnitude of k and j, which means that its rank is roughly $O(N)$. Therefore, we can find a low rank decomposition of $B(k, j)$ as (for 2-D problems this can be precomputed via a SVD)

$$B(k, j) = \sum_{r=1}^{O(N)} U_r(k) V_r(j).$$

Then computing the summation in (A.5) becomes

$$\sum_{j=-N/2}^{N/2-1} B(k,j)\hat{f}_j = \sum_{j=-N/2}^{N/2-1} \sum_{r=1}^{O(N)} U_r(k)V_r(j)\hat{f}_j = \sum_{r=1}^{O(N)} U_r(k)\left[\sum_{j=-N/2}^{N/2-1} V_r(j)\hat{f}_j\right].$$

The cost is reduced to $O(N^3)$.

To summarize, we have two fast spectral solvers for the collision operators $\mathcal{Q}_{el}$ and $\mathcal{Q}_{ee}$ (the cost of $\mathcal{Q}_{ee}$ is dominant due its intrinsic complexity). Both of them provide high-order accuracy and are thus suitable for testing asymptotic properties of our schemes.

References

[1] N.B. Abdallah, P. Degond, On a hierarchy of macroscopic models for semiconductors, J. Math. Phys. 37 (1996) 3306–3333.
[2] N.B. Abdallah, P. Degond, S. Genieys, An energy-transport model for semiconductors derived from the Boltzmann equation, J. Stat. Phys. 84 (1996) 205–231.
[3] N.B. Abdallah, L. Desvillettes, S. Genieys, On the convergence of the Boltzmann equation for semiconductors toward the energy transport model, J. Stat. Phys. 98 (2000) 835–870.
[4] J.F. Bourgat, P.L. Tallec, B. Perthame, Y. Qiu, Coupling Boltzmann and Euler equations without overlapping, in: Domain Decomposition Methods in Science and Engineering, in: Contemp. Math., vol. 157, AMS, 1994, pp. 377–398.
[5] J. Carrillo, I. Gamba, A. Majorana, C.-W. Shu, 2D semiconductor device simulations by WENO-Boltzmann schemes: efficiency, boundary conditions and comparison to Monte Carlo methods, J. Comput. Phys. 214 (2006) 55–80.
[6] Y. Cheng, I. Gamba, A. Majorana, C.-W. Shu, A discontinuous Galerkin solver for Boltzmann–Poisson systems in nano devices, Comput. Methods Appl. Mech. Eng. 198 (2009) 3130–3150.
[7] N. Crouseilles, M. Lemou, An asymptotic preserving scheme based on a micro–macro decomposition for collisional Vlasov equations: diffusion and high field scaling limits, Kinet. Relat. Models 4 (2011) 441–477.
[8] P. Degond, Mathematical Modelling of Microelectronics Semiconductor Devices, AMS/IP Stud. Adv. Math., vol. 15, AMS and International Press, 2000, pp. 77–110.
[9] P. Degond, A. Jüngel, P. Pietra, Numerical discretization of energy-transport models for semiconductors with non-parabolic band structure, SIAM J. Sci. Comput. 22 (2000) 986–1007.
[10] P. Degond, C.D. Levermore, C. Schmeiser, A note on the energy-transport limit of the semiconductor Boltzmann equation, in: N.B. Abdallah, et al. (Eds.), Transport in Transition Regimes, in: IMA Vol. Math. Appl., vol. 135, Springer, New York, 2004, pp. 137–153.
[11] J. Deng, Asymptotic preserving schemes for semiconductor Boltzmann equation in the diffusive regime, Numer. Math. Theor. Meth. Appl. 5 (2012) 278–296.
[12] S.M. Deshpande, Kinetic theory based new upwind methods for inviscid compressible flows, AIAA Paper 86-0275, 1986.
[13] G. Dimarco, L. Pareschi, V. Rispoli, Implicit–explicit Runge–Kutta schemes for the Boltzmann–Poisson system for semiconductors, Commun. Comput. Phys. 15 (2014) 1291–1319.
[14] H. Federer, Geometric Measure Theory, Springer, Berlin, 1969.
[15] F. Filbet, S. Jin, A class of asymptotic-preserving schemes for kinetic equations and related problems with stiff sources, J. Comput. Phys. 229 (2010) 7625–7648.
[16] A. Gnudi, D. Ventura, G. Baccarani, F. Odeh, Two-dimensional MOSFET simulation by means of a multidimensional spherical harmonic expansion of the Boltzmann transport equation, Solid-State Electron. 36 (1993) 575–581.
[17] F. Golse, F. Poupaud, Limite fluide des equations de Boltzmann des semiconducteurs pour une statistique de Fermi–Dirac, Asymptot. Anal. 6 (1992) 135–160.
[18] J. Hu, S. Jin, On kinetic flux vector splitting schemes for quantum Euler equations, Kinet. Relat. Models 4 (2011) 517–530.
[19] J. Hu, S. Jin, L. Wang, An asymptotic-preserving scheme for the semiconductor Boltzmann equation with two-scale collisions: a splitting approach, submitted for publication.
[20] J. Hu, L. Ying, A fast spectral algorithm for the quantum Boltzmann collision operator, Commun. Math. Sci. 10 (2012) 989–999.
[21] S. Jin, Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations, SIAM J. Sci. Comput. 21 (1999) 441–454.
[22] S. Jin, Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review, Riv. Mat. Univ. Parma 3 (2012) 177–216.
[23] S. Jin, L. Pareschi, Discretization of the multiscale semiconductor Boltzmann equation by diffusive relaxation schemes, J. Comput. Phys. 161 (2000) 312–330.
[24] S. Jin, L. Pareschi, G. Toscani, Uniformly accurate diffusive relaxation schemes for multiscale transport equations, SIAM J. Numer. Anal. 38 (2000) 913–936.
[25] S. Jin, L. Wang, Asymptotic-preserving numerical schemes for the semiconductor Boltzmann equation efficient in the high field regime, SIAM J. Sci. Comput. 35 (2013) B799–B819.
[26] W. Jones, N.H. March, Nonequilibrium and Disorder, Theoretical Solid-State Physics, vol. 2, Dover, New York, 1973.
[27] A. Jüngel, Transport Equations for Semiconductors, Lect. Notes Phys., vol. 773, Springer, Berlin, 2009.
[28] A. Jüngel, Energy transport in semiconductor devices, Math. Comput. Model. Dyn. Syst. 16 (2010) 1–22.
[29] A. Jüngel, P. Pietra, A discretization scheme of a quasi-hydrodynamic semiconductor model, Math. Models Methods Appl. Sci. 7 (1997) 935–955.
[30] A. Klar, H. Neunzert, J. Struckmeier, Transition from kinetic theory to macroscopic fluid equations: a problem for domain decomposition and a source for new algorithms, Transp. Theory Stat. Phys. 29 (2000) 93–106.
[31] R.J. LeVeque, Numerical Methods for Conservation Laws, second ed., Birkhäuser Verlag, Basel, 1992.
[32] P.A. Markowich, C. Ringhofer, C. Schmeiser, Semiconductor Equations, Springer-Verlag, Wien, New York, 1990.
[33] R.K. Pathria, P.D. Beale, Statistical Mechanics, third ed., Academic Press, 2011.
[34] F. Poupaud, Diffusion approximation of the linear semiconductor Boltzmann equation: analysis of boundary layers, Asymptot. Anal. 4 (1991) 293–317.
[35] C. Ringhofer, An entropy-based finite difference method for the energy transport system, Math. Models Methods Appl. Sci. 11 (2001) 769–796.
[36] W.V.V. Roosebroeck, Theory of flow of electrons and holes in germanium and other semiconductors, Bell Syst. Tech. J. 29 (1950) 560–607.